Environ Health Perspect

DOI: 10.1289/EHP2920

Supplemental Material

A Suspect Screening Method for Characterizing Multiple Chemical Exposures among a Demographically Diverse Population of Pregnant Women in San Francisco

Aolin Wang, Roy R. Gerona, Jackie M. Schwartz, Thomas Lin, Marina Sirota, Rachel Morello-Frosch, and Tracey J. Woodruff

Table of Contents

Figure S3. Relative concentration of suspect environmental organic acids (EOAs) with detection ≥ 60 that differed by education (A-D) or income (E) before multiple comparison adjustment at 0.1 level using the Kruskal-Wallis Rank Sum test (N=75). Matched chemicals from the EOA database are listed below the chemical formula and mean retention time (RT) of each suspect. The black line in the middle of the box represents median value. The top and bottom lines (also called "hinges") of each box represent the third quartile (75^{th} percentile) and the first quartile (25^{th} percentile) respectively. The height of the box represents the interquartile range (IQR) – distance between the first and third quartiles. The upper whisker extends from the upper hinge to the largest value no further than 1.5×IQR from the hinge. The lower whisker extends from the lower hinge to the smallest value at most 1.5×IQR of the hinge. Data beyond the end of the whiskers are called "outlying" points and are plotted individually as small circles.

Additional file- Code and data zip Document

Supplemental Tables

Table S1. A list of compounds that correspond to multiple chemical classes per chemical formula in our environmental organic acid (EOA) database

Chemical formula (n=9)	Compound name (n=25)	Chemical class	# of chemical classes per formula
C10H12O2	Carbofuran	Phenolic pesticides	2
	Eugenol	Phenols	
C12H10N2O	4-(Phenylazo)phenol	Phenols	2
	Fenazox	Acidic pesticides	
C12H7Br3O2	4-(2,4,6-Tribromo-phenoxy)phenol	Phenols	2
	Monohydroxy tribromodiphenyl ether	OH-BDEs	
C12H7Cl3O2	2,4,4'-Trichloro-2'-hydroxyphenyl ether	Phenols	2
	Hydroxy trichlorobiphenyl	OH-PCBs	
C13H15N3O3	Imazapyr	Acidic pesticides	2
	Pyrolan metabolite	Predicted pesticide metabolites	
C6H4Cl2O	1,4-Dichlorobenzene metabolite	Predicted pesticide metabolites	2
	2,3-Dichlorophenol	Phenols	
	2,4-Dichlorophenol	Phenols	
	2,5-Dichlorophenol	Phenols	
	2,6-Dichlorophenol	Phenols	
	3,4-Dichlorophenol	Phenols	
C8H8O3	2,4-Dihydroxyphenone	Phenols	2
	Methyl paraben	Phenols	
	Phenoxyacetic acid	Acidic pesticides	
C9H10O3	Dioxacarb artifact	Phenolic pesticides	2
	Ethyl-p-hydroxybenzoate (Ethyl paraben)	Phenols	
C12H10O	2-Phenylphenol	Phenolic pesticides	3
	3-Phenylphenol	Phenols	
	4-Phenylphenol	Phenols	
	Biphenyl metabolite	Predicted pesticide metabolites	

Table S2. Raw and adjusted *P* values of suspect environmental organic acids (EOAs) by demographic characteristics, serum samples from pregnant women at delivery (N=75)

