



Published in final edited form as:

Chem Res Toxicol. 2020 August 17; 33(8): 2087–2098. doi:10.1021/acs.chemrestox.0c00064.

Metabolomics Profiles of Smokers from Two Ethnic Groups with Differing Lung Cancer Risk

Romel Dator¹, Peter W. Villalta¹, Nicole Thomson¹, Joni Jensen⁴, Dorothy K. Hatsukami^{1,4}, Irina Stepanov^{1,3}, Benedikt Warth^{2,5,*}, Silvia Balbo^{1,3,*}

¹Masonic Cancer Center, University of Minnesota, Minneapolis, MN 55455

²Department of Food Chemistry and Toxicology, Faculty of Chemistry, University of Vienna, Währingerstraße 38, 1090 Vienna, Austria

³Division of Environmental Health Sciences, School of Public Health, University of Minnesota, Minneapolis, MN 55455, United States

⁴Department of Psychiatry, University of Minnesota, Minneapolis, MN 55455

⁵Scripps Center for Metabolomics, The Scripps Research Institute, La Jolla, CA 92037

Abstract

African American (AA) smokers are at a higher risk of developing lung cancer compared to Whites. The variations in the metabolism of nicotine and tobacco-derived carcinogens in these groups were reported previously with the levels of nicotine metabolites and carcinogen-derived metabolites measured using targeted approaches. While useful, these targeted strategies are not able to detect global metabolic changes for use in predicting the detrimental effects of tobacco use and ultimately lung cancer susceptibility among smokers. To address this limitation, we have performed global untargeted metabolomics profiling in urine of AA and White smokers to characterize the pattern of metabolites, identify differentially regulated pathways, and correlate these profiles with the observed variations in lung cancer risk between these two populations. Urine samples from AA ($n=30$) and White ($n=30$) smokers were used for metabolomics analysis acquired in both positive and negative electrospray ionization modes. LC-MS data were uploaded onto the cloud-based XCMS Online (<http://xcmsonline.scripps.edu>) platform for retention time correction, alignment, feature detection, annotation, statistical analysis, data visualization, and automated systems biology pathway analysis. The latter identified global differences in the metabolic pathways in the two groups including the metabolism of carbohydrates, amino acids, nucleotides, fatty acids, and nicotine. Significant differences in the nicotine degradation pathway (cotinine glucuronidation) in the two groups were observed and confirmed using a targeted LC-MS/MS approach. These results are consistent with previous studies demonstrating AA smokers

*Senior corresponding authors: Phone: (612) 624-4240, Fax: (612) 624-3869, balbo006@umn.edu; benedikt.warth@univie.ac.at.

Conflicts of Interest. The authors declare no potential conflicts of interest.

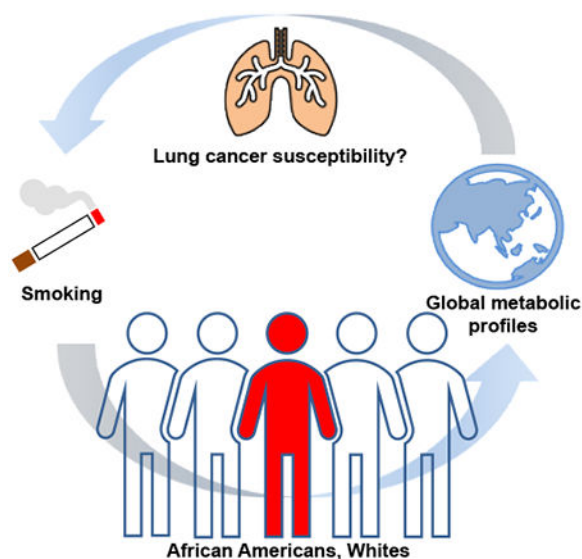
Supporting Information Available

Supporting information available includes PDF file (Figure S1–S6 and Tables S1–S4) with targeted LC-MS/MS data for NNAL and nicotine metabolites, TNE, creatinine, demographics of the 60 subjects, multi-modal pathway analysis results, and extracted ion chromatograms for the associated putative metabolites.

Final version of this manuscript is available online at <https://pubs.acs.org/doi/10.1021/acs.chemrestox.0c00064>

with lower glucuronidation capacity compared to Whites. Furthermore, the D-glucuronate degradation pathway was found to be significantly different between the two populations, with lower amounts of the putative metabolites detected in AA compared to Whites. We hypothesize that the differential regulation of the D-glucuronate degradation pathway is a consequence of the variations in the glucuronidation capacity observed in the two groups. Other pathways including the metabolism of amino acids, nucleic acids, and fatty acids were also identified, however, the biological relevance and implications of these differences across ethnic groups need further investigation. Overall, the applied metabolomics approach revealed global differences in the metabolic networks and endogenous metabolites in AA and Whites, which could be used and validated as new potential panel of biomarkers that could be used to predict lung cancer susceptibility among smokers in population-based studies.

TOC



Keywords

LC-MS; global metabolomics; XCMS Online; systems biology pathway analysis; ethnic differences; lung cancer susceptibility; tobacco smoking

INTRODUCTION

Tobacco smoking is the main cause of lung cancer-related mortalities worldwide. Despite the more than 90% lung cancer incidence associated with this lifestyle habit, only a fraction (11-24%) of smokers will develop lung cancer in their lifetime.¹⁻³ This disparity is hypothesized to be due to inter-individual genetics differences, which result in variations in the uptake and metabolism of nicotine and tobacco-derived carcinogens leading to differing levels of metabolites in biological fluids.⁴⁻⁶ Epigenetics, behavioral, and environmental factors, including diet and lifestyle may also contribute to the observed variations in cancer risk and other major causes of mortality across different ethnic groups.⁷⁻¹⁰ Specifically, AA smokers have been shown to be at a higher risk of developing tobacco-related lung cancer

compared to Whites.^{11, 12} Thus, the global analysis of biological networks and their associated metabolites in populations with differing lung cancer risk as in African Americans (AA) and Whites could identify new potential biomarkers to predict susceptibility to the detrimental effects of tobacco use among smokers. Understanding the differences in the overall metabolic regulation in smokers across different ethnic groups, not only those of nicotine and tobacco-specific carcinogen metabolism, is important in gaining insights into the impact of ethnic and genetic differences on lung cancer susceptibility, which could be used to develop rational strategies for cancer prevention based on targeted surveillance of high-risk and susceptible populations.

Numerous studies have shown differences in genetic backgrounds result into varying capacities to metabolize drugs, nicotine, and tobacco-specific carcinogens.^{9, 13–18} For instance, multi-ethnic studies demonstrated that AA smokers have a higher risk of developing lung cancer than White smokers due in part to the differing metabolic enzymes activities in these groups.^{9, 11–13, 15, 17–21} Inter-individual genetic differences can affect nicotine metabolism, which influence smoking behavior, toxicity, and detoxification capacity and thus ultimately impact tobacco-derived carcinogen exposure.⁹ For example, the gene variants of cytochrome P450 2A6 (CYP2A6) are associated with decreased risk of tobacco smoking-related lung cancer.^{9, 22} CYP2A6 is the main enzyme responsible for metabolizing nicotine. Smokers carrying genetic variants of this gene associated with slower nicotine conversion, are more likely to smoke less, have reduced exposure to tobacco smoke carcinogens, and thus lower risk of developing lung cancer.⁹ The reduced activity of CYP2A6 enzyme in Japanese Americans was observed to be associated with lower risk of developing smoking related-lung cancer in this population compared to other ethnic groups.^{8, 9, 13, 14, 23} In addition, the UDP—glucuronosyltransferases (UGT) are another class of enzymes implicated in ethnic differences in the metabolism and detoxification of tobacco-related compounds such as nicotine and tobacco-derived nitrosamines.^{15, 19, 24, 25} A low glucuronidation capacity in AA compared to Whites was observed in several multi-ethnic studies with the mean urinary cotinine glucuronidation ratio found to be 0.57 in AA over Whites.^{15, 18, 19, 24} AA have high prevalence of UGT2B10 splice variants resulting in lower cotinine glucuronidation.^{19, 26, 27} Similarly, the UGT2B10 splice variants commonly found in AA may increase their exposure to drugs during treatment.²⁶ The predominance of these gene variants including the UGTs and CYP2A6 can lead to differences in the metabolism of nicotine and tobacco-derived carcinogens, which could provide insights into the variations in risk and susceptibility to developing smoking-related cancers in these groups.

Mass spectrometry-based metabolomics have emerged as a powerful tool to investigate global dysregulation of biological networks resulting from specific exposures.²⁸ As more advanced bioinformatics and instrument platforms as well as spectral databases are being developed, its applicability to address various biologically relevant questions is rapidly expanding.^{29–39} For instance, multiple metabolomics platforms have been applied to probe global changes and patterns in altered metabolites between smokers and non-smokers.^{40–42} Metabolomics workflows have also been used in *in vitro* and *in vivo* models of tobacco-smoke-induced perturbations to identify tobacco-related biomarkers for lung cancer and other diseases.^{43–47} In addition, metabolomics analysis has been widely used in human studies to identify biomarkers of smoking habits and assessment of variability in the

metabolism of tobacco-derived compounds to understand their contributions in cancer development.^{40–45, 48, 49} Although relevant information has been deciphered from these studies, limited information is available on the global metabolic dysregulation in populations with differing genetic backgrounds, and in populations from different ethnicities with varying susceptibility to developing lung cancer due to tobacco smoking. To address this limitation, we have performed global untargeted metabolomics profiling in urine of AA and White smokers to characterize the pattern of metabolites, identify potentially dysregulated pathways, and correlate these profiles with the observed variations in lung cancer risk between these two populations. In order to validate the use of these global profiles, the results were compared to those obtained with well-established, traditional targeted MS-based analysis of nicotine-derived metabolites.

Experimental Procedures

Subjects.

Urine samples from AA ($n=30$) and White ($n=30$) smokers were obtained from the study “Ethnic differences in tobacco carcinogen metabolism” at the University of Minnesota. This study was approved by the University of Minnesota Institutional Review Board: Human Subjects Committee. The detailed 24-h urine sample collection and study design has been previously described.⁵⁰ The mean age of AA smokers was 46 ± 8 years, while 40 ± 12 years for White smokers. All subjects used in this study were males and were smokers with smoking frequency of more than 10 cigarettes per day (CPD). The TNE (Total Nicotine Equivalent), an established exposure biomarker of cigarette smoke exposure, and NNAL (4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol) levels (a biomarker of carcinogen exposure), were measured for the 60 subjects as previously described and reported in the Supporting Table 1.^{19, 51}

Caution.

NNAL and [¹³C₆]-NNAL are carcinogenic and must be handled with extreme care and proper personal protective equipment and ventilation.

