Development and Characterization of a Set of Internal Standards, Chemical Libraries and Long Term Reference Samples to Enable Service-Member Related Exposure Studies

Warren Casey[1], Jennifer Kyle[2], Kruttika Dabke[4], Michelle Embry[3], Cara L. Sake[4], Tao Liu[2], Stella Somiari[5], Jerry S.H. Lee[4,6,7], Craig D. Shriver[8] , Jonathan E. Katz[4,6,9]

[1]Division of Translational Toxicology, National Institute of Environmental Health Sciences
[2]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA
[3]Health and Environmental Sciences Institute, Washington, DC
[4]Ellison Institute of Technology, Los Angeles, CA
[5]Chan Soon-Shiong Institute of Molecular Medicine at Windber, Windber, PA
[6]Murtha Cancer Center/Research Program, Uniformed Services University of the Health Sciences/Walter Reed National Military Medical Center, Bethesda, MD
[7]Department of Chemical Engineering and Material Sciences, Viterbi School of Engineering, University of Southern California, Los Angeles, CA
[8]Department of Surgery, Uniformed Services University of the Health Sciences, Bethesda, MD
[9]Department of Medicine/Oncology, Keck School of Medicine, University of Southern California, Los Angeles, CA
Correspondence: jkatz@eit.org
Reprints: https://reprints.katzlab.org/

BACKGROUND

Military service may result in a heightened risk of cancer and other health related issues due to environmental or hazardous exposure. This project aims to develop a robust set of tools to assist in exposure science research.

In biological exposure studies assessing small molecules (xenobiotic or endogenous), well defined materials are required to assess compound "behavior" in an assay as well as for platform and process validation. Towards these purposes we established three products: An "**Exposure Compound Library**" (~300 entries) to identify compound identities across experiments, an "**Exposure Reference Standard**" (~50 entries) of isotopically labeled internal standards for quality assurance/quality control assessment, and a "**Long Term Reference Sample**" for platform/process validation, assessment of technical variability and bridging experiments.

Discussed are the formulations, compound selection rationale, chemical diversity, and initial characterization of these three products.

HIGHLIGHTS

- **Understand the roles of "compound libraries", "internal standards" and "long term reference samples" in molecular analytic workflows and their utility in validating analytic platforms and understanding measurement variability.**
- **Understand the rationale that governed the selection of specific compounds for our exposure libraries and internal standards. Understand the utility of these compounds for class based relative quantitation.**
- **Understand the role of a long term reference sample in validating workflows and bridging datasets measured in different laboratories or at different times.**

Methods

EXPOSURE COMPOUND LIBRARY

Compound libraries, in mass spectrometry (MS), are procured standards, extensively analyzed, that can be used to register m/z, retention time (RT) and fragmentation (ms2) against putative unknown analytes in a mixture (**Figure 1**). Reference chemicals were selected through a review of the literature (**Table 1**) aimed to identify chemically diverse substances with potential cancer risks relevant to military exposures in garrison and while deployed. Chemicals listed included those identified by the National Toxicology Program's Report on Carcinogens as "known to be a human carcinogen" or "reasonably anticipated to be a human carcinogen" [1] as well as those identified by the International Agency for Research on Cancer (IARC)as "carcinogenic to humans (Group1) or "probably/possibly carcinogenic to humans" (Groups 2A and 2B) [2]. Specific consideration was given to other chemicals of concern associated with burn pits, pesticide application, and water contaminated with jet fuels and firefighting foams [3-11]. Chemical classes of interest included but were not limited to: Per- and polyfluoroalkyl substances (PFAS), dioxins and furans (PDCCs and PCDFs) , polycyclic aromatic hydrocarbons (PAHs), volatile organic chemicals (VOCs), phenols, parabens, and phthalates.

LONG TERM REFERENCE SAMPLE

Pooled (separate male and female) human sample materials (both plasma and serum) were procured via BioIVT (Westbury, NY) and described in **Figure 2**. Aliquoting protocols are shown in **Figure 3** and describe how to match "thaw count" with external samples and where to introduce reference compounds.

EXPOSURE INTERNAL STANDARD

MS internal standards are typically compounds in which one or more atoms are replaced by their non-natural isotopes to create unique MS signatures. Our initial list (**Table 2**) is a combination of compounds from **Table 1** augmented by compounds of interest to the authors as well as compounds identified as xenobiotics frequently found in human populations[12]. As these samples are meant to act as internal standards used with every sample, we sought to reduce the complexity of this list to improve handling. Our goal was to create a "class based" exposure standard set in which a majority of compound classes of interest were represented such that each selected standard could act as a surrogate for all similar compounds (e.g., Valine should behave similarly enough to Alanine/Leucine/Isoleucine as to act as a "stand in"). To reduce our list we "fingerprinted" our compounds using Extended Connectivity Fingerprinting (ECFP, [13]) with a radius of 4 atoms (**Figure 4**) and then clustered the compounds based on Tanimoto similarity ([14], **Figure 5**). Individual clusters formed with Butina Clustering [15], similarity 0.5 were inspected and reduced to a single representative (**Figure 6**) resulting in a set of compounds of reduced similarity (**Figure 7**).

Our final set of compounds is shown in **Figure 8**. To validate the landscape of compounds we have captured, we have overlaid these compounds on top of a structural similarity graph created from our **Compound Library** (**Figure 9**) suggesting that we have covered most of the landscape represented by our library.

Results

The **LTRS** has been acquired and distributed to Windber Research Institute (WRI) as well as Pacific Northwest National Laboratory (PNNL) in support of **MCCRP** and **PROMETHEUS** efforts. PNNL is completing its measures of the analytes in the **Exposure Compound Library**. Currently we are in final review of expected biological concentrations for the **Exposure Internal Standard** after which these heavy isotope mixtures will synthesized and made available to the community as "kits" for their own exposomics endeavors.

ACKNOWLEDGEMENTS/CONFLICT OF INTEREST

Funding in part for these efforts was provided from Uniformed Services University of the Health Sciences (USUHS) award from the Defense Health Program to the Murtha Cancer Center Research Program (HU00011820032, JE Katz/JSH Lee) administered by the Henry M. Jackson Foundation for the Advancement of Military Medicine and by USUHS of the DoD through Biomarkers of Cancers and Exposure-related Cancer Risk under Award No. 23.178 (to T. Liu). Disclaimer: The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views, opinions, or policies of the USUHS, the Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., the Department of Defense (DoD), the Departments of the Army, Navy, or Air Force. Mention of trade names, commercial products, or organization does not imply endorsement by the U.S. Government.