Formula, RT	Matched chemical names	Race (n=67)[a]		Education (n=71)		Income (n=72)	
		Raw *P*	Adjusted *P*	Raw *P*	Adjusted *P*	Raw *P*	Adjusted *P*
C6H4Cl2O, 4.067	1,4-Dichlorobenzene metabolite 2,3-Dichlorophenol 2,4-Dichlorophenol 2,5-Dichlorophenol 2,6-Dichlorophenol 3,4-Dichlorophenol	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
C6HF13O3S, 5.058	Perfluorohexanesulfonic acid	<0.001	<0.001	0.025	0.115	0.004	0.079
C9H6O3, 1.204	4-Hydroxycoumarin	0.079	0.536	0.003	0.046	0.566	0.965
C14H18O4, 3.397	Monohexyl phthalate Monoisohexyl phthalate	0.508	0.900	0.006	0.046	0.290	0.965
C14H21NO2S, 0.787	Prosulfocarb metabolite	0.558	0.900	0.006	0.046	0.346	0.965
C15H20O4, 3.594	Monoheptyl phthalate Monoisoheptyl phthalate	0.668	0.906	0.003	0.046	0.198	0.918
C16H18N2O4, 4.207	Difenoxuron metabolite	0.656	0.906	0.004	0.046	0.802	0.984
C15H20O4, 3.876	Monoheptyl phthalate Monoisoheptyl phthalate	0.725	0.906	0.006	0.046	0.895	0.984
C20H33NO2, 8.198	Fenpropimorph metabolite	0.764	0.910	0.002	0.046	0.277	0.965
C13H15N3O3, 0.858	Imazapyr Pyrolan metabolite	1	1	0.003	0.046	0.837	0.984
C6H4N2O5, 1.579	2,4-Dinitrophenol	0.019	0.466	0.016	0.102	0.003	0.079
C16H22O4, 5.139	Mono (2-ethylhexyl) phthalate Monoisooctyl phthalate Monooctyl phthalate	0.228	0.659	0.048	0.160	0.005	0.079
C11H14O3, 5.074	Butyl paraben	0.320	0.686	0.264	0.447	0.005	0.079

Note: Only compounds with detection frequency less than 60 were considered in this analysis. Compounds with detection frequency greater than or equal to 60 were analyzed based on their relative concentration. We further limited this analysis to compounds that were detected in at least 20% of participants with non-missing values for race/ethnicity, education, income, and nativity (U.S.-born status), resulting in 75, 70, 74, and 81 suspect EOAs being included for each demographic comparison respectively. *P*-values were obtained using a Fisher's exact test. We calculated adjusted *p*-values to correct for multiple comparisons for each demographic comparison using the Benjamini–Hochberg procedure (i.e., the false discovery rate method). None of the suspect EOAs differed by nativity due to the limited number of participants who answered this question (n=44) and thus results for nativity were not included in this table.

[a]Analysis was performed on a sample restricted to non-Hispanic whites, Latinas, and Asians only due to small sample sizes of non-Hispanic African Americans and other races.

Table S3. Selected chemical formula, suspect features from the sample, and the corresponding candidate chemicals for confirmation analysis

Formula	RT[a]	DF[a]	Sel[b]	Confirmed	Selected compound for validation	CAS RN	HPV[c]	ToxCast[d]	TSCA[e]	$RT_{Standard}$
C10H12O	3.498	1		No	Estragole	140-67-0	√	√	√	NP
	3.823	1		No						
	4.144	29	√	No						
	4.348	1		No						
C10H12O2	1.499	1		No	Eugenol	97-53-0		√	√	NP
	3.887	10		No						
	4.072	36	√	No						
	4.543	3		No						
C10H14O	3.815	2		No	2-tert-Butylphenol	88-18-6	√	√	√	5.6
	4.029	2		No	4-sec-Butylphenol	99-71-8	√	√	√	5.4
	4.246	43	√	No						
C10H14O2	2.19	1		No	4-Butoxyphenol	122-94-1			√	5.4
	2.403	1		No						
	2.926	2		No						
	4.029	70	√(A)	No						
C11H14O2	4.008	2		No	Methyl eugenol	93-15-2		√	√	NP
	4.38	2		No						
	4.629	3		No						
	5.129	63	√(A)	No						
C12H18O2	0.831	2		No	4-Hexyloxyphenol	18979-55-0		√	√	5.8
	3.985	9		No						
	4.767	58	√	No						
	5.27	2		No						
C13H20O2	1.113	1		No	4-Heptyloxyphenol	13037-86-0			√	6.2
	3.863	11		No						
	4.331	21	√	No						
	4.793	19		No						
	5.276	21		No						
C14H22O	4.321	5		No	2,6-Di-tert-butylphenol	128-39-2	√	√	√	7.25
	4.433	9		No						
	4.788	17		No						
	5.066	33		No						