Chemicals and Reagents.

NNAL, [¹³C₆]-NNAL, NNAL-O-glucuronide, NNAL-N-glucuronide, NNK-N-oxide and NNAL-N-oxide were obtained from Toronto Research Chemicals (Ontario, Canada). Oasis HLB (3 cc) and Oasis MCX (2 mg solid phase extraction 96-well plates) were purchased from Waters (Milford, MA). All acids and organic solvents were MS grade.

Sample preparation.

For the untargeted global metabolomics and targeted metabolite analysis, the same set of samples were used, that is the same aliquots of urine from each subject were processed and analyzed for both the analyses. Urine samples were centrifuged at 14,000 rpm and 4°C for 10 min to remove particulates and 500 μ L were used for untargeted metabolomics analysis (Fig. 1a). The samples were cleaned-up using Oasis HLB (3 cc) (Waters, Milford, MA). The cartridge was conditioned with 3 mL methanol and then with 3 mL water. The urine samples were loaded into the cartridge and washed with 3 mL water. After which, the metabolites

were eluted with 3 mL methanol and the fractions were dried in a refrigerated vacuum centrifuge ($T < 10^{\circ}\text{C}$). The fractions were re-suspended in acetonitrile/water (20/80, v/v) and the volume (30 μL) normalized based on creatinine concentration. The creatinine concentration (mg/mL) was measured using a colorimetric microplate assay (CRE34-K01) obtained from Eagle Bioscience (<http://stores.eaglebio.com/creatinine-microplate-assay-kit>). The total nicotine equivalents (TNE: nmol/mL, sum of total nicotine, total cotinine, total 3-hydroxycotinine and nicotine-*N*-oxide) were determined for these samples as previously described (Table S1).¹⁹ Sample normalization was performed based on the creatinine concentration of each urine sample and the TNE (total nicotine equivalent) to normalize the levels of specific nicotine metabolites detected in the metabolomics analysis to account for the differences in tobacco smoke exposure.

Untargeted LC-HRMS-Based Metabolomics Analysis.

LC-HRMS analyses were performed on Agilent 1200 series micro-flow HPLC (Agilent Technologies, Santa Clara, CA) coupled to a Bruker Impact II quadrupole time-of-flight (Q-TOF) mass spectrometer (Bruker Daltonics, Billerica, MA) in both positive and negative ionization mode. Reversed-phase chromatography was performed on a Waters Atlantis T3 column (3 μm , 1.0 x 150 mm) (Waters, Milford, MA) equipped with a VanGuard pre-column (2.1 x 5 mm). Separation was performed at room temperature and flow rate of 65 $\mu\text{L min}^{-1}$ using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in acetonitrile as mobile phase B. Four microliters of the sample was injected on-column. Gradient elution was carried out starting with 2% B for the first 5 min and a linear gradient from 2% to 40% B over 15 min and to 100% B for 7 min followed by a constant 100% B for 4 min. Finally, a linear gradient from 100% to 2% B over 2 min was performed and the column was re-equilibrated at 2% B for another 6 min. The total run time was 39 min. For the MS analysis, full scan data acquisition was performed with a mass range of m/z 50-1000 with a mass resolving power (FWHM) of 30,000. The Funnel 1 RF was set to 150 Vpp, Funnel 2 RF to 200 Vpp, Hexapole RF to 50 Vpp, quadrupole ion energy was 4.0 eV, and collision energy of 7.0 eV. For data-dependent MS/MS, the isolation width was ± 0.5 Da, 3 ions per full scan were subjected to MS/MS, with exclusion of ions from subsequent analysis for 1 min. Samples were measured in a randomized manner with pooled QC samples injected after every 6 samples. The QC sample was made by pooling 10 random urine samples from the 60 subjects. A mixture of authentic reference standards (NNAL, [$^{13}\text{C}_6$]-NNAL, NNAL-O-glucuronide, NNAL-N-glucuronide, NNK-N-oxide and NNAL-N-oxide) was utilized as additional QC measure. Internal calibration was performed by injecting sodium formate around 35 min within each run. After the samples were analyzed, they were un-blinded and classified into two groups corresponding to AA and Whites. LC-HRMS mass spectral data were uploaded to the cloud-based XCMS Online platform (<http://xcmsonline.scripps.edu>) for retention time correction, alignment, feature detection, annotation, and statistical analysis.⁵² Feature detection parameters include 5 ppm mass tolerance, minimum peak width of 10 s, maximum peak width of 60 s, obiwarp was used for retention time correction with profStep set to 1, mzwid of 0.015, minfrac of 0.5, bw of 5, allowable retention time deviation of 20 s, unpaired parametric t-test (Welch t-test) and post-hoc analysis for statistical analysis with p-value threshold of 0.001 for highly significant features, fold change greater than 1.5, adducts considered for the database search include $[\text{M} + \text{H}]^+$, $[\text{M} +$

$\text{Na}]^+$, $[\text{M} + \text{H} + \text{Na}]^{2+}$, $[\text{M} + \text{NH}_3 + \text{H}]^+$, $[\text{M} - \text{NH}_3 + \text{H}]^+$ and $[\text{M} + \text{H} - \text{H}_2\text{O}]^+$ for ESI (+) mass spectral data and $[\text{M} - \text{H}]^-$, $[\text{M} + \text{Na} - 2\text{H}]^-$ and $[\text{M} - \text{H}_2\text{O}]^-$ for ESI (-) mass spectral data. Isotopes were searched with m/z absolute error of 0.005 Da, and 5 ppm mass error and the sample bio-source was set to human. Systems biology pathway analysis was performed to identify differentially regulated pathways in the two groups of smokers as previously described.⁵⁴ Putative metabolites were searched against the METLIN and HMDB databases.^{36, 53} Multivariate principal component analysis (PCA) was performed in the XCMS Online platform to identify features that show maximum variation in the two groups.⁵⁵ Finally, autonomous multi-modal pathway analysis was performed with p-value set to 0.01 and 5 ppm mass tolerance on the XCMS Online processed datasets (positive and negative modes) as previously described.⁵⁶ Metabolomics data have been deposited to the EMBL-EBI MetaboLights database (DOI: 10.1093/nar/gkz1019, PMID:31691833) with the identifier MTBLS1705. The complete dataset can be accessed here (<https://www.ebi.ac.uk/metabolights/MTBLS1705>).⁷¹

Targeted LC-MS/MS quantitation of nicotine metabolites.

To determine the levels of nicotine metabolites in urine of the 60 subjects, we performed targeted LC-MS/MS analysis using selected reaction monitoring (SRM) of free and total cotinine and trans-3-hydroxycotinine as previously described.^{19, 57} Briefly, the diluted urine samples (1:10, 10 μL total) from the sixty subjects were combined with 400 μL of 100 mM ammonium acetate, pH 5.0, and methyl- d_3 internal standards (1 ng each). The mixture from each sample was added to paired 96-well plates, one for the analysis of free cotinine and trans-3-hydroxycotinine and the second for total (free + glucuronide) cotinine and trans-3-hydroxycotinine (Fig. 1b). One of the plates (Plate 2) was incubated overnight at 37°C with β -glucuronidase (500 units) (recombinant β -glucuronidase from over-expressing *E. coli* BL21). Samples were cleaned up using mixed-mode cation solid-phase extraction 96-well plates using Oasis MCX (2 mg solid phase extraction 96-well plates) (Waters, Milford, MA). LC-MS/MS analysis was performed on an Agilent 1100 capillary HPLC system coupled to a Thermo Scientific TSQ Vantage mass spectrometer in positive electrospray ionization mode (Thermo Fisher, San Jose, CA). The samples were re-suspended in 25 μL 100 mM ammonium acetate:methanol and 4 μL were injected on an Atlantis HILIC column (300 $\mu\text{m} \times 100 \text{ mm}$) (Waters, Milford, MA). Cotinine (t_R : 5.03 min) and trans-3-hydroxycotinine (t_R : 4.29 min) were eluted with acetonitrile:water:formic acid (95/3.5/1.5). The flow rate was 20 $\mu\text{L min}^{-1}$. SRM transitions used were m/z 177.1 \rightarrow 80.1 (confirmation) and m/z 177.1 \rightarrow 98.1 (quantitation) for d_0 -cotinine; m/z 180.1 \rightarrow 80.1 (confirmation) and m/z 180.1 \rightarrow 101.1 (quantitation) for d_3 -cotinine; m/z 193.1 \rightarrow 80.1 (confirmation) and m/z 193.1 \rightarrow 134.1 (quantitation) for d_0 -trans-3-hydroxycotinine; m/z 196.1 \rightarrow 80.1 (confirmation) and m/z 196.1 \rightarrow 134.1 (quantitation) for d_3 -trans-3-hydroxycotinine. Peak areas were integrated using Xcalibur 3.0 (Thermo Scientific, Sunnyvale, CA) and the ratio of d_0/d_3 for each of the metabolites was determined. Total cotinine (free cotinine + cotinine glucuronide) and total trans-3-hydroxycotinine (free trans-3-hydroxycotinine + trans-3-hydroxycotinine glucuronide) were measured after beta-glucuronidase treatment, while free cotinine and free trans-3-hydroxycotinine glucuronide were measured without beta-glucuronidase treatment. The levels of cotinine glucuronide and trans-3-hydroxycotinine glucuronide were obtained by subtraction of the free metabolites from the total metabolites (\pm beta glucuronidase).

Statistical analysis was performed using SigmaPlot 12.5 (Systat Software, Inc. San Jose, CA). Group comparisons using t-test and Shapiro-Wilk ($P < 0.050$) for normality test were performed. Mann-Whitney Rank Sum Test was then performed for datasets that failed the normality test. P -value < 0.05 was considered statistically significant in the group comparisons.

RESULTS

To identify variations in biological pathways and their associated putative metabolites in smokers from two ethnic groups with differing lung cancer risk, sixty 24-hr urine samples (AA=30 and Whites=30) were used for the LC-HRMS-based metabolomics analysis acquired in both positive and negative modes. The subjects were all males and smoking at least 10 cigarettes per day (CPD). The TNE, which is the sum of smokers' urinary nicotine, cotinine, 3-hydroxycotinine, their corresponding glucuronides, plus nicotine N-oxide accounts for $>85\%$ of the nicotine dose consumed and considered as an excellent biomarker of cigarette smoke exposure was measured for these subjects.^{58, 59} The TNE levels for AA ranged from 30-159 nmol/mL, while for Whites, ranged from 20-173 nmol/mL (Table S1). In addition, the level of urinary NNAL (4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol), a biomarker of tobacco-specific nitrosamine exposure with longer half-life compared to TNE, was measured for the 60 subjects. The average concentration of urinary NNAL in AA was $218 \text{ pg mL}^{-1} (\pm 119)$ and $224 \text{ pg mL}^{-1} (\pm 110)$ for Whites. Both TNE and urinary NNAL levels showed no significant differences ($p=0.587$) between the two populations indicating that the cohort was exposed to similar levels of tobacco smoke and tobacco-specific carcinogens (Fig. S1 and Table S1).