References

1. NTP (National Toxicology Program). 2021. Report on Carcinogens, Fifteenth Edition. Research Triangle Park, NC: U.S. Department of Health and Human Services, Public Health Service. https://ntp.niehs.nih.gov/go/roc15 (EndNote XML) DOI: https://doi.org/10.22427/NTP-OTHER-1003
2. IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Man. Lyon :International Agency for Research on Cancer, 19721987.
3. Xia X, Carroll-Haddad A, Brown N, Utell MJ, Mallon CT, Hopke PK. **Polycyclic Aromatic Hydrocarbons and Polychlorinated Dibenzo-p-Dioxins/Dibenzofurans in Microliter Samples of Human Serum as Exposure Indicators.** J Occup Environ Med. 2016 Aug;58(8 Suppl 1):S72-9. doi: 10.1097/JOM.0000000000000743. PMID: 27501107; PMCID: PMC4977992.
4. Smith MR, Uppal K, Walker DI, Utell MJ, Hopke PK, Mallon TM, Krahl PL, Rohrbeck P, Go YM, Jones DP. **Environmental Chemicals Altered in Association With Deployment for High Risk Areas**. J Occup Environ Med. 2019 Dec;61 Suppl 12(Suppl 12):S15-S24. doi: 10.1097/JOM.0000000000001647. PMID: 31800447; PMCID: PMC7299071
5. Smith MR, Uppal K, Walker DI, Utell MJ, Hopke PK, Mallon TM, Krahl PL, Rohrbeck P, Go YM, Jones DP. **Environmental Chemicals Altered in Association With Deployment for High Risk Areas**. J Occup Environ Med. 2019 Dec;61 Suppl 12(Suppl 12):S15-S24. doi: 10.1097/JOM.0000000000001647. PMID: 31800447; PMCID: PMC7299071
6. NASEM 2011. National Academies of Sciences, Engineering, and Medicine. 2011. **Long-Term Health Consequences of Exposure to Burn Pits in Iraq and Afghanistan**. Washington, DC: The National Academies Press. https://doi.org/10.17226/13209
7. NASEM 2017. National Academies of Sciences, Engineering, and Medicine. 2017. **Assessment of the Department of Veterans Affairs Airborne Hazards and Open Burn Pit Registry**. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/23677
8. NASEM 2020. National Academies of Sciences, Engineering, and Medicine. 2020. **Respiratory Health Effects of Airborne Hazards Exposures in the Southwest Asia Theater of Military Operations**. Washington, DC: The National Academies Press. https://doi.org/10.17226/25837.
9. NASEM 2022. National Academies of Sciences, Engineering, and Medicine. 2022. **Reassessment of the Department of Veterans Affairs Airborne Hazards and Open Burn Pit Registry.** Washington, DC: The National Academies Press. https://doi.org/10.17226/26729
10. Anderson RH, Thompson T, Stroo HF, Leeson A. **US Department of Defense-Funded Fate and Transport Research on Per- and Polyfluoroalkyl Substances at Aqueous Film-Forming Foam-Impacted Sites**. Environ Toxicol Chem. 2021 Jan;40(1):37-43. doi: 10.1002/etc.4694. Epub 2020 Jun 2. PMID: 32077141; PMCID: PMC7984261.
11. EPA NPDWR (PFAS) **National Primary Drinking Water Regulations: PFAS** https://www.epa.gov/sdwa/and-polyfluoroalkyl-substances-pfas
12. Katherine A. Phillips, Alex Chao, Rebecca L. Church, Kristin Favela, Stavros Garantziotis, Kristin K. Isaacs, Brian Meyer, Annette Rice, Risa Sayre, Barbara A. Wetmore, Alice Yau and John F. Wambaugh. **Suspect Screening Analysis of Pooled Human Serum Samples Using GC × GC/TOF-MS**. Environ Science and Tech. 2024 Jan;58(4):1802-1812
13. David Rogers and Mathew Hahn **Extended-Connectivity Fingerprints**, *Journal of Chemical Information and Modeling* 2010 50(5):742-754
14. David Bajusz, Anita Racz, Karoly Heberger **Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?** J Cheminform 2015 20:7
15. J. Chem. Inf. Comput. Sci., 1999, 39, 747

Design and Formulation of an **EXPOSURE COMPOUND LIBRARY**

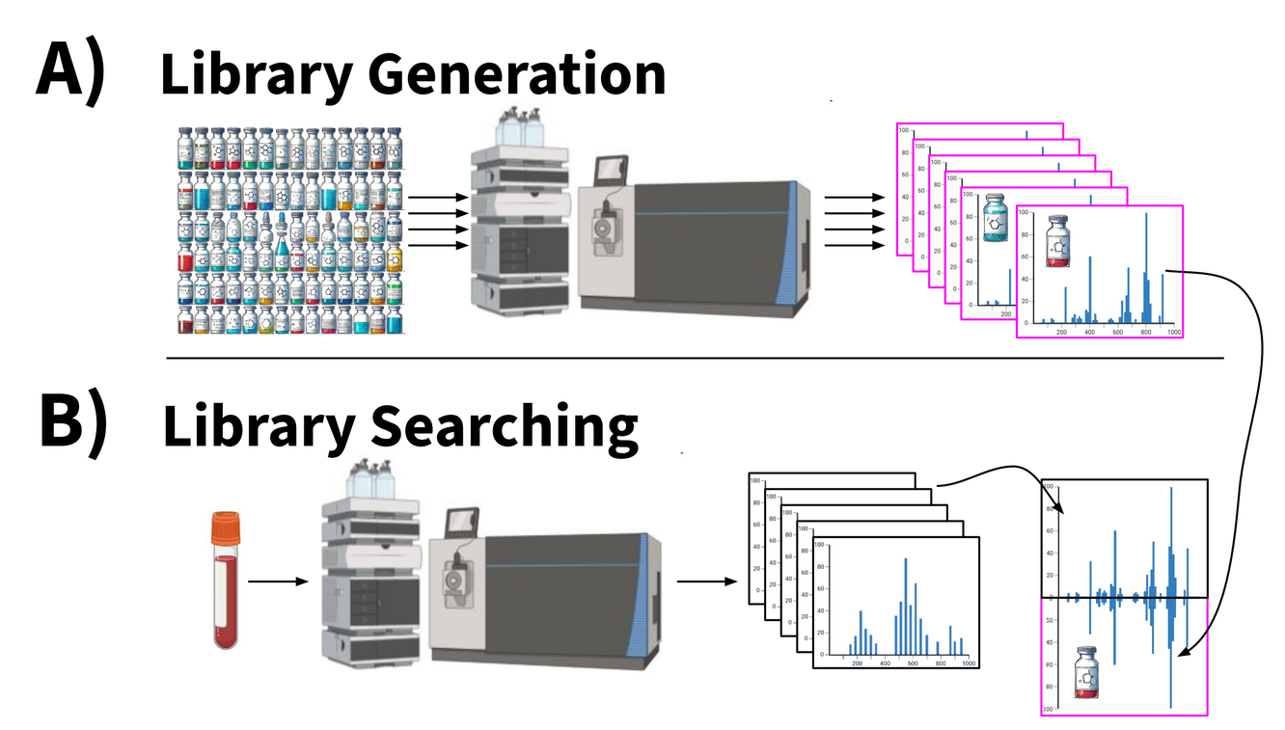


Figure 1: Mass Spectral Compound Libraries. Mass spectral libraries allow for high-confidence identifications of specific compounds in unknown samples. This is a two part process. In the initial process (library building), shown in **A**, collections of known compounds are run through the mass spectrometer and their retention times (RT), mass to charge (m/z) and fragmentation spectra (ms2) are collected (shown with a pink border). Compounds will be ionized with liquid chromatography coupled to combination of Electrospray Ionization (ESI) and Atmospheric pressure photoionization ionization (APPI) with follow up gas chromatography coupled to Electron Impact (EI) ionization as required. Subsequently, when unknown samples are measured in the mass spectrometer (**B**), the RT, m/z and ms2 spectra are compared against the library and matches are used to confirm the presence of these compounds.

