Kristin Isaacs, Jon Sobus, Caroline Ring, John Wambaugh, and Katherine Phillips

U.S. EPA Office of Research and Development, Center for Computational Toxicology and Exposure, Research Triangle Park, NC

Kristin Isaacs (ORCID: 0000-0001-9547-1654) I Isaacs.kristin@epa.gov I 919-541-2785

Introduction

- New non-targeted analysis (NTA) technologies and supporting data analysis platforms hold great potential for identifying thousands of chemicals in environmental and biological samples.

- High-value applications of NTA include associating environmental chemical exposures with disease or health endpoints and screening priority environmental samples to identify unknown or uncharacterized exposure sources.

- Challenges in linking identified chemical substances with downstream endogenous biological pathways and upstream exogenous exposure sources must be addressed.

- There is a need to annotate exposure sources and pathways associated with chemicals tentatively identified or confirmed via NTA.

- The Office of Research and Development (ORD) of the U.S. Environmental Protection Agency (EPA) has been collecting and curating exposure-pathway relevant information including chemical use, release, and monitoring information.

- These data can be mined to develop hypotheses related to exposure sources for chemical features tentatively identified or confirmed in NTA experiments.

Approach

- An exposure pathway is the path a chemical takes from its source to a human or ecological receptor.

Determination of exposure pathway for a chemical identified in a sample using NTA can be informed by known information about its use and its occurrence in environmental and biological media.
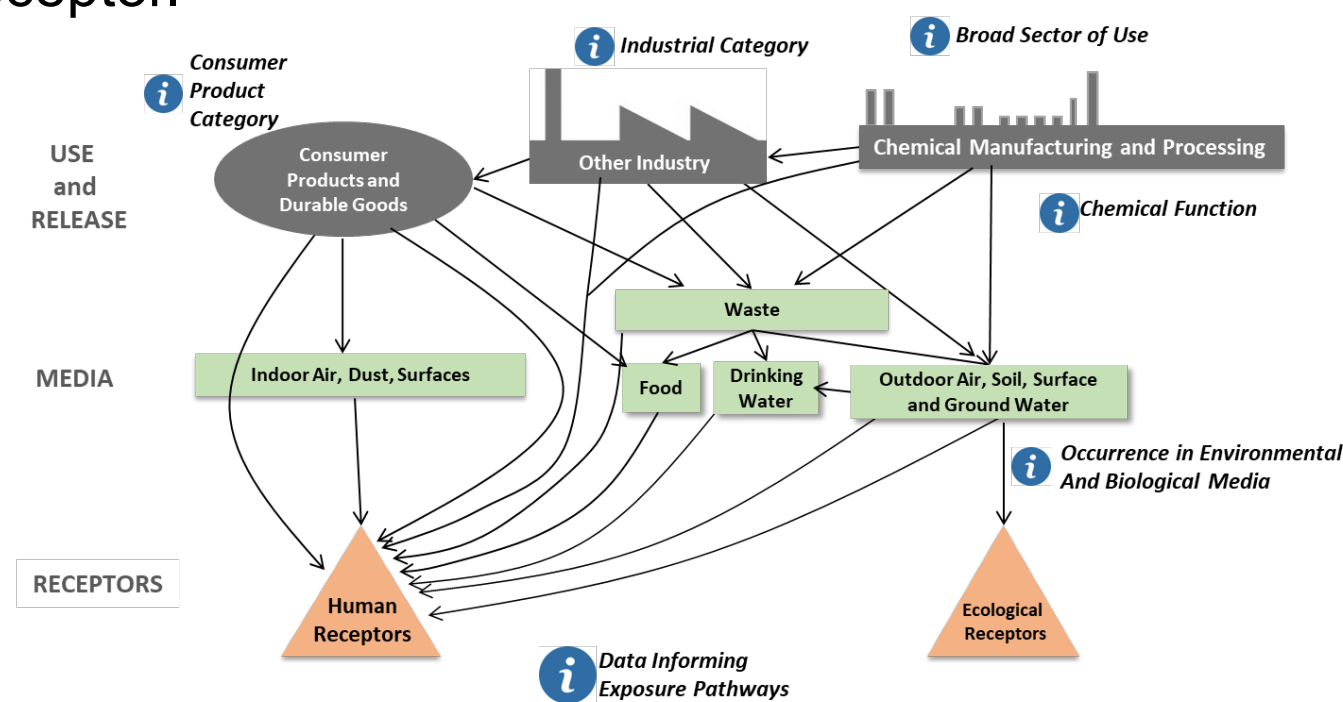