Metabolic profiles of smokers' urine from AA and Whites.

The XCMS Online platform identified features that are significantly different between the two groups. The metabolic cloud plot in positive mode showed 114 features with fold change ≥ 1.5 and p -value ≤ 0.001 (Fig. 2a). Likewise, in negative mode, 36 features were detected with fold change ≥ 1.5 and p -value ≤ 0.001 (Fig. 2b). Minimal retention time shifts (< 0.6 min in both modes) were observed in the chromatographic runs of the 60 samples, which indicate good run-to-run reproducibility (Fig. S2). The multivariate principal component analysis (PCA) in both positive and negative modes showed modest separation between the two groups (Fig. 2c-d). This modest clustering was expected as both groups included only smokers, the sample size was relatively small, and other factors such as diet/lifestyle factors were not matched. Overall, a robust and reproducible LC-HRMS-based metabolomics analysis of the urine samples from AA and Whites yielded features or putative metabolites associated with differentially regulated metabolic pathways, with the majority of these pathways being down-regulated and resulting in lower levels of associated metabolites in AA compared to Whites.

Putative pathway analysis for identifying differentially regulated biological networks.

To identify differences in metabolic pathways in the two groups of smokers, an automated pathway analysis tool was performed as previously described.⁵⁴ Significant and differentially regulated features identified by XCMS Online were mapped onto known

biological pathways, and differentially regulated biological networks were identified using the Fisher's exact test based on the processed accurate mass spectral data.⁵⁴ The systems biology results show metabolic pathways differentially regulated between the two populations in both positive and negative mode (Table 1). In addition, an autonomous multi-modal pathway analysis was performed using both the positive and negative mode-acquired and processed dataset as previously described.⁵⁶ Using this integrated approach, metabolic pathways involving carbohydrate/sugar metabolism, amino acids, nucleic acids, fatty acids, and nicotine were identified (Tables S3 and S4). The nicotine degradation (p-value = 0.0002) and D-glucuronate degradation (p-value = 0.0002) pathways were significantly and differentially regulated between the two groups with reduced amounts of the putative metabolites detected in AA compared to Whites. The metabolites implicated in the nicotine degradation pathway using the multi-modal pathway analysis were: 3-pyridylacetate, 4-(3-pyridyl)-butanoate, cotinine methonium ion, cotinine glucuronide, and trans-3-hydroxycotinine glucuronide (Table 2; Table S4; Fig. S6). The metabolites, cotinine glucuronide, and trans-3-hydroxycotinine glucuronide, were down-regulated in AA compared to Whites (Table 2). To confirm the results obtained by our untargeted metabolomics analysis, the levels of cotinine glucuronide and trans-3-hydroxycotinine glucuronide were quantified by LC-MS/MS (Table S2).

In addition, the D-glucuronate degradation pathway was the top pathway identified being differentially regulated in AA in both positive (p=1.10E-08) and negative mode (p=3.60E-03) as well as in the multi-modal pathway analysis (p=2.0E-04) (Table 2 and Table S3). Figure 3 shows the pathway of the degradation of D-glucuronate in humans including the enzymes responsible for each step of the process. The metabolites associated in this pathway are 3-keto-L-gulonate, aldehyde-D-glucuronate, L-gulonate, and L-xylulose. The levels of these metabolites were lower in AA compared to Whites (Table 2; Table S4; Fig. S6). Figure 4 shows a representative EIC, zoomed precursor ion MS spectrum, and box plot for L-gulonate. Overall, metabolic pathways involving carbohydrate/sugar metabolisms that are differentially regulated in AA compared to Whites, were identified.

Targeted LC-MS analysis of nicotine metabolites.

To confirm the results of our metabolomics analysis, in particular when considering the metabolites implicated in the nicotine degradation pathway, the levels of cotinine glucuronide and trans-3-hydroxycotinine glucuronide were measured in the same samples used for the global untargeted metabolomics analysis. Total cotinine, total trans-3-hydroxycotinine, free cotinine, free trans-3-hydroxycotinine, cotinine glucuronide, and trans-3-hydroxycotinine glucuronide were measured in the 60 samples. Figure 5 shows the levels of nicotine metabolites (cotinine glucuronide, trans-3-hydroxycotinine glucuronide; free cotinine, free trans-3-hydroxycotinine, total cotinine, and total trans-3-hydroxycotinine) in the urine samples of AA and Whites. Figure S3a–b and Figure S4 a–b show representative chromatograms of the LC-MS/MS analysis of total cotinine and total trans-3-hydroxycotinine detected in the urine of a heavy smoker with their corresponding *d*₃-labeled internal standard. The levels of cotinine glucuronide are significantly lower (p<0.001) in AA compared to Whites, while the levels of free cotinine were significantly higher (p=0.026) in AA compared to White smokers. These results are consistent with previous studies where

the levels of cotinine glucuronide are lower and the free cotinine higher in the urine of AA compared to White smokers.¹⁵ The other metabolites did not show any significant difference in the levels between the two groups.

DISCUSSION

The overall goal of this work is to identify differentially regulated metabolic pathways and their associated metabolites using global MS-based metabolomics in a cohort with established differences in the levels of tobacco-related metabolites previously measured by targeted approaches. The identification of other pathways and biomarkers that may be relevant in cancer development may be used and validated as new potential panel of biomarkers to better understand and therefore potentially predict cancer susceptibility in smokers in future population-based studies. Previous studies demonstrated that African American (AA) smokers are at a higher risk of developing lung cancer compared to Whites.^{11,12} The metabolic pathways of nicotine and tobacco-specific carcinogens are well studied in these two groups; therefore, providing an ideal cohort to test the use of untargeted metabolomics methods to uncover other biologically relevant and novel pathways that are valuable in understanding the association of genetics, ethnic differences, and environmental and lifestyle factors such as smoking in disease development. The use of untargeted mass spectrometry-based metabolomics has been widely used to identify differentially regulated biological pathways resulting from tobacco exposure. For example, metabolomics analyses were used to compare the global metabolic profiles between smokers and non-smokers, and current smokers before and after smoking, and using mentholated cigarettes.^{40, 42–45} These metabolomics investigations serve as proof-of-principle for using metabolomics to identify novel tobacco-exposure biomarkers and provide important information on the dysregulated pathways and associated metabolites in smokers, which maybe relevant in cancer development while confirming the ability to detect the known differences in these groups.⁴⁵ Hsu et al. (2017) reported the identification of unique metabolites in smokers' plasma that are affected by acute smoking including menthol glucuronide, the reduction of glutamate, oleamide, and 13 glycerophospholipids.⁴⁵ However, global metabolomics studies comparing the variations in metabolic pathways in current smokers from different ethnic groups are limited or non-existent. There are metabolomics studies comparing the differences between non-smokers and smokers of the same ethnic background, but limited studies on the combined effects of smoking and race/ethnic group. Furthermore, because these studies have only used methods focusing on specific pathways related to a few toxicants, the overall changes in an individual's metabolic pathways could not be explored. Here, we present a unique study that relates global metabolic changes in smokers' urine from two populations of different ethnic backgrounds using untargeted metabolomics to test the possibility of identifying a panel of biomarkers, including those traditionally measured in relation to tobacco smoke, that may enable identification and stratification of highly susceptible population to the detrimental effects of tobacco use.

Metabolic information can be derived from a number of biological sources such as saliva, urine, and blood. However, urine provides a non-invasive and accessible biofluid for longitudinal studies. In addition, urine is a rich source of cellular metabolites and has been extensively used for diagnostic and clinical applications. So far, there are about 4500

metabolites detected in urine associated with approximately 600 human diseases/conditions such as obesity, cancer, inflammation, neurodegeneration, infectious disease, and diet to name a few.⁶⁰ Therefore, it is an ideal biofluid for global metabolomics studies as it reflects the overall metabolic network regulation of an individual resulting from specific and complex exposures. Using global LC-HRMS-based metabolomics, we have identified differentially regulated biological pathways and metabolites in urine of AA and White smokers. One pathway we have identified is the nicotine degradation pathway, which is different in AA compared to Whites. The metabolites associated with the pathway are down-regulated in AA compared to Whites according to our analysis. We have confirmed the low cotinine glucuronidation in AA using a targeted LC-MS/MS approach in the same aliquots of urine and is consistent with previous multi-ethnic studies.^{8, 13, 15, 19} In addition, we also found that trans-3-hydroxycotinine glucuronide is significantly lower in AA compared to Whites in our metabolomics analysis. However, we found no significant difference in the levels of trans-3-hydroxycotinine glucuronide in AA using a targeted approach (Fig. 5). Previous studies have shown that trans-3-hydroxycotinine glucuronide is not different within these groups.^{22, 57} When the levels of cotinine glucuronide and trans-3-hydroxycotinine glucuronide were adjusted to TNE, only cotinine glucuronide ($p=0.027$) showed significantly lower amounts, while trans-3-hydroxycotinine glucuronide ($p=0.061$) was not significant anymore in the metabolomics analysis. These results are consistent with our targeted analysis and previous studies and support the need to normalize the levels of nicotine metabolites to TNE, rather than to CPD, to account for the differences in the exposure levels or nicotine dose in smokers.^{15, 22} We have used the indirect measurement of glucuronide metabolites of nicotine such as cotinine glucuronide and trans-3-hydroxycotinine glucuronide (beta-glucuronidase treatment) as reported previously by others to allow appropriate comparison of our current results with previous studies that used such indirect approaches.^{15, 57} The targeted analysis of nicotine metabolites in the two groups confirmed the results of our metabolomics analysis, and therefore provided more confidence in the identification of the other differentially regulated pathways even those not directly related to nicotine. The biological relevance of these pathways, however, is still unknown and warrants further investigation.