CASRN_Tar	Name_Tar	CASRN_Tar	Name_Tar	CASRN_Tar	Name_Tar	CASRN_Tar	Name_Tar	CASRN_Tar	Name_Tar	CASRN_Tar	Name_Tar
83-32-9	Acenaphthene	3165-93-3	4-Chloro-2-methylaniline hydrochloride	57-41-0	5,5-Diphenylhydantoin	73250-68-7	Mefenacet	116714-46-6	Novaluron	283594-90-1	Spiromesifen
208-96-8	Acenaphthylene	107-30-2	Chloromethyl methyl ether	630-93-3	Diphenylhydantoin sodium	135590-91-9	Mefenpyr-diethyl	39001-02-0	Octachlorodibenzofuran	203313-25-1	Spirotetramat
30560-19-1	Acephate	563-47-3	3-Chloro-2-methylpropene	122-66-7	1,2-Diphenylhydrazine	4376-20-9	MEHP	3268-87-9	Octachlorodibenzo-para-dioxin	118134-30-8	Spiroxamine
75-07-0	Acetaldehyde	15545-48-9	Chlorotoluron	330-54-1	Diuron	55814-41-0	Mepronil	1806-26-4	4-Octylphenol	18883-66-4	Streptozotocin
135158-54-2	Acibenzolar-S-methyl	218-01-9	Chrysene	155569-91-8	Emamectin benzoate	57837-19-1	Metalaxyl	101-80-4	4,4'-Oxydianiline	100-42-5	Styrene
107-02-8	Acrolein	569-61-9	C.I. Basic Red 9 monohydrochloride	100-41-4	Ethylbenzene	125116-23-6	Metconazole	56-38-2	Parathion	96-09-3	Styrene oxide
79-06-1	Acrylamide	99129-21-2	Clethodim	96-45-7	Ethylene thiourea	18691-97-9	Methabenzthiazuron	57117-41-6	1,2,3,7,8-Pentachlorodibenzofuran	80-09-1	4,4'-Sulfonyldiphenol
107-13-1	Acrylonitrile	210880-92-5	Clothianidin	21245-02-3	2-Ethylhexyl 4-(dimethylamino)benzoate	10265-92-6	Methamidophos	57117-31-4	2,3,4,7,8-Pentachlorodibenzofuran	10540-29-1	Tamoxifen
25316-40-9	Adriamycin hydrochloride	26444-49-5	Cresyl diphenyl phosphate	759-73-9	1-Ethyl-1-nitrosourea	16752-77-5	Methomyl	40321-76-4	1,2,3,7,8-Pentachlorodibenzo-p-dioxin (PCDD)	107534-96-3	Tebuconazole
834-12-8	Ametryn	135-20-6	Cupferron	120-47-8	Ethylparaben	134-29-2	2-Methoxyaniline hydrochloride	87-86-5	Pentachlorophenol	112410-23-8	Tebufenozide
117-79-3	2-Aminoanthraquinone	120116-88-3	Cyazofamid	161326-34-7	Fenamidone	120-71-8	2-Methoxy-5-methylaniline	3296-90-0	Pentaerythritol dibromide	34014-18-1	Tebuthiuron
97-56-3	2-Amino-5-azotoluene	50-18-0	Cyclophosphamide	60168-88-9	Fenarimol	91-23-6	1-Methoxy-2-nitrobenzene	375-73-5	Perfluorobutanesulfonic acid	83121-18-0	Teflubenzuron
82-28-0	1-Amino-2-methylanthraquinone	59865-13-3	Cyclosporin A	114369-43-6	Fenbuconazole	298-81-7	8-Methoxypsoralen	375-22-4	Perfluorobutanoic acid	79-94-7	3,3',5,5'-Tetrabromobisphenol A
61-82-5	Amitrole	57966-95-7	Cymoxanil	67564-91-4	(2R,6S)-Fenpropimorph	95-53-4	2-Methylaniline	335-76-2	Perfluorodecanoic acid	51207-31-9	2,3,7,8-Tetrachlorodibenzofuran
90-04-0	2-Anisidine	66215-27-8	Cyromazine	101-42-8	Fenuron	101-61-1	4,4'-Methylenebis(N,N-dimethylaniline)	375-85-9	Perfluoroheptanoic acid	1746-01-6	2,3,7,8-Tetrachlorodibenzo-p-dioxin
120-12-7	Anthracene	134-62-3	DEET	142459-58-3	Flufenacet	93-15-2	Methyleugenol	355-46-4	Perfluorohexanesulfonic acid	127-18-4	Tetrachloroethylene
320-67-2	5-Azacytidine	101-77-9	4,4'-Diaminobiphenyl methane	206-44-0	Fluoranthene	66-27-3	Methyl methanesulfonate	307-24-4	Perfluorohexanoic acid	22248-79-9	Tetrachlorvinphos
446-86-6	Azathioprine	95-80-7	2,4-Diaminotoluene	86-73-7	Fluorene	99-76-3	Methylparaben	375-95-1	Perfluorononanoic acid	112281-77-3	Tetraconazole
71626-11-4	Benalaxyl	333-41-5	Diazinon	77279-89-1	4-Fluoro-3-phenoxybenzoic acid	3060-89-7	Metobromuron	1763-23-1	Perfluorooctane sulfonic acid	509-14-8	Tetranitromethane
82560-54-1	Benfuracarb	136-35-6	Diazoaminobenzene	248-580-6	6:2 fluorotelomer sulfonate (62FTS)	21087-64-9	Metribuzin	335-67-1	Perfluorooctanoic acid	148-79-8	Thiabendazole
56-55-3	Benz(a)anthracene	53-70-3	Dibenz(a,h)anthracene	85509-19-9	Flusilazole	443-48-1	Metronidazole	2706-90-3	Perfluoropentanoic acid	111988-49-9	Thiacloprid
71-43-2	Benzene	96-12-8	1,2-Dibromo-3-chloropropane	76674-21-0	Flutriafol	90-94-8	Michler's ketone	2058-94-8	Perfluoroundecanoic acid	51707-55-2	Thidiazuron
92-87-5	Benzidine	106-93-4	1,2-Dibromoethane	173159-57-4	Foramsulfuron	2385-85-5	Mirex	62-44-2	Phenacetin	62-55-5	Thioacetamide
50-32-8	Benzo[a]pyrene	84-74-2	Dibutyl phthalate	23422-53-9	Formetanate hydrochloride	2528-16-7	Monobenzyl phthalate	85-01-8	Phenanthrene	139-65-1	4,4'-Thiodianiline
205-99-2	Benzo(b)fluoranthene	106-46-7	1,4-Dichlorobenzene	3878-19-1	Fuberidazole	131-70-4	Monobutyl phthalate	136-40-3	Phenazopyridine hydrochloride	52-24-4	Thiotepa