Formula	RT[a]	DF[a]	Sel[b]	Confirmed	Selected compound for validation	CAS RN	HPV[c]	ToxCast[d]	TSCA[e]	$RT_{Standard}$
	5.153	40		No						
	5.283	38		No						
	6.719	74	√(A)	Yes[f]	2,4-Di-tert-butylphenol (2,4,-DTBP)	96-76-4	√	√	√	6.75
				Yes[f]	4-tert-Octylphenol	140-66-9	√	√	√	6.5
C15H22O3	3.791	1		No						
	4.537	8		No						
	5.132	64	√	Yes	3,5-Di-tert-Butylsalicylic acid	19715-19-6			√	5.3
C16H26O2	4.191	6		No	4-tert-Octylphenol monoethoxylate solution	2315-67-5				NP
	4.545	3		No						
	4.779	5		No						
	6.153	61	√(A)	No						
C6H4N2O5	0.579	8		No						
	1.579	52	√(A)	Yes	2,4-Dinitrophenol (2,4-DNP)	51-28-5	√	√	√	1.6~2.0
C6H6O2	0.942	33		No						
	1.352	39	√	Yes	Pyrocatechol	120-80-9	√	√	√	1.49
	2.364	2		No						
C7H8O	0.893	3		No	2-Methylphenol	95-48-7	√	√	√	4.2
	1.75	1		No	4-Methylphenol	106-44-5	√	√	√	4.4
	1.999	70	√	No						
	3.239	1		No						
	3.587	1		No						
	4.357	1		No						
C8H8O2	1.039	54	√	No						
	1.555	1		No						
	2.489	2		No						
	3.027	17		No						
	3.422	6		Yes[g]	2-Hydroxyacetophenone	118-93-4		√	√	3.3
				Yes[g]	3'-Hydroxyacetophenone	121-71-1			√	3.3
	4.051	15		No						
	5.024	6		No						
C8H8O3	1.187	15		No	Phenoxyacetic Acid	122-59-8		√	√	0.86
	1.931	66	√(A)	No						
	2.393	2		No						
	3.888	35		No						
	4.273	1		No						

Formula	RT[a]	DF[a]	Sel[b]	Confirmed	Selected compound for validation	CAS RN	HPV[c]	ToxCast[d]	TSCA[e]	$RT_{Standard}$
C9H6O3	0.927	1		Yes	4-Hydroxycoumarin	1076-38-6			√	0.8
	1.204	20	√(A)	No						

Abbreviations: EOA, environmental organic acid; DF, detection frequency; RT, retention time (mean); HPV, high production volume; $RT_{standard}$, retention time of the corresponding reference standard obtained from the QTOF validation analysis; NP, not present.

Note: confirmed compounds are highlighted in blue. Information on each suspect EOA (unique combination of formula and mean RT) is listed in the first four columns. The columns that were shaded in purple contain information on the selected candidate compound(s) for validation. There can be multiple candidate compounds being selected for validation and we will compare the $RT_{standard}$, accurate mass and MS/MS spectral pattern of a reference standard to that of each of the suspects with the same formula. We confirmed the presence of the suspect EOA if it has the same RT, mass, and MS/MS spectral pattern as that of the corresponding reference standard.

[a]RT (mean) and DF for suspect EOAs detected in our samples.

[b]Suspect EOA selected initially for compound confirmation process. "(A)" annotates compounds that were chosen at the first round because of their high detection (DF≥60) and showing potential demographic differences in relative concentration based on raw *p*-values < 0.1 (Supplementary Figure S3) or in detection frequency based on adjusted *p*-values < 0.1 (Supplementary Table S4). Compounds without annotation were chosen at subsequent rounds according to their rank order of detection frequency.

[c]Chemicals being manufactured and/or imported into the US with an aggregate volume of 1 million to 10 million pounds/year, according to the US HPV list by US Environmental Protection Agency (2004).

[d]Toxicity ForeCaster (ToxCast™) Data: Chemical information database, available at: ftp://newftp.epa.gov/comptox/High_Throughput_Screening_Data/DSSTox_Oct_2015

[e]The Toxic Substances Control Act (TSCA) Chemical Substance Inventory, available at: https://www.epa.gov/tsca-inventory

[f] 4-tert-Octylphenol has RT very close to 2,4,-DTBP and is currently biomonitored by NHANES. Thus, it is not counted towards confirmed novel compounds.