The low levels of cotinine glucuronidation in AA is associated with relatively higher frequency of UGT2B10 splice variants in this population.^{19, 26, 27} For instance, UGT2B10 is the enzyme responsible for cotinine glucuronidation.^{19, 61} The majority (mean = 78% in two studies) of the UGT2B10-null individuals were AA in a multi-ethnic study based on the racial/ethnic specific frequency of the UGT2B10 splice variant.¹⁵ Genetic variations in UGT enzyme activities are associated with increased risk of developing solid cancers including colon, GI, lung, liver, oral, orolaryngeal and prostate cancer, which indicates that these variants are likely involved in detoxification of various carcinogens.⁶² Furthermore, the UGT2B10 splice variant common in AA have been shown to greatly increase drug exposure in this population and therefore should be considered in the treatment regimen.²⁶

In addition to the nicotine degradation pathway/cotinine glucuronidation, we have identified other pathways differentially regulated in the two groups. The D-glucuronate degradation pathway was highly and significantly dysregulated in AA in both, positive and negative ESI mode and in the multi-modal pathway analysis. The majority of the metabolites in the D-

glucuronate degradation pathway were down-regulated in AA. We hypothesize that the altered D-glucuronate degradation pathway could be influenced by the low glucuronidation in AA, or it could be due to differentially regulated carbohydrate metabolic pathways. The upstream pathways such as the sugar/carbohydrate degradation influence the level of the metabolites associated with downstream pathways such as the D-glucuronate degradation pathway. D-glucuronic acid is important in cellular processes including glucuronidation in xenobiotic metabolism and as a precursor for vitamin C biosynthesis.^{63, 64} These results demonstrate that even without using the nicotine or tobacco-specific biomarkers to stratify these two groups with differing lung cancer risk, the variations in the pattern and profile of metabolites as a result of differentially regulated biological networks between AA and Whites can still be ascertained. For instance, excluding the nicotine metabolites in the metabolomics analysis, unique patterns of metabolites in smokers were observed and were used to differentiate this population from non-smokers based on the global profiles and patterns of metabolites implicated in other affected pathways.⁴⁰ The differential regulation of the D-glucuronate degradation pathway is not likely smoke-induced as both groups of smokers showed no significant differences in TNE, an established biomarker of tobacco smoke exposure. We speculate that analysis of populations of AA and White non-smokers or healthy controls using our metabolomics approach would allow for the observation of the same metabolic alterations as observed for smokers, except for the nicotine degradation pathway, which is specific to tobacco use. Further studies are warranted to elucidate and understand the consequences of the differentially regulated D-glucuronate degradation pathway as it relates to ethnic and genetics differences and potentially to lung cancer susceptibility among smokers.

Variations in other pathways including the metabolism of amino acids, nucleic acids, and fatty acids were also identified (Table 1 **and** Table S3–S4). The fatty acid biosynthesis pathway was found to be significantly decreased in AA compared to Whites with putative metabolites, palmitate and oleate, being lower in AA. In addition, the amino acid degradation pathways (lysine degradation I and II; saccharopine and pipercolate pathways) were differentially regulated with lower amounts of putative metabolites in AA compared to Whites. Furthermore, the carbohydrate degradation and biosynthesis pathways (sucrose degradation, lactose degradation, D-galactose degradation, trehalose degradation, sorbitol degradation, myo-inositol de novo biosynthesis, D-myo-inositol (1,4,5)-trisphosphate biosynthesis, glycogenolysis, UDP-N-acetyl-D-galactosamine biosynthesis II) were also decreased in AA compared to Whites, with levels of putative metabolites implicated in the pathways being lower in AA (Table S4). Because the majority of the putative metabolites associated with the carbohydrate metabolic pathways are isobaric and associated with other pathways as well, it is difficult to decipher the exact nature of the differentially regulated metabolic pathway/s without synthetic standards. For example, β -D-glucose is implicated in multiple pathways including the sucrose degradation, lactose degradation, trehalose degradation, glycogenolysis, and UDP-N-acetyl-D-galactosamine biosynthesis II (Table S4). Further studies are warranted to investigate the biological relevance and implications of these differences between the two groups.

This study has its strengths and limitations. This work, for the first time, provides a comprehensive and global urinary metabolome of smokers from two ethnic groups with differing lung cancer risk. Because of the well-established differences in the metabolism of a few specific nicotine metabolites and tobacco-derived carcinogens in AA and Whites, we were able to further confirm these variations in the metabolomics analysis and by using a targeted approach to quantify selected nicotine metabolites in the same samples. In addition, the multi-modal pathway analysis (positive and negative mode data acquisition and data processing) identified other potentially relevant and novel pathways and putative metabolites not related to smoking, which may still be important in understanding differences in susceptibility to the detrimental effects of tobacco use. Although we were able to confirm the levels and structures of a few nicotine metabolites in the same samples using a targeted MS-based approach, other significant and differentially regulated metabolites detected need further confirmation in the future using synthetic standards. In addition, highly-curated and annotated metabolic pathways should be developed. For instance, the metabolic pathways of several tobacco-derived carcinogens such as NNK(4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone), NNN (N'-nitrosonornicotine), PAHs (polyaromatic hydrocarbons) are yet to be integrated into online pathway databases such as BioCyc (<https://biocyc.org>). Previous work of our colleagues and others have comprehensively characterized the metabolism of these tobacco-derived carcinogens as well as the metabolism of nicotine in both human and animal models.^{9, 15, 19, 65–67} The results from this extensive body of work need to be incorporated into pathway databases or in-house metabolomics workflows to enable more comprehensive “tobacco-focused” metabolomics studies in population-based settings to assess exposure and effects of tobacco use. Because the metabolomics analysis used here measures global metabolite profiles and typically the abundant ones in the sample, trace level compounds due to carcinogen-specific compounds such as NNK, NNN, and PAHs could be missed during the analysis. Furthermore, a minimal cleanup (desalting) step was performed before the metabolomics analysis to capture most of the metabolites in urine, and resulting however in being less likely to detect trace level metabolites, which need thorough sample cleanup. In fact, we have previously developed a “focused” metabolomics analysis of known and unknown NNK metabolites in rat urine and the workflow involved extensive cleanup to detect all known and novel NNK metabolites.⁶⁵ We plan to evaluate this “focused” metabolomics workflow in human smokers’ urine to investigate the NNK metabolite profiles across ethnic groups.

Another limitation of the present work is that we only used male subjects. This gender-specific cohort was used to avoid confounding factors that might complicate the metabolomics data analysis. Since previous studies have shown differences in lung cancer susceptibility between gender, females being at higher risk than males of the same ethnic group.^{68, 69} Furthermore, the differences in the diet between the two groups of smokers could also affect the observed variations in the urinary metabolite profiles. For instance, in a randomized, controlled, crossover trial on healthy volunteers, taking four different type of diets, differences in the urinary metabolic profiles were observed between the four groups.⁷⁰ While it is possible that the variations we observed in the untargeted metabolomics analysis could be an indirect or direct effects of diet on the overall global urinary metabolome in the two groups of smokers, we did not detect diet-specific metabolites such as hippurate (a

marker of fruit and vegetable consumption), (N-acetyl-)S-methyl-L-cysteine-sulfoxide (cruciferous vegetables), dimethylamine and TMAO (fish), and 1-methylhistidine and 3-methyl-histidine (oily fish and chicken) or observed differences in the levels of these compounds in smokers' urine.⁷⁰ Future investigation is warranted to evaluate these metabolomics results in a large cohort of subjects including females, influence of dietary intake, as well as in other ethnic groups where disparities in the levels of urinary metabolites are not consistent with the observed cancer risk for these populations. Finally, the metabolomics analysis was performed on a relatively small number of subjects (60 subjects total; AA, $n=30$ and White, $n=30$) resulting in reduced ability to perform any stratification. Nevertheless, the significant differences in the levels of the metabolites detected in such a small cohort, and the consistency of these findings with previous studies performed on selected metabolites in larger cohorts support the robustness of the method and the need to use this approach on a larger study.

Conclusions

The LC-HRMS-based metabolomics analysis of smokers' urine from AA and Whites with differing lung cancer risk identified differentially regulated biological pathways and metabolites, which may be used and validated in future studies as potential biomarkers for predicting the detrimental effects of tobacco use and ultimately contribute in the development of bioanalytical tools to predict lung cancer susceptibility among smokers in population-based studies. We have identified differentially regulated pathways including decreased nicotine degradation, in particular cotinine glucuronidation in AA, and the D-glucuronate degradation. Other pathways including the metabolism of amino acids, nucleic acids, and fatty acids were also identified. Further studies are warranted to investigate the biological relevance and implications of these differences in the metabolic pathway regulations between the two groups. Finally, our metabolomics analysis provides an alternative approach of characterizing the global patterns of metabolites in various ethnic groups where the disparities and differences in risks are not explained by the metabolites currently being measured and illustrate the importance of global profiling of all metabolites to identify other relevant biomarkers and dysregulated biological networks to ultimately develop tools to identify susceptible population to smoking-related lung cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank the American Society for Mass Spectrometry Postdoctoral Career Development Award to RD; Prof. Gary Siuzdak and group members (Aries Aisporna, Tao Han, Erica Forsberg, H. Paul Benton, Xavi Domingo, Rafael Montenegro, Carlos Guijas, Brian Hilmers, Bill Web, Linh Hoang, Winnie Heim Duane Rinehart, Bill Webb) at the Scripps Center for Metabolomics and Mass Spectrometry, The Scripps Research Institute, CA for hosting RD to conduct this work; Prof. Sharon E. Murphy and group members; Laura A. Maertens and Anshu Jain for the sample inventory. BW acknowledges the Austrian Science Fund (FWF) for funding (Erwin-Schrödinger Fellowship J-3808). P.V. acknowledges funding from grant NCI R50 CA211256. Targeted mass spectrometry was carried out in the Analytical Biochemistry Shared Resource of the Masonic Cancer Center, University of Minnesota, funded in part by Cancer Center Support Grant P30 CA-077598. This work was supported by grant CA-138338 from the National Cancer Institute.