191-24-2	Benzo[g,h,i]perylene	612-83-9	3,3'-Dichlorobenzidine dihydrochloride	556-52-5	Glycidol	40809-41-4	Mono(5-carboxy-2-ethylpentyl) Phthalate	77-09-8	Phenolphthalein	62-56-6	Thiourea
207-08-9	Benzo(k)fluoranthene	50-29-3	Dichlorodiphenyltrichloroethane	112226-61-6	Halofenozide	936022-02-5	Mono-(7-carboxy-4-methylheptyl) Phthalate	63-92-3	Phenoxybenzamine hydrochloride	584-84-9	Toluene-2,4-diisocyanate
98-07-7	Benzotrichloride	107-06-2	1,2-Dichloroethane	67562-39-4	1,2,3,4,6,7,8-Heptachlorodibenzo[b,d]furan	40321-98-0	Mono(2-ethyl-5-oxohexyl)phthalate	51-03-6	Piperonyl butoxide	91-08-7	Toluene-2,6-diisocyanate
29104-30-1	Benzoximate	75-09-2	Dichloromethane	35822-46-9	1,2,3,4,6,7,8-Heptachlorodibenzodioxin	2306-33-4	Monoethyl phthalate	23103-98-2	Pirimicarb	55219-65-3	Triadimenol
94-18-8	Benzylparaben	94-75-7	2,4-Dichlorophenoxyacetic acid	55673-89-7	1,2,3,4,7,8,9-Heptachlorodibenzofuran	1746-81-2	Monolinuron	24634-61-5	Potassium (E,E)-hexa-2,4-dienoate	52-68-6	Trichlorfon
57-57-8	beta-Propiolactone	542-75-6	1,3-Dichloropropene	72918-21-9	1,2,3,7,8,9-Hexachlorodibenzo[b,d]furan	91-20-3	Naphthalene	671-16-9	Procarbazine	95-95-4	2,4,5-Trichlorophenol
1464-53-5	2,2'-Bioxirane	15165-67-0	Dichlorprop-P	60851-34-5	2,3,4,6,7,8-Hexachlorodibenzo[b,d]furan	91-59-8	2-Naphthylamine	366-70-1	Procarbazine hydrochloride	88-06-2	2,4,6-Trichlorophenol
188425-85-6	Boscalid	62-73-7	Dichlorvos	70648-26-9	1,2,3,4,7,8-Hexachlorodibenzofuran	555-37-3	Neburon	67747-09-5	Prochloraz	93-76-5	2,4,5-Trichlorophenoxyacetic acid
41483-43-6	Bupirimate	87130-20-9	Diethofencarb	57117-44-9	1,2,3,6,7,8-Hexachlorodibenzofuran	139-13-9	Nitrilotriacetic acid	57-83-0	Progesterone	96-18-4	1,2,3-Trichloropropane
69327-76-0	Buprofezin	117-81-7	Di(2-ethylhexyl) phthalate	57653-85-7	1,2,3,6,7,8-Hexachlorodibenzo-p-dioxin	98-95-3	Nitrobenzene	1610-18-0	Prometon	101-20-2	Triclocarban
55-98-1	Busulfan	56-53-1	Diethylstilbestrol	19408-74-3	1,2,3,7,8,9-Hexachlorodibenzo-p-dioxin	1836-75-5	Nitrofen	7287-19-6	Prometryn	78-40-0	Triethyl phosphate
106-99-0	1,3-Butadiene	64-67-5	Diethyl sulfate	79983-71-4	Hexaconazol	75-52-5	Nitromethane	24579-73-5	Propamocarb	68694-11-1	Triflumizole
25013-16-5	Butylated hydroxyanisole	84-69-5	Diisobutyl phthalate	680-31-9	Hexamethylphosphoramide	100-02-7	4-Nitrophenol	1120-71-4	1,3-Propane sultone	126-71-6	Triisobutyl phosphate
2425-06-1	Captafol	28553-12-0	Diisononyl phthalate	10034-93-2	Hydrazine sulfate	79-46-9	2-Nitropropane	114-26-1	Propoxur	512-56-1	Trimethyl phosphate
10605-21-7	Carbendazim	110488-70-5	Dimethomorph	99-96-7	4-Hydroxybenzoic acid	88-72-2	2-Nitrotoluene	51-52-5	6-Propyl-2-thiouracil	68937-41-7	Triphenyl phosphates isopropylated
1563-66-2	Carbofuran	119-90-4	3,3'-Dimethoxybenzidine	138261-41-3	Imidacloprid	64091-91-4	4-(N-Methyl-N-nitrosamino)-1-(3-pyridyl)-1-butanone	123312-89-0	Pymetrozine	513-08-6	Tripropyl phosphate
5234-68-4	Carboxine	119-93-7	3,3'-Dimethylbenzidine	193-39-5	Indeno(1,2,3-cd)pyrene	924-16-3	N-Nitrosodibutylamine	129-00-0	Pyrene	78-51-3	Tris(2-butoxyethyl) phosphate
154-93-8	Carmustine	79-97-0	3,3'-Dimethylbisphenol A	2631-40-5	Isoprocarb	55-18-5	N-Nitrosodiethylamine	96489-71-3	Pyridaben	115-96-8	Tris(2-chloroethyl) phosphate
305-03-3	Chlorambucil	79-44-7	Dimethylcarbamoyl chloride	34123-59-6	Isoproturon	62-75-9	N-Nitrosodimethylamine	53112-28-0	Pyrimethanil	126-72-7	Tris(2,3-dibromopropyl) phosphate
56-75-7	Chloramphenicol	57-14-7	1,1-Dimethylhydrazine	143-50-0	Kepone	621-64-7	N-Nitrosodipropylamine	50-55-5	Reserpine	78-32-0	Tris(4-methylphenyl) phosphate
115-28-6	Chlorendic acid	77-78-1	Dimethyl sulfate	143390-89-0	Kresoxim-methyl	59-89-2	N-Nitrosomorpholine	83-79-4	Rotenone	64532-95-2	Tris[2-(propan-2-yl)phenyl] phosphate
74-11-3	4-Chlorobenzoic acid	513-37-1	2,2-Dimethylvinyl chloride	58-89-9	Lindane	684-93-5	N-Nitroso-N-methylurea	94-59-7	Safrole	131983-72-7	Triticonazole
95-83-0	4-Chloro-1,2-diaminobenzene	117-84-0	Di-n-octyl phthalate	330-55-2	Linuron	930-55-2	N-Nitrosopyrrolidine	13909-09-6	Semustine	51-79-6	Urethane
74-87-3	Chloromethane	88-85-7	Dinoseb	13010-47-4	Lomustine	104-40-5	4-Nonylphenol	1982-49-6	Siduron	106-87-6	4-Vinyl-1-cyclohexene dioxide
95-69-2	4-Chloro-2-methylaniline	165252-70-0	Dinotefuran	121-75-5	Malathion	68-22-4	Norethindrone	131-52-2	Sodium pentachlorophenate	2425-10-7	Xylylcarb
										156052-68-5	Zoxamide