[g]2-Hydroxyacetophenone or 3'-Hydroxyacetophenone being isomers with very close structures and cannot be distinguished by RT solely in the current LC-QTOF/MS analysis

Supplemental Figures

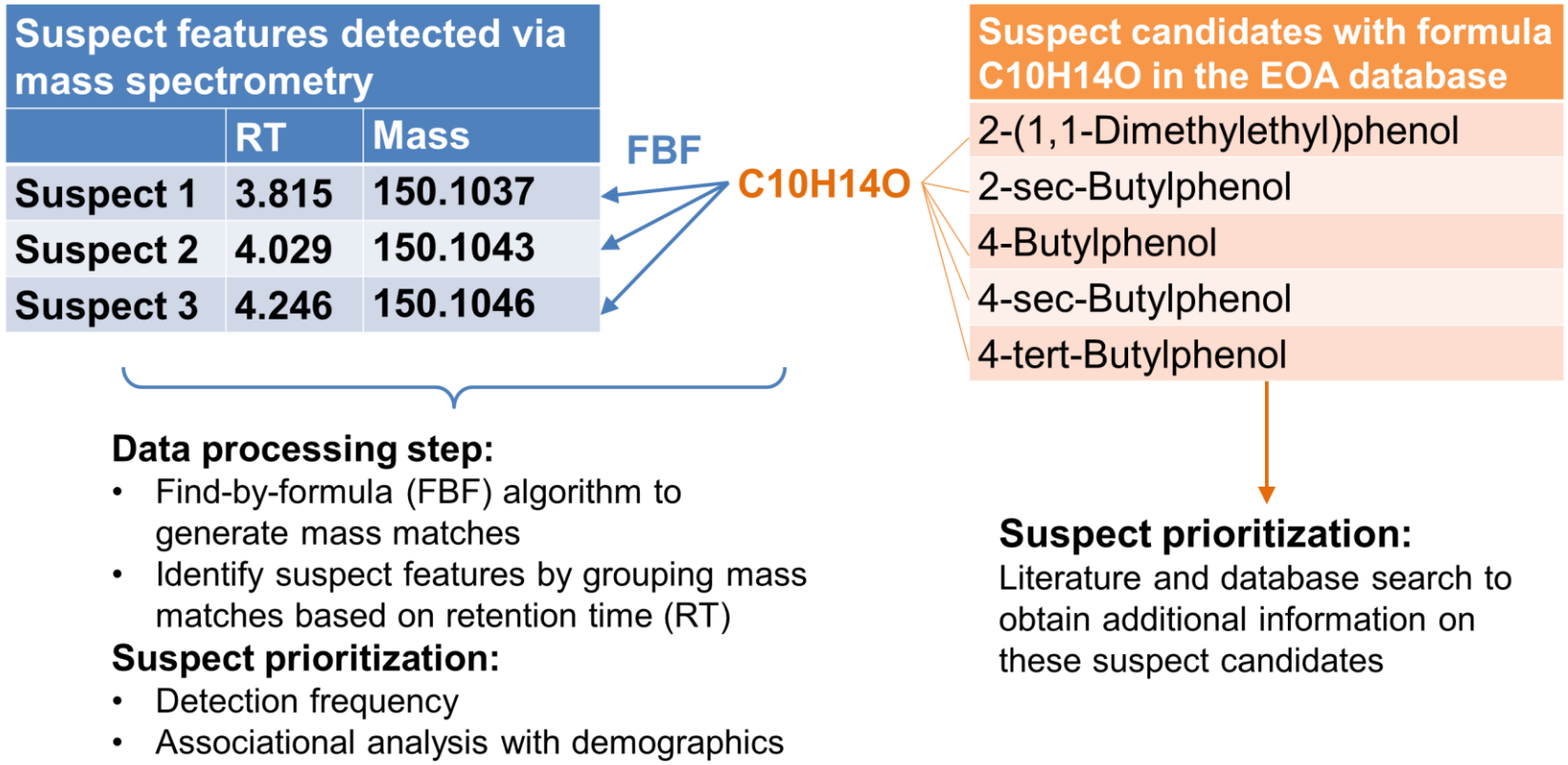

Figure S1. Illustration of many (suspect features)-to-many (chemicals) matching. When isomers are present, we could detect multiple chemical features via high-resolution mass spectrometry, indicated by the same (allowing for small measurement error) mass and different retention time. Each of these features can be matched to multiple suspect candidates with the same chemical mass (*i.e.*, formula) included in our EOA database (with equal probability). The exact match between a specific suspect feature and its chemical identity will need to be confirmed using synthetic reference standards.