References

1. Hecht SS, Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat Rev Cancer* 2003, 3 (10), 733–44. [PubMed: 14570033]
2. Hecht SS, Progress and challenges in selected areas of tobacco carcinogenesis. *Chem Res Toxicol* 2008, 21 (1), 160–71. [PubMed: 18052103]
3. Siegel RL; Miller KD; Jemal A, Cancer Statistics, 2017. *CA Cancer J Clin* 2017, 67 (1), 7–30. [PubMed: 28055103]
4. Rodriguez-Antona C; Gomez A; Karlgren M; Sim SC; Ingelman-Sundberg M, Molecular genetics and epigenetics of the cytochrome P450 gene family and its relevance for cancer risk and treatment. *Hum Genet* 2010, 127 (1), 1–17. [PubMed: 19823875]
5. Schwartz AG; Prysak GM; Bock CH; Cote ML, The molecular epidemiology of lung cancer. *Carcinogenesis* 2007, 28 (3), 507–18. [PubMed: 17183062]
6. Taioli E, Gene-environment interaction in tobacco-related cancers. *Carcinogenesis* 2008, 29 (8), 1467–74. [PubMed: 18550573]
7. Willett WC, Balancing life-style and genomics research for disease prevention. *Science* 2002, 296 (5568), 695–8. [PubMed: 11976443]
8. Murphy SE; Park SL; Balbo S; Haiman CA; Hatsukami DK; Patel Y; Peterson LA; Stepanov I; Stram DO; Tretyakova N; Hecht SS; Le Marchand L, Tobacco biomarkers and genetic/epigenetic analysis to investigate ethnic/racial differences in lung cancer risk among smokers. *NPJ Precis Oncol* 2018, 2, 17. [PubMed: 30155522]
9. Murphy SE, Nicotine Metabolism and Smoking: Ethnic Differences in the Role of P450 2A6. *Chem Res Toxicol* 2017, 30 (1), 410–419. [PubMed: 28092945]
10. Rowland A; Miners JO; Mackenzie PI, The UDP-glucuronosyltransferases: their role in drug metabolism and detoxification. *Int J Biochem Cell Biol* 2013, 45 (6), 1121–32. [PubMed: 23500526]
11. Le Marchand L; Wilkens LR; Kolonel LN, Ethnic differences in the lung cancer risk associated with smoking. *Cancer Epidemiol Biomarkers Prev* 1992, 1 (2), 103–7. [PubMed: 1306091]
12. Haiman CA; Stram DO; Wilkens LR; Pike MC; Kolonel LN; Henderson BE; Le Marchand L, Ethnic and racial differences in the smoking-related risk of lung cancer. *N Engl J Med* 2006, 354 (4), 333–42. [PubMed: 16436765]
13. Wang H; Park SL; Stram DO; Haiman CA; Wilkens LR; Hecht SS; Kolonel LN; Murphy SE; Le Marchand L, Associations Between Genetic Ancestries and Nicotine Metabolism Biomarkers in the Multiethnic Cohort Study. *Am J Epidemiol* 2015, 182 (11), 945–51. [PubMed: 26568573]
14. Park SL; Tiirikainen MI; Patel YM; Wilkens LR; Stram DO; Le Marchand L; Murphy SE, Genetic determinants of CYP2A6 activity across racial/ethnic groups with different risks of lung cancer and effect on their smoking intensity. *Carcinogenesis* 2016, 37 (3), 269–279. [PubMed: 26818358]
15. Murphy SE; Sipe CJ; Choi K; Raddatz LM; Koopmeiners JS; Donny EC; Hatsukami DK, Low Cotinine Glucuronidation Results in Higher Serum and Saliva Cotinine in African American Compared to White Smokers. *Cancer Epidemiol Biomarkers Prev* 2017, 26 (7), 1093–1099. [PubMed: 28264876]
16. Hecht SS; Carmella SG; Murphy SE; Akerkar S; Brunnemann KD; Hoffmann D, A tobacco-specific lung carcinogen in the urine of men exposed to cigarette smoke. *N Engl J Med* 1993, 329 (21), 1543–6. [PubMed: 8413477]
17. Benowitz NL; Dains KM; Dempsey D; Wilson M; Jacob P, Racial differences in the relationship between number of cigarettes smoked and nicotine and carcinogen exposure. *Nicotine Tob Res* 2011, 13 (9), 772–83. [PubMed: 21546441]
18. Wassenaar CA; Conti DV; Das S; Chen P; Cook EH; Ratain MJ; Benowitz NL; Tyndale RF, UGT1A and UGT2B genetic variation alters nicotine and nitrosamine glucuronidation in european and african american smokers. *Cancer Epidemiol Biomarkers Prev* 2015, 24 (1), 94–104. [PubMed: 25277794]
19. Murphy SE; Park SS; Thompson EF; Wilkens LR; Patel Y; Stram DO; Le Marchand L, Nicotine N-glucuronidation relative to N-oxidation and C-oxidation and UGT2B10 genotype in five ethnic/racial groups. *Carcinogenesis* 2014, 35 (11), 2526–33. [PubMed: 25233931]

20. Park SL; Kotapati S; Wilkens LR; Tiirikainen M; Murphy SE; Tretyakova N; Le Marchand L, 1,3-Butadiene exposure and metabolism among Japanese American, Native Hawaiian, and White smokers. *Cancer Epidemiol Biomarkers Prev* 2014, 23 (11), 2240–9. [PubMed: 25368399]
21. Benowitz NL; St Helen G; Dempsey DA; Jacob P; Tyndale RF, Disposition kinetics and metabolism of nicotine and cotinine in African American smokers: impact of CYP2A6 genetic variation and enzymatic activity. *Pharmacogenet Genomics* 2016, 26 (7), 340–50. [PubMed: 27035242]
22. Derby KS; Cuthrell K; Caberto C; Carmella SG; Franke AA; Hecht SS; Murphy SE; Le Marchand L, Nicotine metabolism in three ethnic/racial groups with different risks of lung cancer. *Cancer Epidemiol Biomarkers Prev* 2008, 17 (12), 3526–35. [PubMed: 19029401]
23. Patel YM; Park SL; Han Y; Wilkens LR; Bickeböller H; Rosenberger A; Caporaso N; Landi MT; Brüske I; Risch A; Wei Y; Christiani DC; Brennan P; Houlston R; McKay J; McLaughlin J; Hung R; Murphy S; Stram DO; Amos C; Le Marchand L, Novel Association of Genetic Markers Affecting CYP2A6 Activity and Lung Cancer Risk. *Cancer Res* 2016, 76 (19), 5768–5776. [PubMed: 27488534]
24. Murphy SE; Spina DA; Nunes MG; Pullo DA, Glucuronidation of 4-(hydroxymethyl)nitrosamino-1-(3-pyridyl)-1-butanone, a metabolically activated form of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, by phenobarbital-treated rats. *Chem Res Toxicol* 1995, 8 (5), 772–9. [PubMed: 7548761]
25. Murphy SE; Nunes MG; Hatala MA, Effects of phenobarbital and 3-methylcholanthrene induction on the formation of three glucuronide metabolites of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, NNK. *Chem Biol Interact* 1997, 103 (3), 153–66. [PubMed: 9134006]
26. Fowler S; Kletzl H; Finel M; Manevski N; Schmid P; Tuerck D; Norcross RD; Hoener MC; Spleiss O; Iglesias VA, A UGT2B10 splicing polymorphism common in african populations may greatly increase drug exposure. *J Pharmacol Exp Ther* 2015, 352 (2), 358–67. [PubMed: 25503386]
27. Patel YM; Stram DO; Wilkens LR; Park SS; Henderson BE; Le Marchand L; Haiman CA; Murphy SE, The contribution of common genetic variation to nicotine and cotinine glucuronidation in multiple ethnic/racial populations. *Cancer Epidemiol Biomarkers Prev* 2015, 24 (1), 119–27. [PubMed: 25293881]
28. Warth B; Spangler S; Fang M; Johnson CH; Forsberg EM; Granados A; Martin RL; Domingo-Almenara X; Huan T; Rinehart D; Montenegro-Burke JR; Hilmers B; Aisporna A; Hoang LT; Uritboonthai W; Benton HP; Richardson SD; Williams AJ; Siuzdak G, Exposome-Scale Investigations Guided by Global Metabolomics, Pathway Analysis, and Cognitive Computing. *Anal Chem* 2017, 89 (21), 11505–11513. [PubMed: 28945073]
29. Forsberg EM; Huan T; Rinehart D; Benton HP; Warth B; Hilmers B; Siuzdak G, Data processing, multi-omic pathway mapping, and metabolite activity analysis using XCMS Online. *Nat Protoc* 2018, 13 (4), 633–651. [PubMed: 29494574]
30. Tautenhahn R; Patti GJ; Rinehart D; Siuzdak G, XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal Chem* 2012, 84 (11), 5035–9. [PubMed: 22533540]
31. Allen F; Pon A; Wilson M; Greiner R; Wishart D, CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. *Nucleic Acids Res* 2014, 42 (Web Server issue), W94–9. [PubMed: 24895432]
32. Wolf S; Schmidt S; Müller-Hannemann M; Neumann S, In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinformatics* 2010, 11, 148. [PubMed: 20307295]
33. Wang L; Xing X; Chen L; Yang L; Su X; Rabitz H; Lu W; Rabinowitz JD, Peak Annotation and Verification Engine for Untargeted LC-MS Metabolomics. *Anal Chem* 2019, 91 (3), 1838–1846. [PubMed: 30586294]
34. Blaženovi I; Kind T; Sa MR; Ji J; Vaniya A; Wancewicz B; Roberts BS; Torbašinovi H; Lee T; Mehta SS; Showalter MR; Song H; Kwok J; Jahn D; Kim J; Fiehn O, Structure Annotation of All Mass Spectra in Untargeted Metabolomics. *Anal Chem* 2019, 91 (3), 2155–2162. [PubMed: 30608141]
35. Tsugawa H; Cajka T; Kind T; Ma Y; Higgins B; Ikeda K; Kanazawa M; VanderGheynst J; Fiehn O; Arita M, MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods* 2015, 12 (6), 523–6. [PubMed: 25938372]