- A fingerprinting approach was developed that encodes available exposure pathway information from existing EPA databases as hierarchical binary descriptor sets.

- These descriptors can be queried for enrichment using lists of chemicals identified in NTA studies to provide evidence of exposure sources and pathways.

Methods

### ExpoPrint Fingerprints for Chemicals

- Exposure pathways fingerprints ("ExpoPrints") encoding information about chemical use and media occurrence were developed from existing EPA data sources.

- These descriptors represented occurrence of chemicals in broad commercial use sectors (e.g., consumer use), industrial categories (e.g., specific utilities or manufacturing types), consumer product categories (e.g., personal care, solvent), chemical functional use categories (e.g., solvent), and environmental media.

- For a given set of samples, the subset of exposure pathway components (uses, media) that are enriched compared to all chemicals with information can be determined (using a Fisher's exact test).

- ExpoPrints could be developed for 26,859 chemicals (DTXSIDs, which are unique chemical substance identifiers used by EPA's CompTox Chemicals Dashboard[1]).

A set of 475 binary descriptors ("ExpoPrints") encoding information about exposure pathways for chemicals was developed based on existing chemical use and environmental monitoring information, including EPA's Chemical and Products Database (CPDat),[2,3] Chemical Data Reporting (CDR),[4] and Multimedia Monitoring Database.[5]
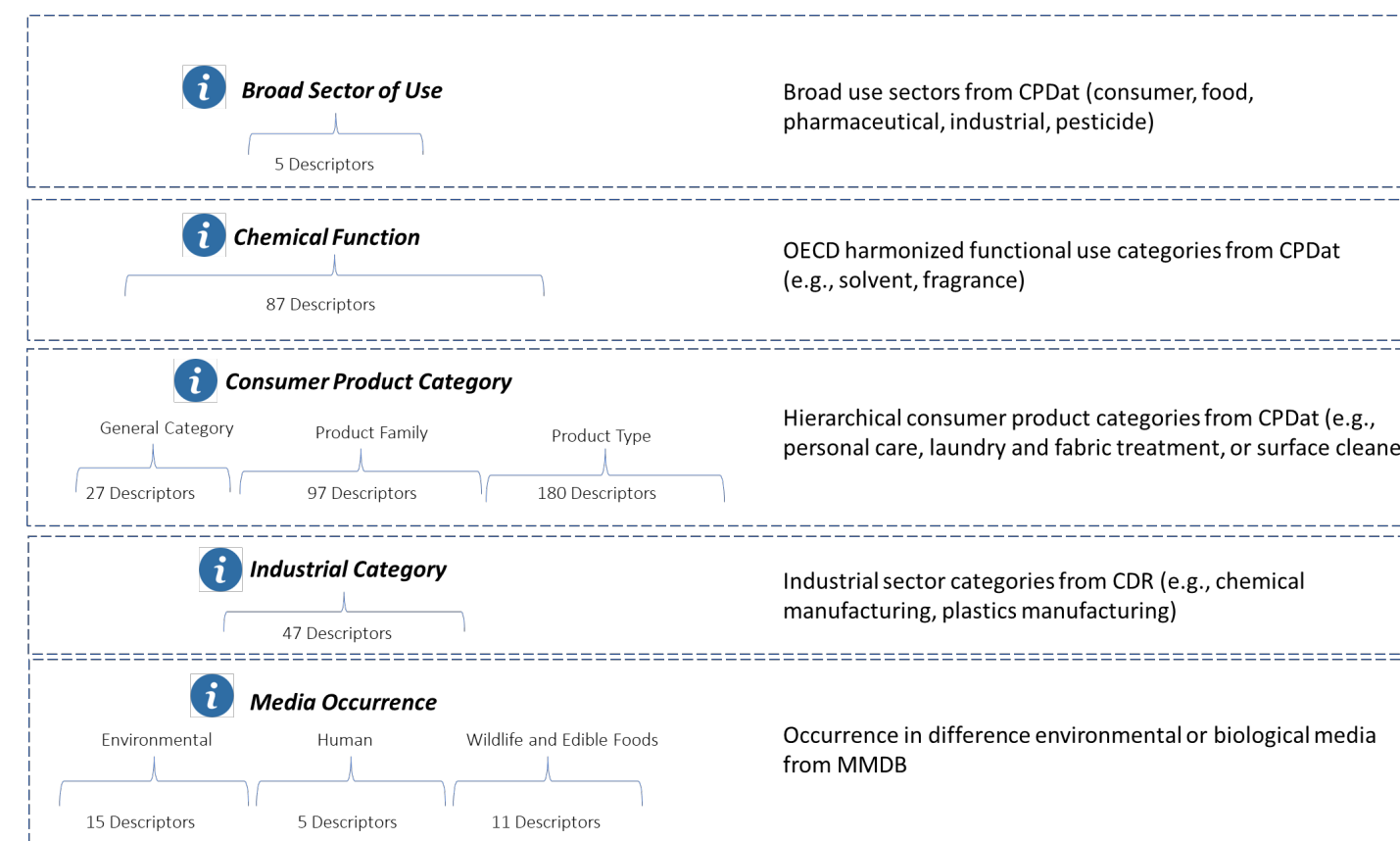
Visualizing Exposure Pathways as Networks