Table 1: Target Exposure Compounds. Listed are the CAS numbers and the compound names for the 301 compounds we have identified as targets for our Exposure Compound Library. Please see the methods for methodology of selection of theses compounds.

Design and Formulation of a **LONG TERM REFERENCE SAMPLE**

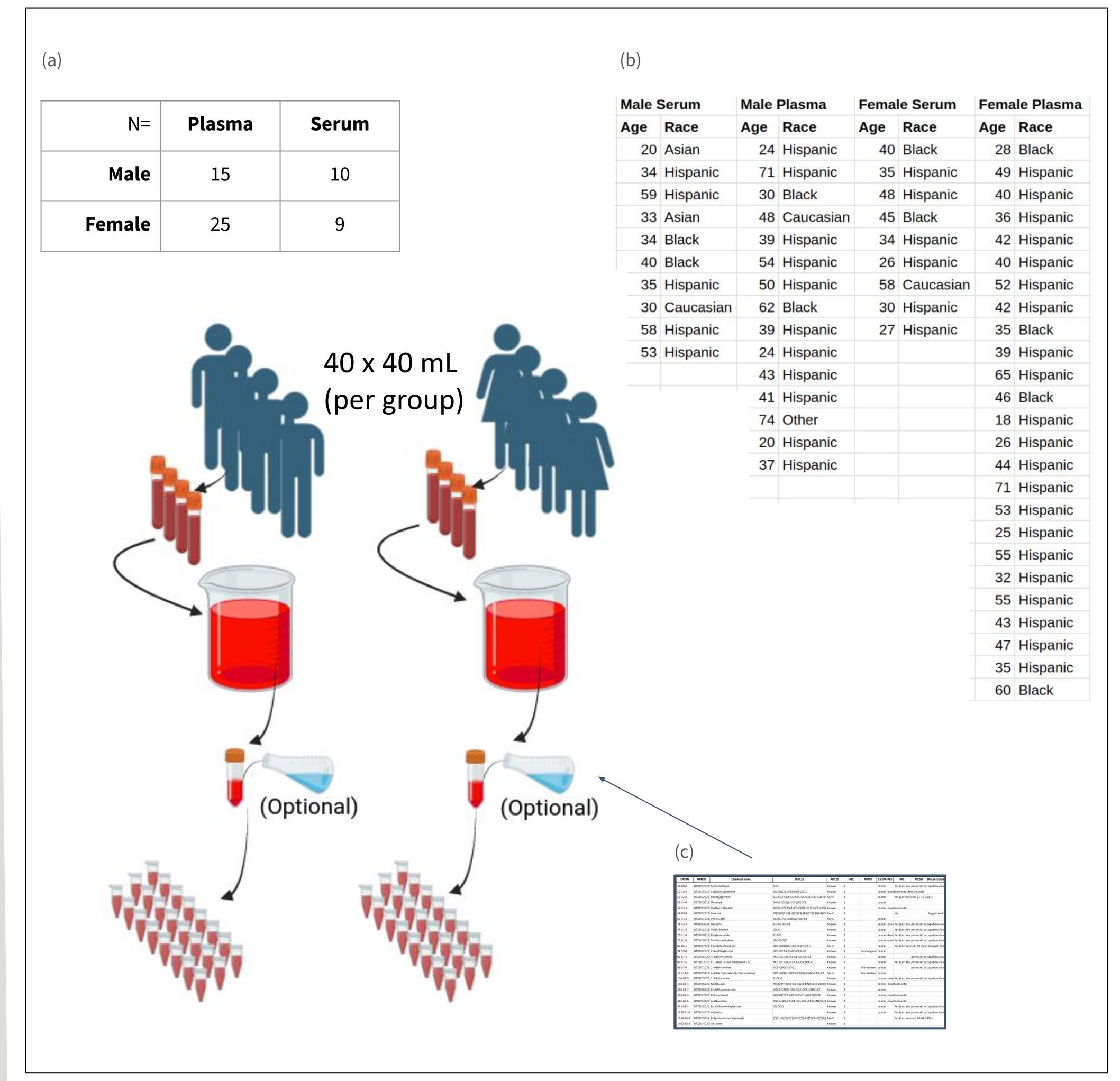


Figure 2: Long Term Reference Sample (LTRS). Our LTRS was acquired to support efforts in PROMETHEUS and for the MCCRP. There are 4 separate pools for plasma and serum, male and female acquired from BioIVT. Each 1.6 L pool was comprised of N individuals as shown in (a) with demographics as shown in (b). Below is discussed our plan for an EXPOSURE INTERNAL REFERENCE, inset (c) demonstrates how can can be used to dope the LTRS as it is aliquoted.

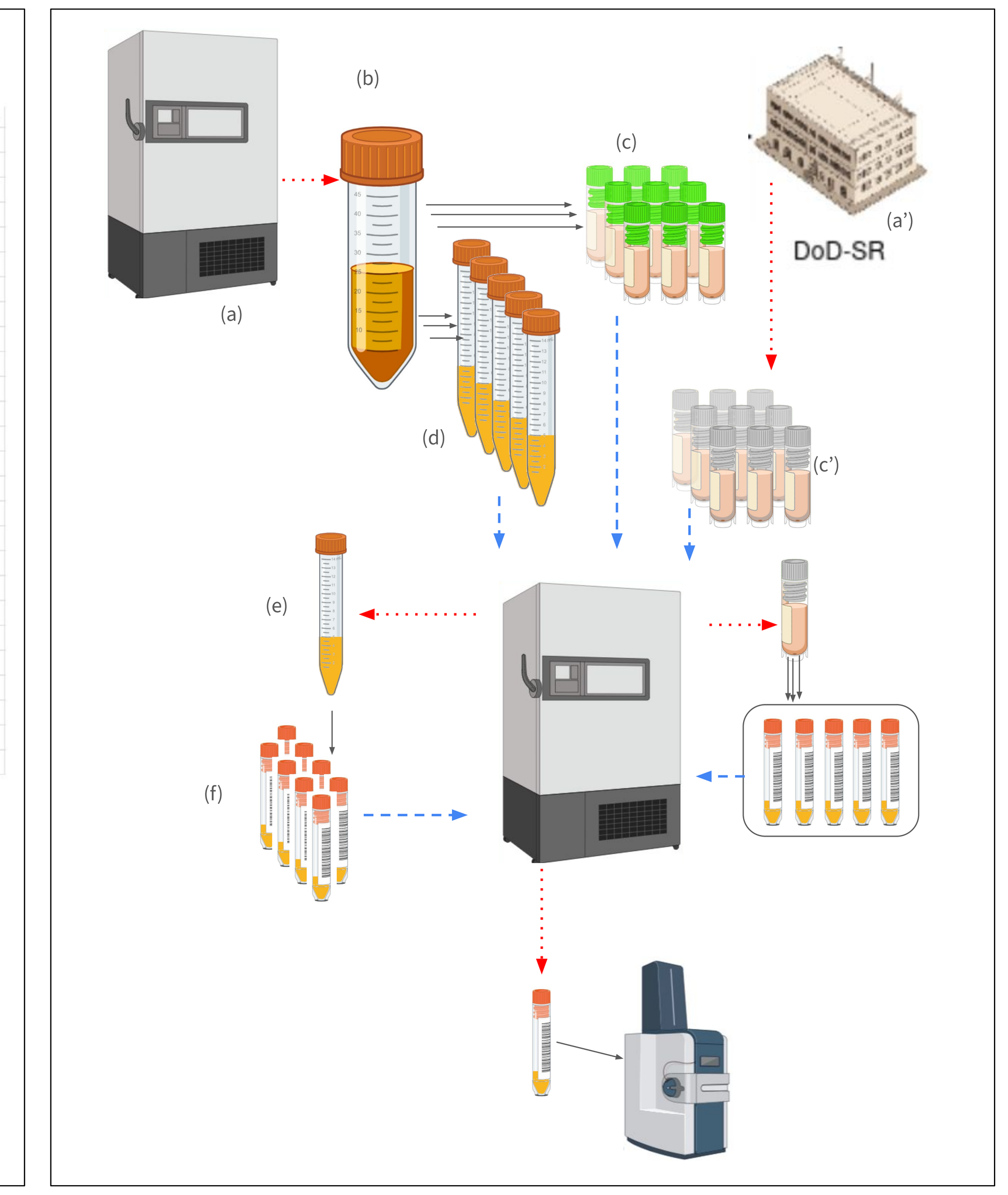


Figure 3: LTRS Aliquoting Schedule. Aliquoting schedule for the LTRS is designed to match handling of real samples. All thaw steps are red and dotted, all freeze steps are in blue and dashed and all transfer steps are in black and solid. Samples from DodSR come in 2 mL cryotubes (c, c'). All our samples are analyzed from microplate format tubes (f). We designed our handling of LTRS material (b) for convenient volumes (c,d) and to match the thaw count of "real" samples - we do not go directly from (b) or (d) to (f) without a freeze/thaw producing (e)