36. Guijas C; Montenegro-Burke JR; Domingo-Almenara X; Palermo A; Warth B; Hermann G; Koellensperger G; Huan T; Uritboonthai W; Aisporna AE; Wolan DW; Spilker ME; Benton HP; Siuzdak G, METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Anal Chem* 2018, 90 (5), 3156–3164. [PubMed: 29381867]
37. Djoumbou-Feunang Y; Pon A; Karu N; Zheng J; Li C; Arndt D; Gautam M; Allen F; Wishart DS, CFM-ID 3.0: Significantly Improved ESI-MS/MS Prediction and Compound Identification. *Metabolites* 2019, 9 (4).
38. Gil-de-la-Fuente A; Godzien J; Saugar S; Garcia-Carmona R; Badran H; Wishart DS; Barbas C; Otero A, CEU Mass Mediator 3.0: A Metabolite Annotation Tool. *J Proteome Res* 2019, 18 (2), 797–802. [PubMed: 30574788]
39. Wishart DS; Feunang YD; Marcu A; Guo AC; Liang K; Vázquez-Fresno R; Sajed T; Johnson D; Li C; Karu N; Sayeeda Z; Lo E; Assempour N; Berjanskii M; Singhal S; Arndt D; Liang Y; Badran H; Grant J; Serra-Cayuela A; Liu Y; Mandal R; Neveu V; Pon A; Knox C; Wilson M; Manach C; Scalbert A, HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res* 2018, 46 (D1), D608–D617. [PubMed: 29140435]
40. Garcia-Perez I; Lindon JC; Minet E, Application of CE-MS to a metabonomics study of human urine from cigarette smokers and non-smokers. *Bioanalysis* 2014, 6 (20), 2733–49. [PubMed: 25413705]
41. Kaluarachchi MR; Boulangé CL; Garcia-Perez I; Lindon JC; Minet EF, Multiplatform serum metabolic phenotyping combined with pathway mapping to identify biochemical differences in smokers. *Bioanalysis* 2016, 8 (19), 2023–43. [PubMed: 27635669]
42. Ramakrishnan P; Nair S; Rangiah K, A method for comparative metabolomics in urine using high resolution mass spectrometry. *J Chromatogr A* 2016, 1443, 83–92. [PubMed: 27012786]
43. Hsu PC; Zhou B; Zhao Y; Resson HW; Cheema AK; Pickworth W; Shields PG, Feasibility of identifying the tobacco-related global metabolome in blood by UPLC-QTOF-MS. *J Proteome Res* 2013, 12 (2), 679–91. [PubMed: 23240883]
44. Hsu PC; Lan RS; Brasky TM; Marian C; Cheema AK; Resson HW; Loffredo CA; Pickworth WB; Shields PG, Menthol Smokers: Metabolomic Profiling and Smoking Behavior. *Cancer Epidemiol Biomarkers Prev* 2017, 26 (1), 51–60. [PubMed: 27628308]
45. Hsu PC; Lan RS; Brasky TM; Marian C; Cheema AK; Resson HW; Loffredo CA; Pickworth WB; Shields PG, Metabolomic profiles of current cigarette smokers. *Mol Carcinog* 2017, 56 (2), 594–606. [PubMed: 27341184]
46. Vulimiri SV; Misra M; Hamm JT; Mitchell M; Berger A, Effects of mainstream cigarette smoke on the global metabolome of human lung epithelial cells. *Chem Res Toxicol* 2009, 22 (3), 492–503. [PubMed: 19161311]
47. Barupal DK; Pinkerton KE; Hood C; Kind T; Fiehn O, Environmental Tobacco Smoke Alters Metabolic Systems in Adult Rats. *Chem Res Toxicol* 2016, 29 (11), 1818–1827. [PubMed: 27788581]
48. Cross AJ; Boca S; Freedman ND; Caporaso NE; Huang WY; Sinha R; Sampson JN; Moore SC, Metabolites of tobacco smoking and colorectal cancer risk. *Carcinogenesis* 2014, 35 (7), 1516–22. [PubMed: 24648381]
49. Ren X; Zhang J; Fu X; Ma S; Wang C; Wang J; Tian S; Liu S; Zhao B; Wang X, LC-MS based metabolomics identification of novel biomarkers of tobacco smoke-induced chronic bronchitis. *Biomed Chromatogr* 2016, 30 (1), 68–74. [PubMed: 26390017]
50. Stepanov I; Upadhyaya P; Carmella SG; Feuer R; Jensen J; Hatsukami DK; Hecht SS, Extensive metabolic activation of the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in smokers. *Cancer Epidemiol Biomarkers Prev* 2008, 17 (7), 1764–73. [PubMed: 18628430]
51. Hecht SS; Carmella SG; Chen M; Dor Koch JF; Miller AT; Murphy SE; Jensen JA; Zimmerman CL; Hatsukami DK, Quantitation of urinary metabolites of a tobacco-specific lung carcinogen after smoking cessation. *Cancer Res* 1999, 59 (3), 590–6. [PubMed: 9973205]
52. Warth B; Levin N; Rinehart D; Teijaro J; Benton HP; Siuzdak G, Metabolizing Data in the Cloud. *Trends Biotechnol* 2017, 35 (6), 481–483. [PubMed: 28117091]

53. Smith CA; O'Maille G; Want EJ; Qin C; Trauger SA; Brandon TR; Custodio DE; Abagyan R; Siuzdak G, METLIN: a metabolite mass spectral database. *Ther Drug Monit* 2005, 27 (6), 747–51. [PubMed: 16404815]
54. Huan T; Forsberg EM; Rinehart D; Johnson CH; Ivanisevic J; Benton HP; Fang M; Aisporna A; Hilmers B; Poole FL; Thorgersen MP; Adams MWW; Krantz G; Fields MW; Robbins PD; Niedernhofer LJ; Ideker T; Majumder EL; Wall JD; Rattray NJW; Goodacre R; Lairson LL; Siuzdak G, Systems biology guided by XCMS Online metabolomics. *Nat Methods* 2017, 14 (5), 461–462. [PubMed: 28448069]
55. Gowda H; Ivanisevic J; Johnson CH; Kurczy ME; Benton HP; Rinehart D; Nguyen T; Ray J; Kuehl J; Arevalo B; Westenskow PD; Wang J; Arkin AP; Deutschbauer AM; Patti GJ; Siuzdak G, Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses. *Anal Chem* 2014, 86 (14), 6931–9. [PubMed: 24934772]
56. Huan T; Palermo A; Ivanisevic J; Rinehart D; Edler D; Phommavongsay T; Benton HP; Guijas C; Domingo-Almenara X; Warth B; Siuzdak G, Autonomous Multimodal Metabolomics Data Integration for Comprehensive Pathway Analysis and Systems Biology. *Anal Chem* 2018, 90 (14), 8396–8403. [PubMed: 29893550]
57. Murphy SE; Wickham KM; Lindgren BR; Spector LG; Joseph A, Cotinine and trans 3'-hydroxycotinine in dried blood spots as biomarkers of tobacco exposure and nicotine metabolism. *J Expo Sci Environ Epidemiol* 2013, 23 (5), 513–8. [PubMed: 23443235]
58. Scherer G; Engl J; Urban M; Gilch G; Janket D; Riedel K, Relationship between machine-derived smoke yields and biomarkers in cigarette smokers in Germany. *Regul Toxicol Pharmacol* 2007, 47 (2), 171–83. [PubMed: 17034917]
59. Wang J; Liang Q; Mendes P; Sarkar M, Is 24h nicotine equivalents a surrogate for smoke exposure based on its relationship with other biomarkers of exposure? *Biomarkers* 2011, 16 (2), 144–54. [PubMed: 21323604]
60. Miller IJ; Peters SR; Overmyer KA; Paulson BR; Westphall MS; Coon JJ, Real-time health monitoring through urine metabolomics. *NPJ Digit Med* 2019, 2, 109. [PubMed: 31728416]
61. Berg JZ; von Weymarn LB; Thompson EA; Wickham KM; Weisensel NA; Hatsukami DK; Murphy SE, UGT2B10 genotype influences nicotine glucuronidation, oxidation, and consumption. *Cancer Epidemiol Biomarkers Prev* 2010, 19 (6), 1423–31. [PubMed: 20501767]
62. Mackenzie PI; Hu DG; Gardner-Stephen DA, The regulation of UDP-glucuronosyltransferase genes by tissue-specific and ligand-activated transcription factors. *Drug Metab Rev* 2010, 42 (1), 99–109. [PubMed: 20070244]
63. Mazerska Z; Mróz A; Pawłowska M; Augustin E, The role of glucuronidation in drug resistance. *Pharmacol Ther* 2016, 159, 35–55. [PubMed: 26808161]
64. Linster CL; Van Schaftingen E, Vitamin C Biosynthesis, recycling and degradation in mammals. *FEBS J* 2007, 274 (1), 1–22.
65. Dator R; von Weymarn LB; Villalta PW; Hooymann CJ; Maertens LA; Upadhyaya P; Murphy SE; Balbo S, In Vivo Stable-Isotope Labeling and Mass-Spectrometry-Based Metabolic Profiling of a Potent Tobacco-Specific Carcinogen in Rats. *Anal Chem* 2018, 90 (20), 11863–11872. [PubMed: 30086646]
66. Hecht SS; Young R; Chen CB, Metabolism in the F344 rat of 4-(N-methyl-N-nitrosamino)-1-(3-pyridyl)-1-butanone, a tobacco-specific carcinogen. *Cancer Res* 1980, 40 (11), 4144–50. [PubMed: 7471055]
67. Hecht SS, Approaches to cancer prevention based on an understanding of N-nitrosamine carcinogenesis. *Proc Soc Exp Biol Med* 1997, 216 (2), 181–91. [PubMed: 9349687]
68. Dresler CM; Fratelli C; Babb J; Everley L; Evans AA; Clapper ML, Gender differences in genetic susceptibility for lung cancer. *Lung Cancer* 2000, 30 (3), 153–60. [PubMed: 11137199]
69. Zang EA; Wynder EL, Differences in lung cancer risk between men and women: examination of the evidence. *J Natl Cancer Inst* 1996, 88 (3-4), 183–92. [PubMed: 8632492]
70. Garcia-Perez I; Posma JM; Gibson R; Chambers ES; Hansen TH; Vestergaard H; Hansen T; Beckmann M; Pedersen O; Elliott P; Stamler J; Nicholson JK; Draper J; Mathers JC; Holmes E; Frost G, Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. *Lancet Diabetes Endocrinol* 2017, 5 (3), 184–195. [PubMed: 28089709]

71. Haug K, Cochrane K, Nainala VC, Williams M, Chang J, Jayaseelan KV, and O'Donovan C
MetaboLights: a resource evolving in response to the needs of its scientific community. *Nucleic
Acids Res.* 2020, 48, D440–D444. [PubMed: 31691833]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