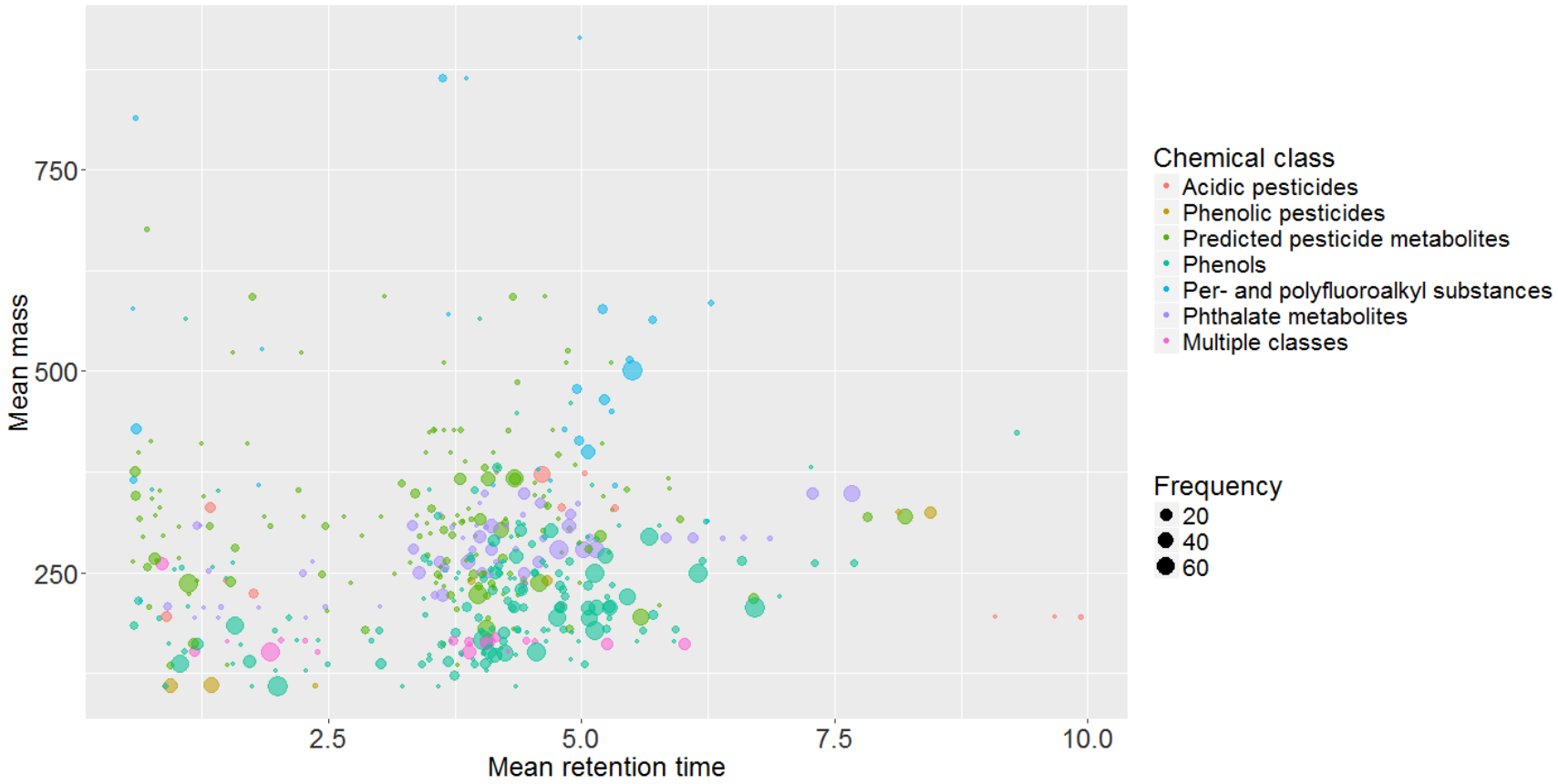

Figure S2. Overview of 455 suspect environmental organic acids (EOAs) detected in 75 serum samples by retention time and mass, colored by matched chemical class.