Design and Formulation of an **EXPOSURE INTERNAL STANDARD**

Acrylamide	4-Fluoro-3-phenoxy-benzoic acid	L-Palmitoylcarnitine
Adenosine	6:2-fluorotelomer sulfonate (62FTS)	1-palmitoyl-2-hydroxy-sn-glycero-3-phosphocholine (Lyso-PC (16:0))
Adenosine-5'-triphosphate (ATP)	Fructosamine	Paraben (Methylparaben)
L-Alanine	gamma-Caprolactone	2,3,3',4,4'-Pentachlorobiphenyl (PCB 105)
Ametryn	Glucose	1,2,3,7,8-Pentachlorodibenzofuran (PeCDF)
Anthracene	L-Glutamic Acid	2,3,4,7,8-Pentachlorodibenzofuran (PeCDF)
Anthracene	D-Glyceraldehyde	1,2,3,7,8-Pentachlorodibenzo-p-dioxin (PCDD)
L-Arginine	1,2,3,4,6,7,8-Heptachlorodibenzofuran (HpCDF)	perfluorobutanesulfonic acid (PFBS)
L-Aspartic Acid	1,2,3,4,7,8,9-Heptachlorodibenzofuran (HpCDF)	perfluorobutanoic acid (PFBA)
Benz[a]anthracene	1,2,3,4,6,7,8-Heptachlorodibenzo-p-dioxin (HpCDD)	Perfluorodecanoic acid (PFDA)
Benzene	1,2,3,4,7,8-Hexachlorodibenzofuran (HxCDF)	perfluoroheptanoic acid (PFHpA)
Benzo[a]pyrene (BaP)	1,2,3,6,7,8-Hexachlorodibenzofuran (HxCDF)	Perfluorohexanesulfonic acid (PFHxS)
Benzo[b]fluoranthene	1,2,3,7,8,9-Hexachlorodibenzofuran (HxCDF)	perfluorohexanoic acid (PFHxA)
Benzo(ghi)perylene	2,3,4,6,7,8-Hexachlorodibenzofuran (HxCDF)	Perfluorononanoic acid (PFNA)
Benzoximate	1,2,3,6,7,8-Hexachlorodibenzo-p-dioxin (HxCDD)	Perfluorooctanesulfonic acid (PFOS)
Benzyl paraben	1,2,3,7,8,9-Hexachlorodibenzo-p-dioxin (HxCDD)	Perfluorooctanoic acid (PFOA)
beta-Ionone	Hexadecane	perfluoropentanoic acid (PFPeA)
Bisphenol A (BPA)	Hexafluoropropylene Oxide Dimer Acid (HFPO DA)	Perfluoroundecanoic acid (PFUnDA /PFUnA)
Boscalid	3-hydroxybutyric acid (BHB)	Phenanthrene
1,3-Butadiene	Hypoxanthine	Pirimicarb
Caffeine	Ibuprofen	Prometon/Secbumeton/Terbumeton
Carboxin	(Iso)propyl paraben	Pyrene
L-Carnitine	L-Kynurenine	Rotenone
4-Chlorobenzoic acid	Lactic acid	S-Phenylmercapturic acid (S-PMA) (S-Phenyl-N-acetylcysteine)
Chloromethane	L-Leucine	Sphingosine
Chlorotoluron	Linoleic Acid	Styrene
Cholesterol	L-Lysine	Taurine
Choline	Mefenacet	Tebuthiuron
Chrysene	Melanin	Teflubenzuron
Citric Acid	Methylene chloride	2,3,7,8-Tetrachlorodibenzofuran (TCDF)
Cotinine	Methyl salicylate	2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD)
Cotinine	Methyl (9Z)-octadec-9-enoate	Tetrachloroethene (PCE)
Creatinine	Mono-(2-ethyl-5-oxohexyl) phthalate (MEOHP)	Tetraconazole
Dibenz[a,h]anthracene	Mono-n-butyl phthalate (MnBP)	α-Tocopherol (Vitamin E)
Diclofenac	Muconic acid (trans,trans)	Trichlorfon (metrifonate)
Di(2-ethylhexyl) phthalate (DEHP)	Naphthalene	Triclosan
Diisobutyl phthalate	Naphthalene	Trimethyl Phosphate
Dinoseb	Neburon	tris(2-Chloroethyl) Phosphate (TCEP)
Dinotefuran	Nicotine	tris(2-Isopropylphenyl) Phosphate
Diuron	N,N-Diethyl-meta-toluamide (DEET)	Urea
Docosahexaenoic Acid (DHA; Cervonic Acid)	Nonanoic acid	Warfarin
Dodecylbenzene	Nonylphenol (suggested correction: 4-NONYLPHENOL)	Xylylcarb

Table 2: Target Exposure Internal Standards Compounds. Listed are the compound names for the short list (N=131) of compounds curated for for potential inclusion in an exposure internal standard.

For Poster Reprints and Spreadsheets of Compound Lists.

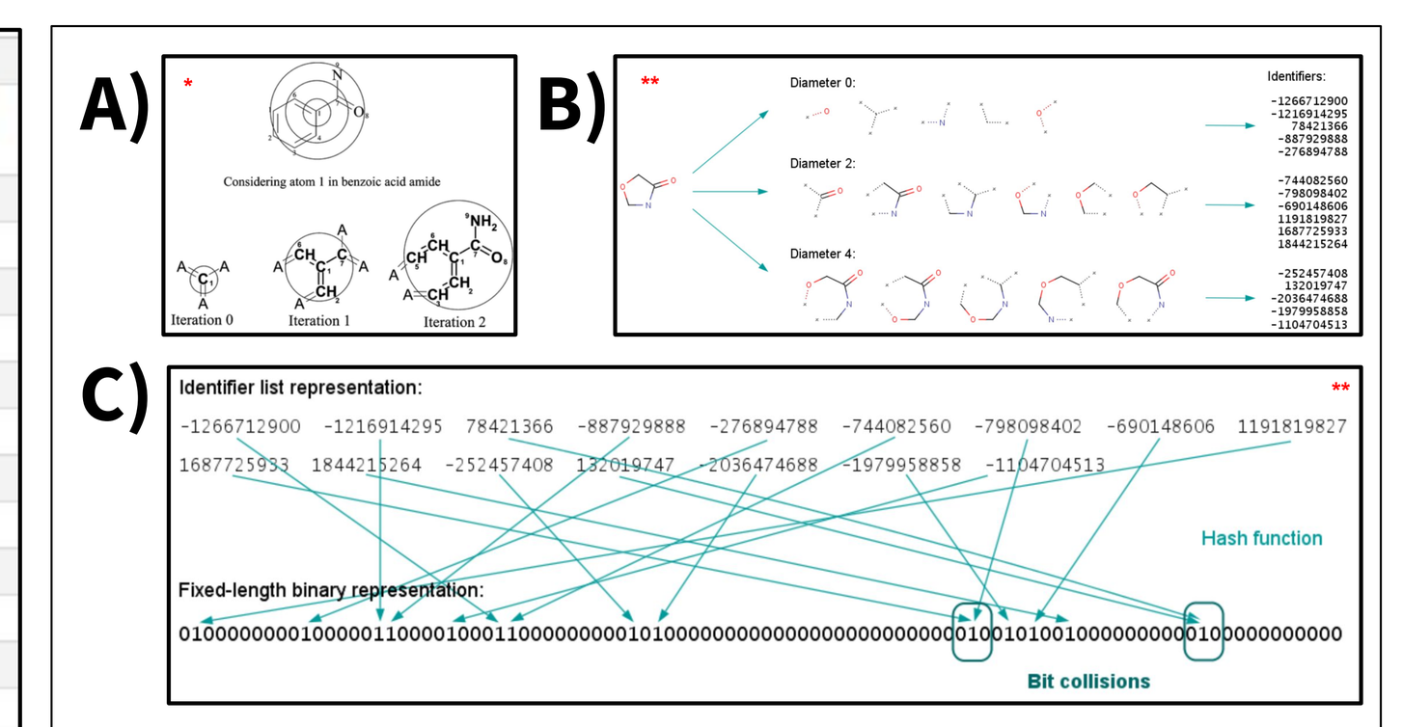


Figure 4: Extended Connectivity Fingerprinting (ECFP). Each atom in a compound was iteratively accessed at a diameter of 4 neighbors (**A**) to produce a unique binary representation representing that structure (**B**). The totality of representations were collapsed into a 2048 bit space creating the final hash for that molecule (or "fingerprint")

* J. Chem. Inf. Model. 2010, 50, 5, 742-754
** https://docs.chemaxon.com/display/docs/fingerprints_extended-connectivity-fingerprint-ecfp.md

Figure 5: Visualization of Tanimoto Similarity. The Tanimoto similarity between ECFP's are calculated as the proportion of hits out of the total possible. False colored here light to dark corresponding to 0 to 1 and clustered according to similarity.