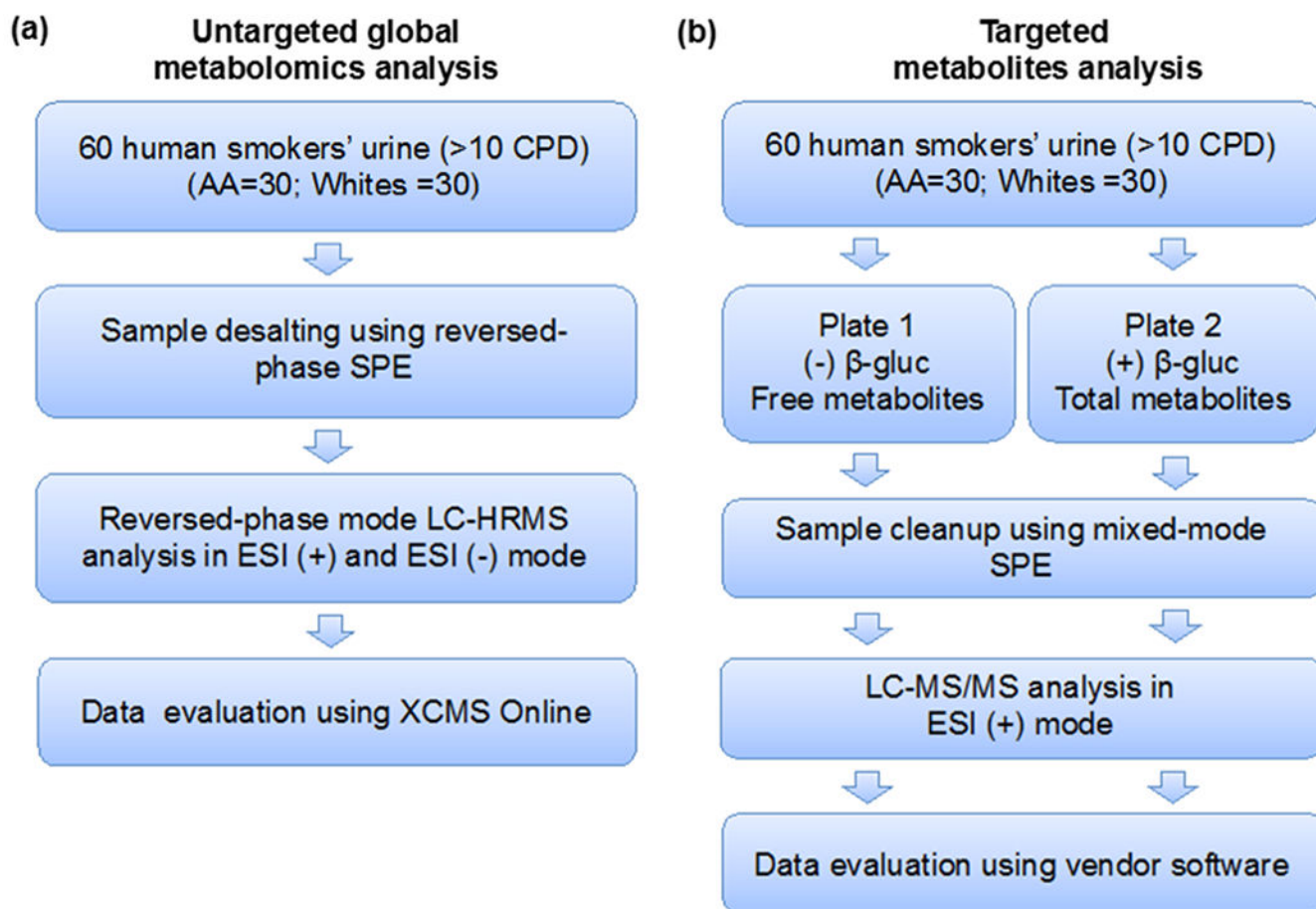


Fig. 1. Experimental workflow for global untargeted and targeted metabolomics analyses of smokers' urine from two ethnic groups. (a) Untargeted approach (b) Targeted approach.

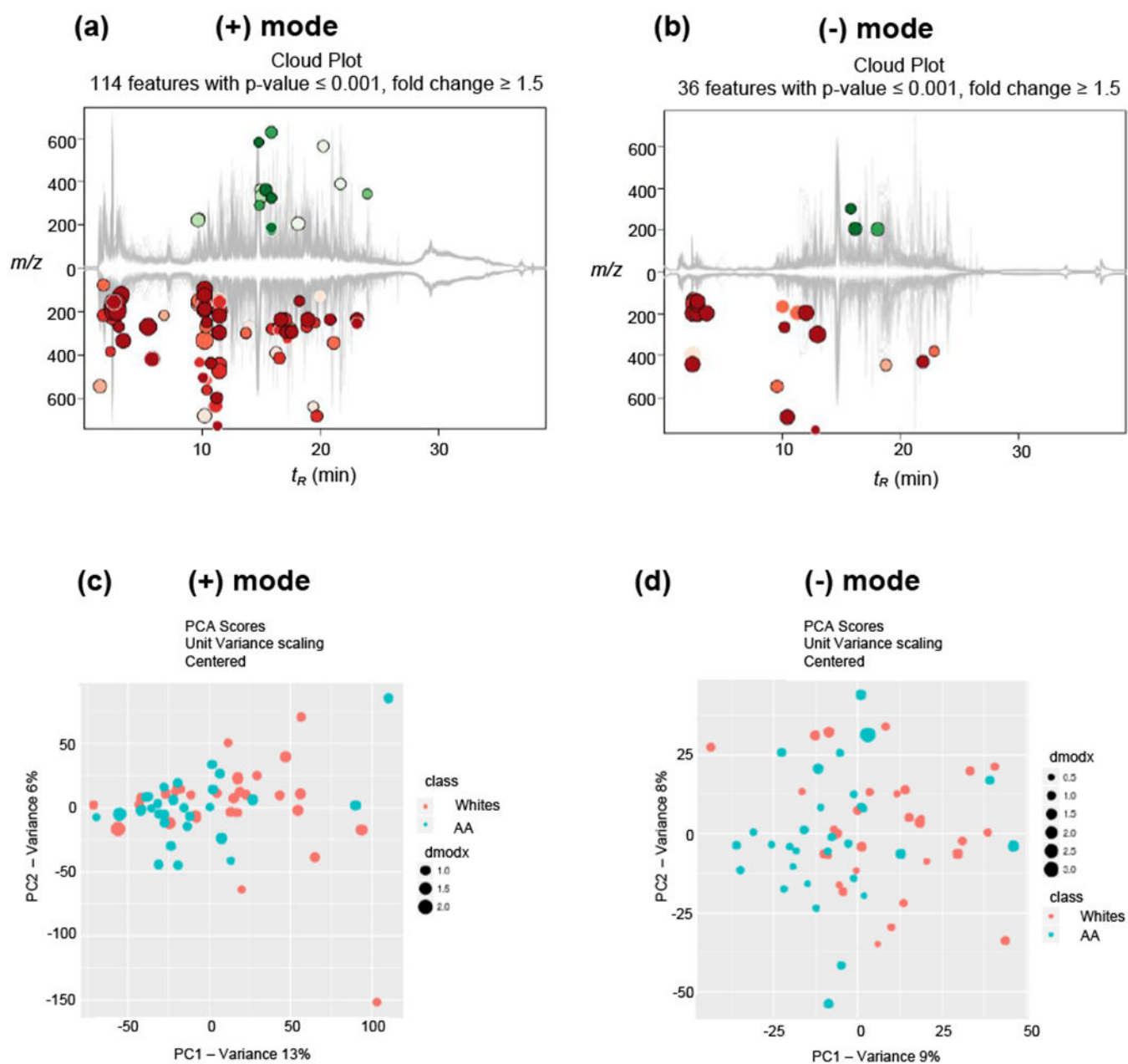


Fig. 2. Metabolic cloud plots showing significantly upregulated features (green circles) and down-regulated features (red circles) between the two groups in (a) positive mode and (b) negative mode. PCA analysis in both (c) positive mode and (d) negative mode showing modest separation between the two groups.

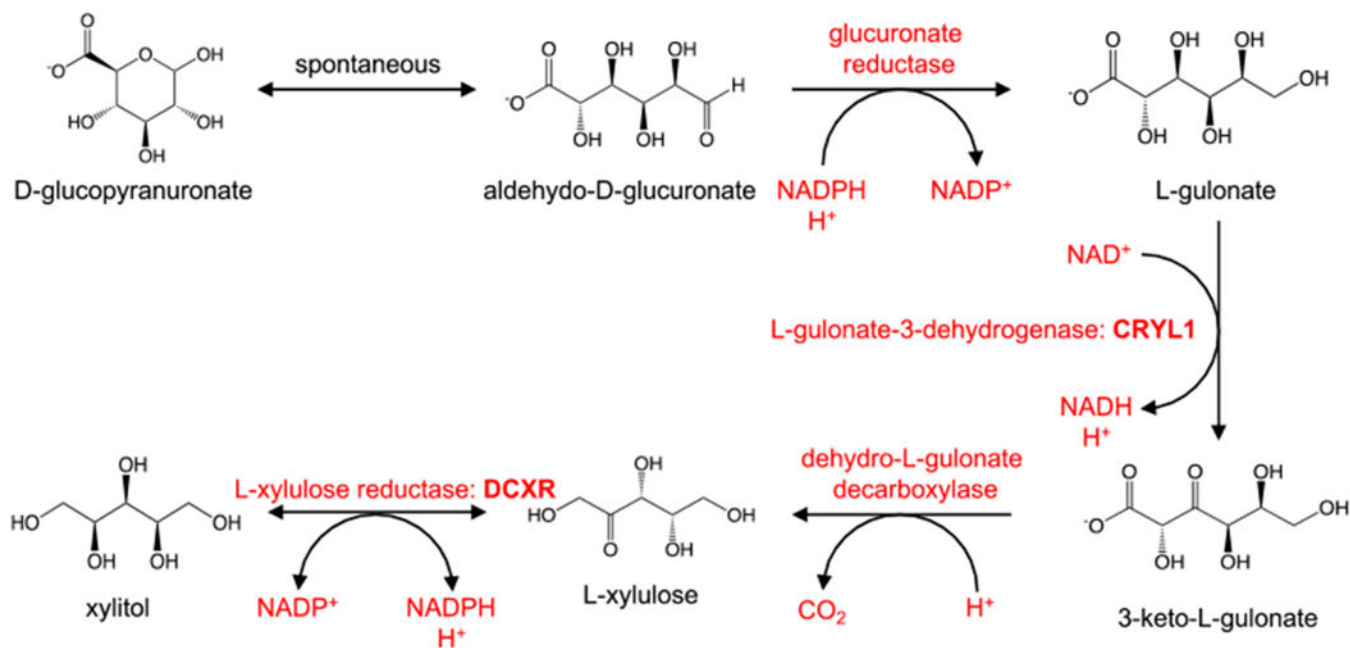


Fig. 3.