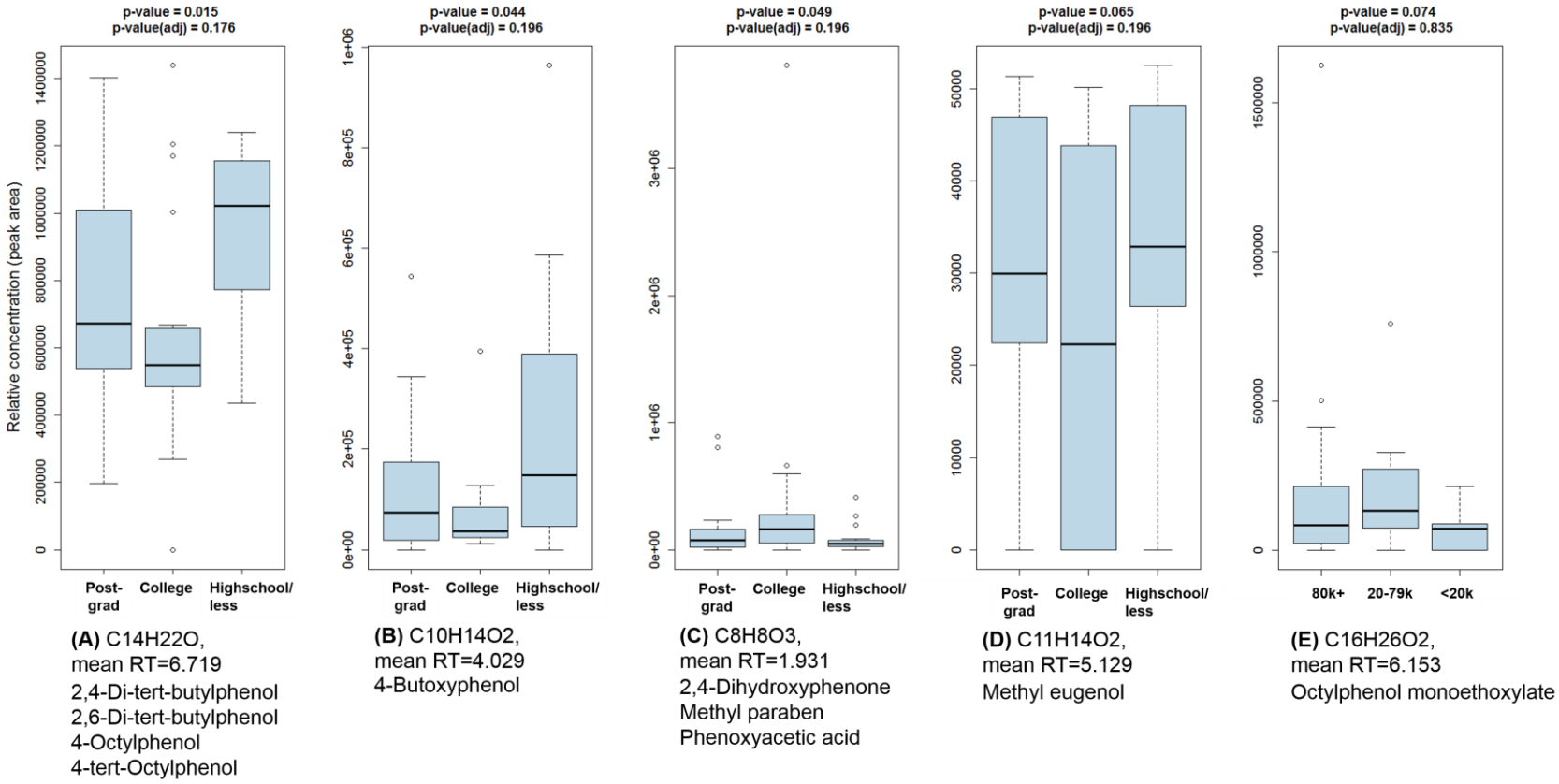

Figure S3. Relative concentration of suspect environmental organic acids (EOAs) with detection ≥ 60 that differed by education (A-D) or income (E) before multiple comparison adjustment at 0.1 level using the Kruskal-Wallis Rank Sum test (N=75). Matched chemicals from the EOA database are listed below the chemical formula and mean retention time (RT) of each suspect. The black line in the middle of the box represents median value. The top and bottom lines (also called "hinges") of each box represent the third quartile (75^{th} percentile) and the first quartile (25^{th} percentile) respectively. The height of the box represents the interquartile range (IQR) – distance between the first and third quartiles. The upper whisker extends from the upper hinge to the largest value no further than 1.5×IQR from the hinge. The lower whisker extends from the lower hinge to the smallest value at most 1.5×IQR of the hinge. Data beyond the end of the whiskers are called "outlying" points and are plotted individually as small circles.