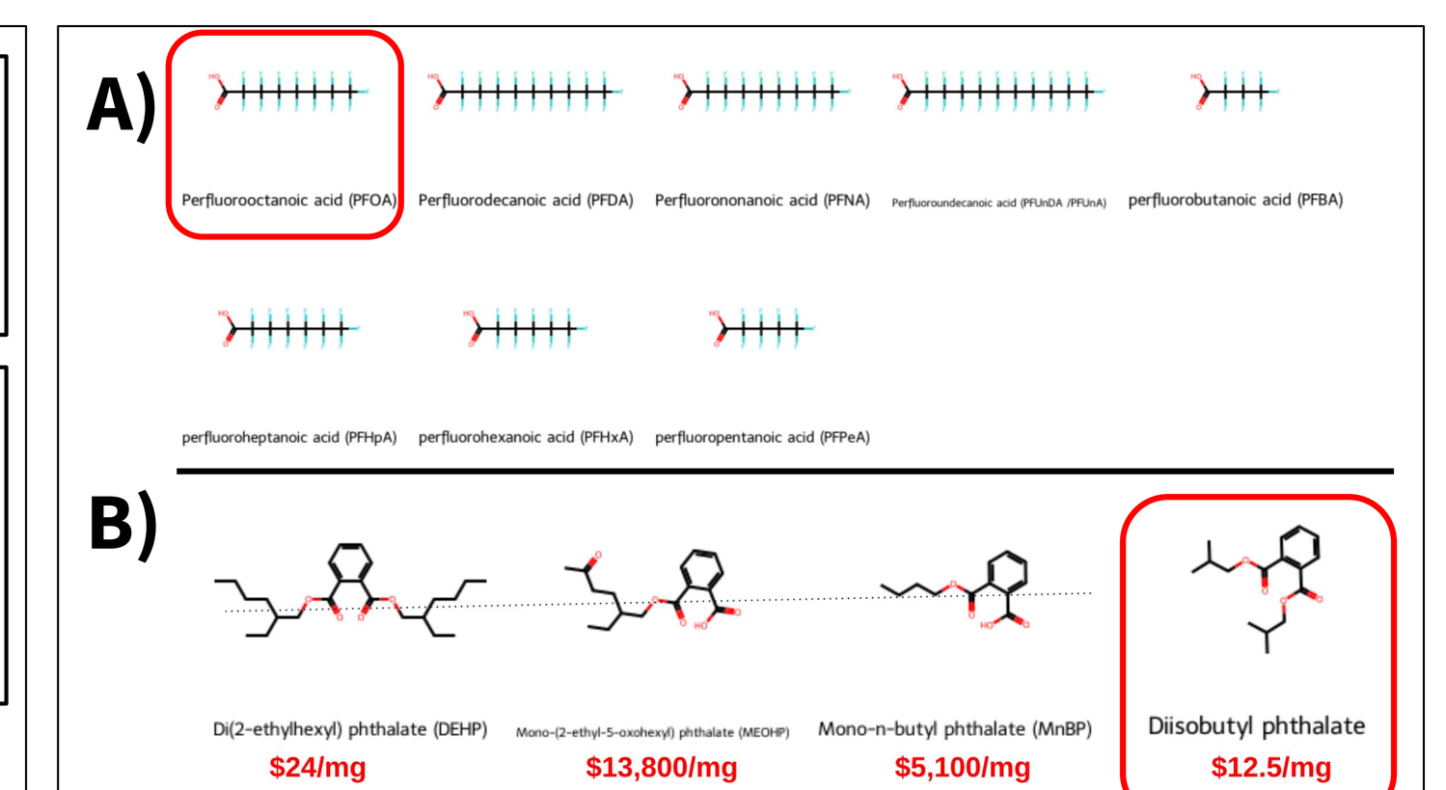


Figure 6: Set Reduction. Butina Clustering (0.5) created compound sets from which a representative unit from each was selected. In some cases (**A**) the most "iconic" compound was chosen. In other cases (**B**) availability or price guided choice. Sometimes multiple similar compounds were kept for biological rationale (e.g., nicotine, cotinine)

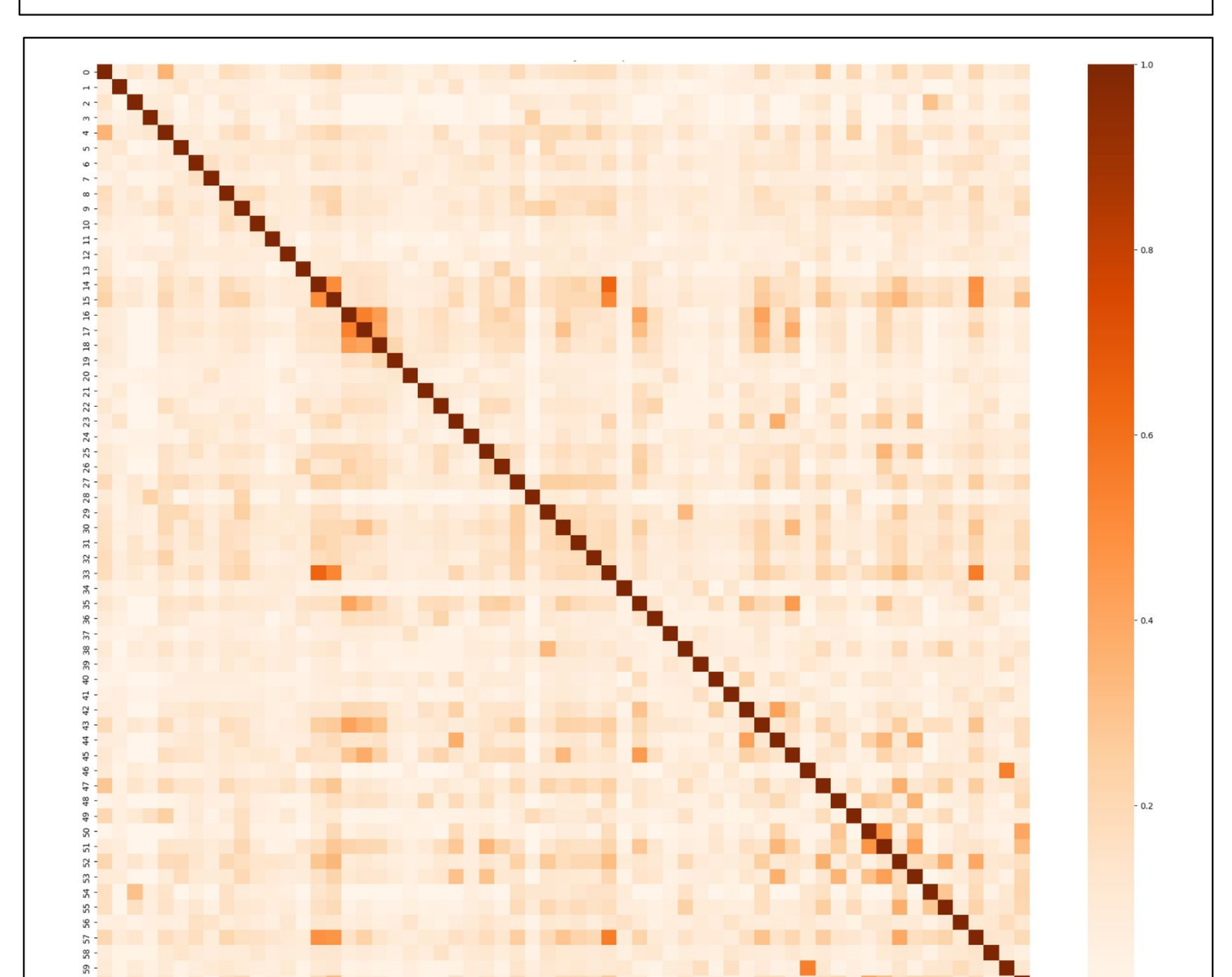
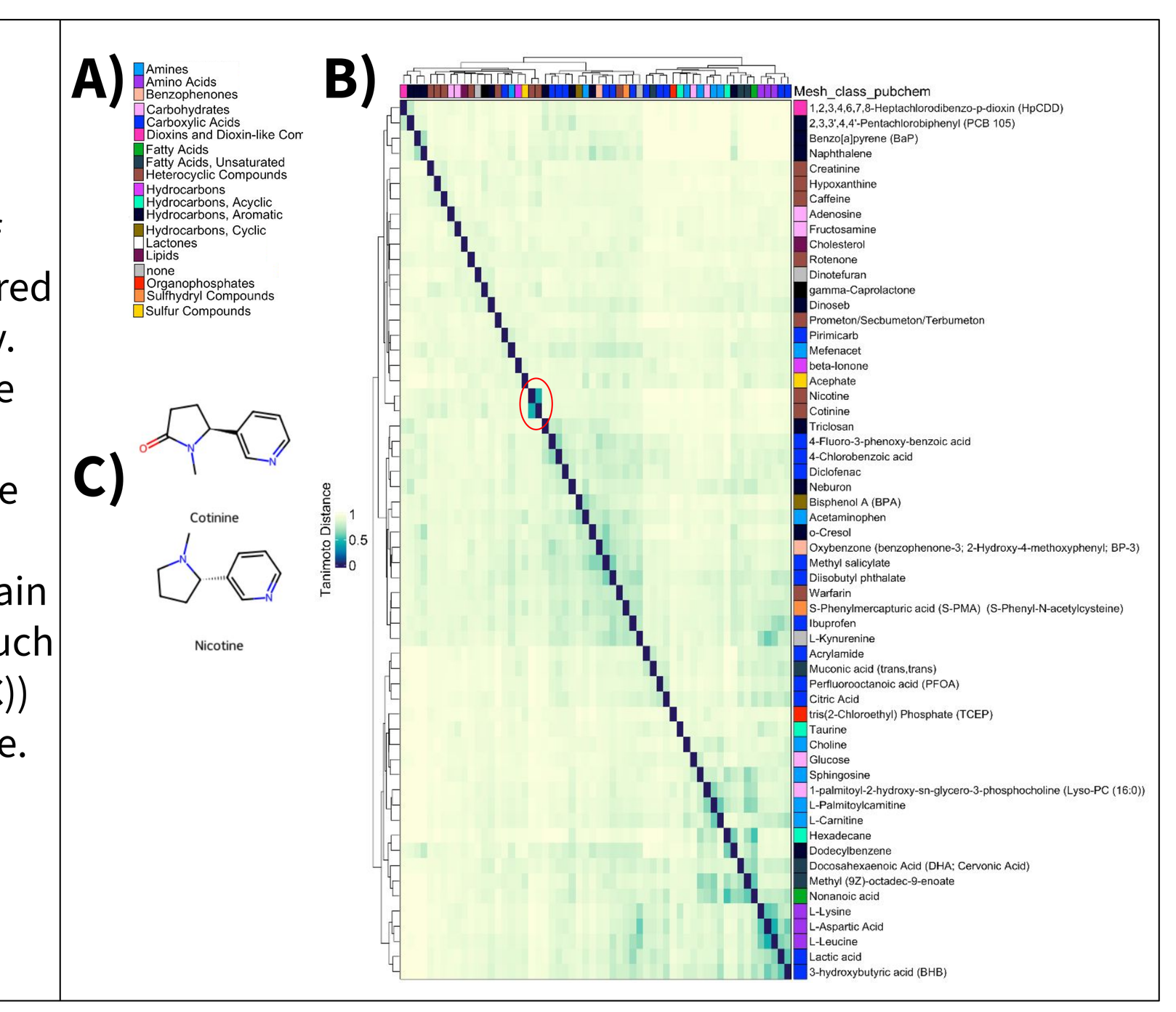