The D-glucuronate degradation pathway in *Homo sapiens* illustrating the different metabolites associated with the pathway (<https://biocyc.org>). The metabolites implicated in the pathway are down-regulated in AA compared to Whites. (<https://humancyc.org/HUMAN/NEW-IMAGE?type=PATHWAY&object=PWY-5525&detail-level=3>)

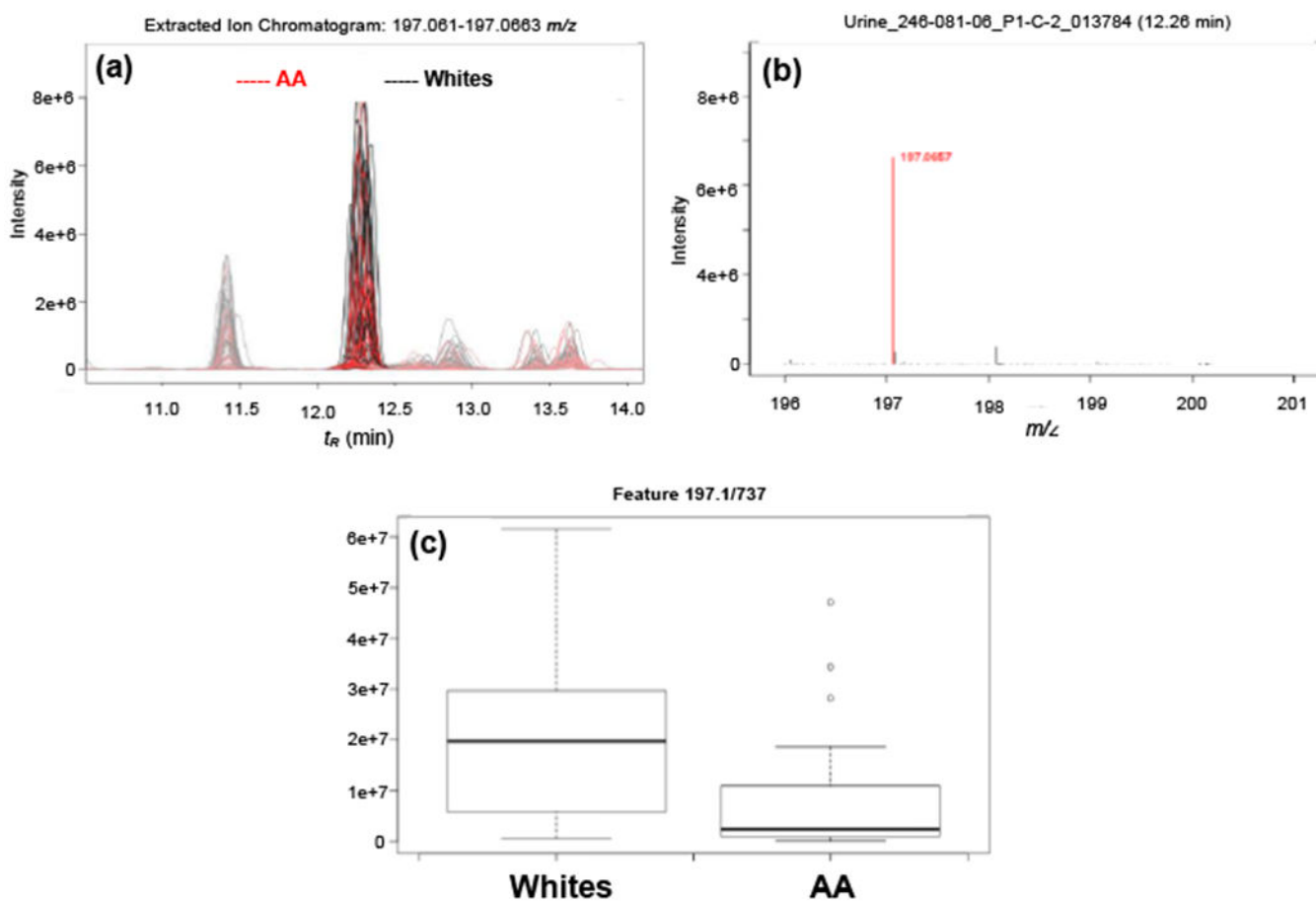


Fig. 4. Representative (a) extracted ion chromatogram (EIC), (b) zoomed precursor ion full scan MS spectrum in positive mode, and (c) box plot of the putative metabolite, L-gulonate (fold change =2.4; p-value = 0.0024).

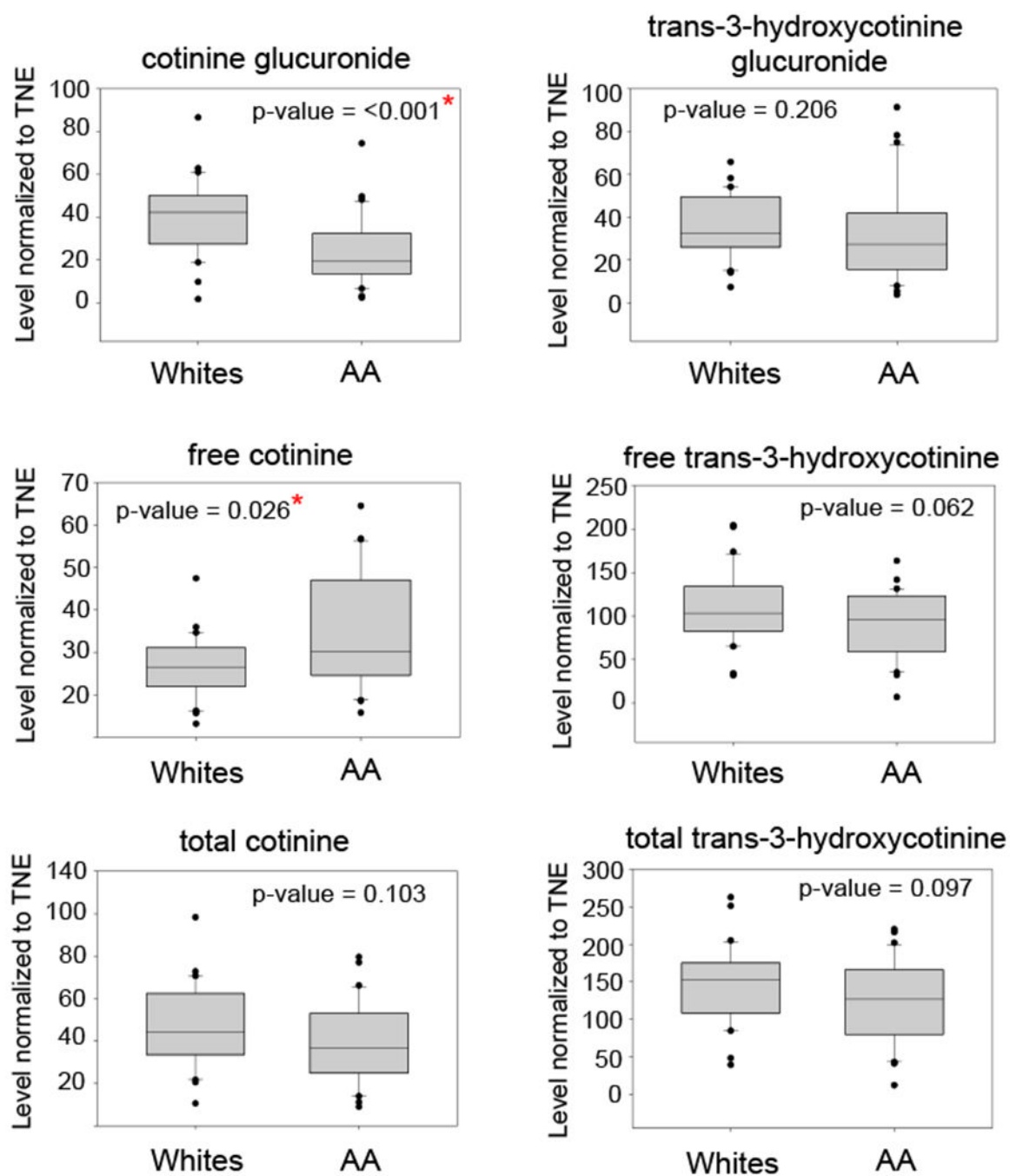


Fig. 5. Quantitation of nicotine metabolites (ng nmol^{-1}) in the 60 subjects using targeted LC-MS approach.

Table 1.

Differentially regulated metabolic pathways identified in positive and negative mode.

Pathway				
(+) mode	Overlapping putative metabolites	All metabolites	p-value	
D-glucuronate degradation	4	4	1.10E-08	
lysine degradation I (saccharopine pathway)	3	6	4.20E-08	
lactose degradation III	2	2	1.60E-07	
D-galactose degradation V (Leloir pathway)	2	2	1.60E-07	
trehalose degradation	2	3	5.20E-07	
bupropion degradation	3	4	1.50E-06	
sucrose degradation	3	5	4.20E-06	
lysine degradation II (pipecolate pathway)	2	8	5.50E-05	
tRNA charging	2	11	4.10E-04	
nicotine degradation V	2	18	9.80E-03	
(-) mode				
D-glucuronate degradation	4	4	3.60E-03	
tryptophan degradation via tryptamine	4	4	3.60E-03	
gluconeogenesis	2	2	3.10E-02	
sorbitol degradation I	2	2	3.10E-02	
taurine biosynthesis	2	2	3.10E-02	
lactose degradation III	2	2	3.10E-02	
D-galactose degradation V (Leloir pathway)	2	2	3.10E-02	
trehalose degradation	2	2	3.10E-02	
urate biosynthesis/inosine 5-phosphate degradation	2	2	3.10E-02	
adenosine nucleotides degradation	2	2	3.10E-02	
glycolysis	2	2	3.10E-02	
putrescine degradation III	2	2	3.10E-02	
lysine degradation II (pipecolate pathway)	3	6	4.60E-02	

Table 2.

Representative metabolites associated with the nicotine degradation and D-glucuronate degradation pathways in multi-modal pathway analysis.

Pathway/Metabolite	METLIN ID	Dysregulation *	Fold Change	p-value	m/z	t_R (min)	Adduct Type
nicotine degradation V							
3-pyridylacetate	NA	DOWN	2.1	4.50E-03	121.0279	12.9	[M-NH ₃ +H] ⁺¹
4-(3-pyridyl)-butanoate	NA	UP	1.7	6.60E-03	166.0855	13.7	[M+H] ⁺¹
cotinine-gluc	NA	DOWN	1.5	7.20E-03	373.1022	14.6	[M+Na-2H] ⁻¹
cotinine methonium ion	NA	DOWN	1.5	5.00E-03	209.1516	18.8	[M+NH ₃ +H] ⁺¹
trans-3-hydroxycotinine-gluc	NA	DOWN	2.1	6.00E-03	196.0593	12.8	[M+H+Na] ⁺²
nicotine degradation IV							
3-pyridylacetate	NA	DOWN	2.1	4.50E-03	121.0279	12.9	[M-NH ₃ +H] ⁺¹
4-(3-pyridyl)-butanoate	NA	UP	1.7	6.60E-03	166.0855	13.7	[M+H] ⁺¹
D-glucuronate degradation pathway							
L-xylulose	139	DOWN	2.2	3.00E-03	151.0604	9.1	[M+H] ⁺¹
L-gulonate	63144	DOWN	2.4	2.40E-03	197.0652	12.3	[M+H] ⁺¹
L-gulonate	63144	DOWN	3	1.50E-04	195.0503	12.0	[M-H] ⁻¹
aldehyde-D-glucuronate	NA	DOWN	2	1.70E-03	177.0392	13.2	[M-H ₂ O+H] ⁺¹
3-keto-L-gulonate	58394	DOWN	2	1.70E-03	177.0392	13.2	[M-H ₂ O+H] ⁺¹

* Dysregulation (fold change) relative to Whites; NA: not applicable