Figure 7: Visualization of Tanimoto Similarity - Post Reduction. Shown is the reduced similarity of our compound list following set reduction based on the strategy of set reduction shown in **Figure 6**.

Figure 8: Final set. Visualized, sorted by compound class (**A**) is our current final set of compounds (**B**) clustered by structural similarity. In this visualization are apparent some of the biological decisions we have made to keep rather than reject certain similar compounds (such as nicotine/cotinine (**C**)) shown by the red circle.



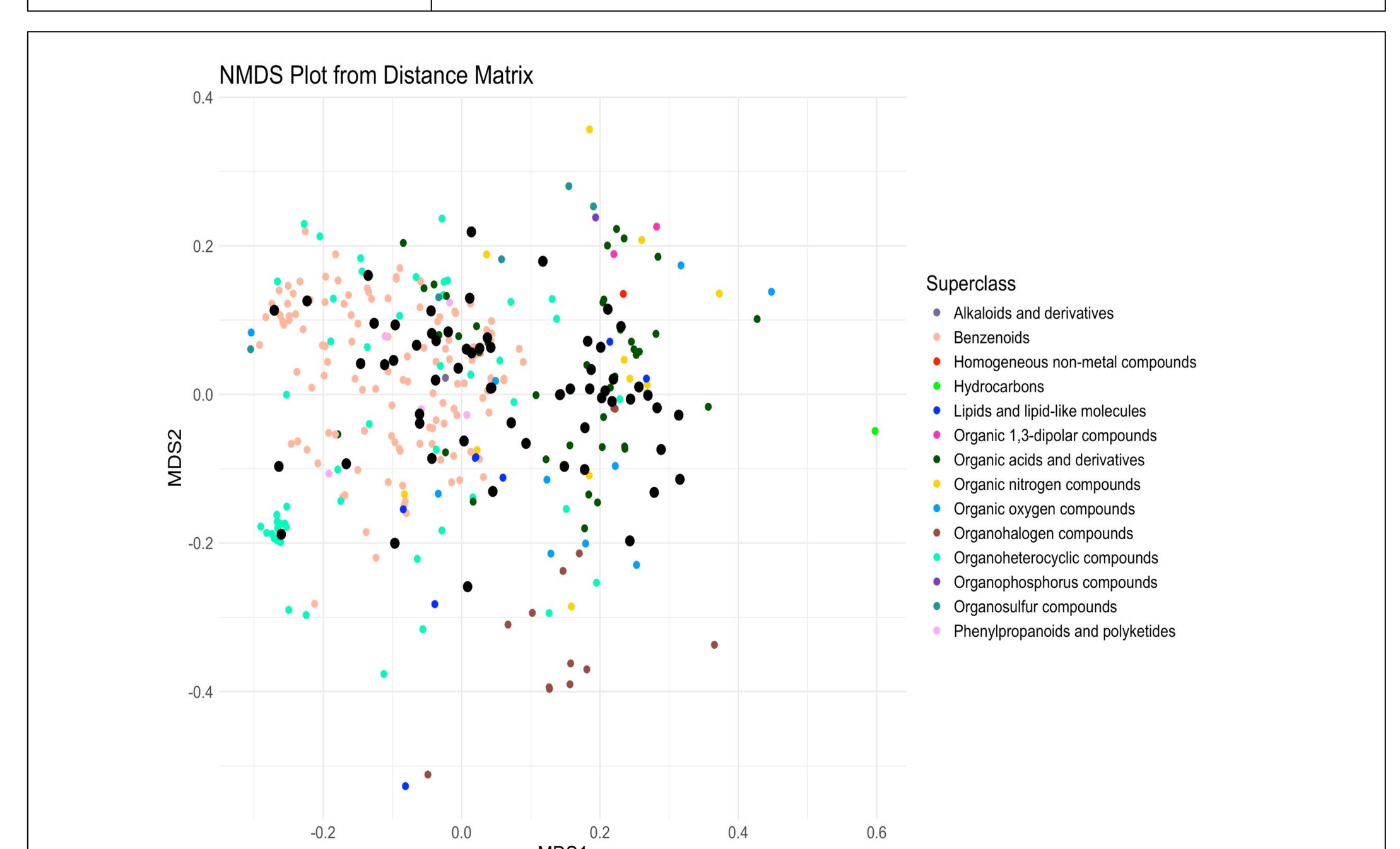


Figure 9: Visualized Compound Structure Coverage : Non-metric multidimensional scaling of paired Tanimoto distance scores for compounds in the chemical library (**Table 1**, shown by their Superclass) and exposure internal standards (**Figure 8**, in black) indicates that our selected internal standards cover the broad range of exposure compounds that our team has identified as highly relevant for human exposure studies.