- ExpoPrint descriptors (nodes) were linked where possible with edges (using known reported information or reasonable assumptions) to form directed networks representing exposure pathways.

- Co-occurrence networks can also be built by constructing weighted edges between ExpoPrints based on the number of chemicals they share.
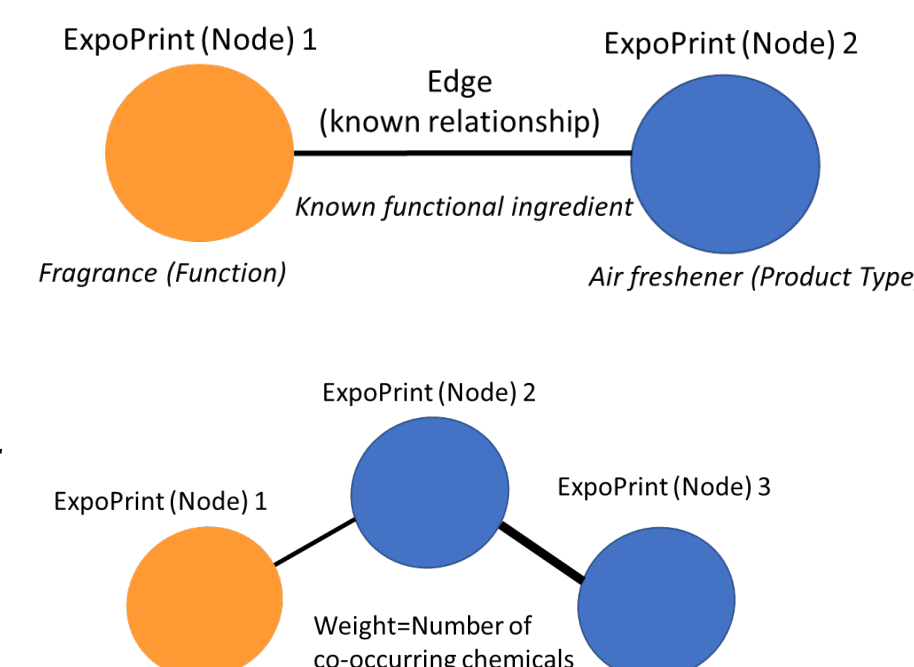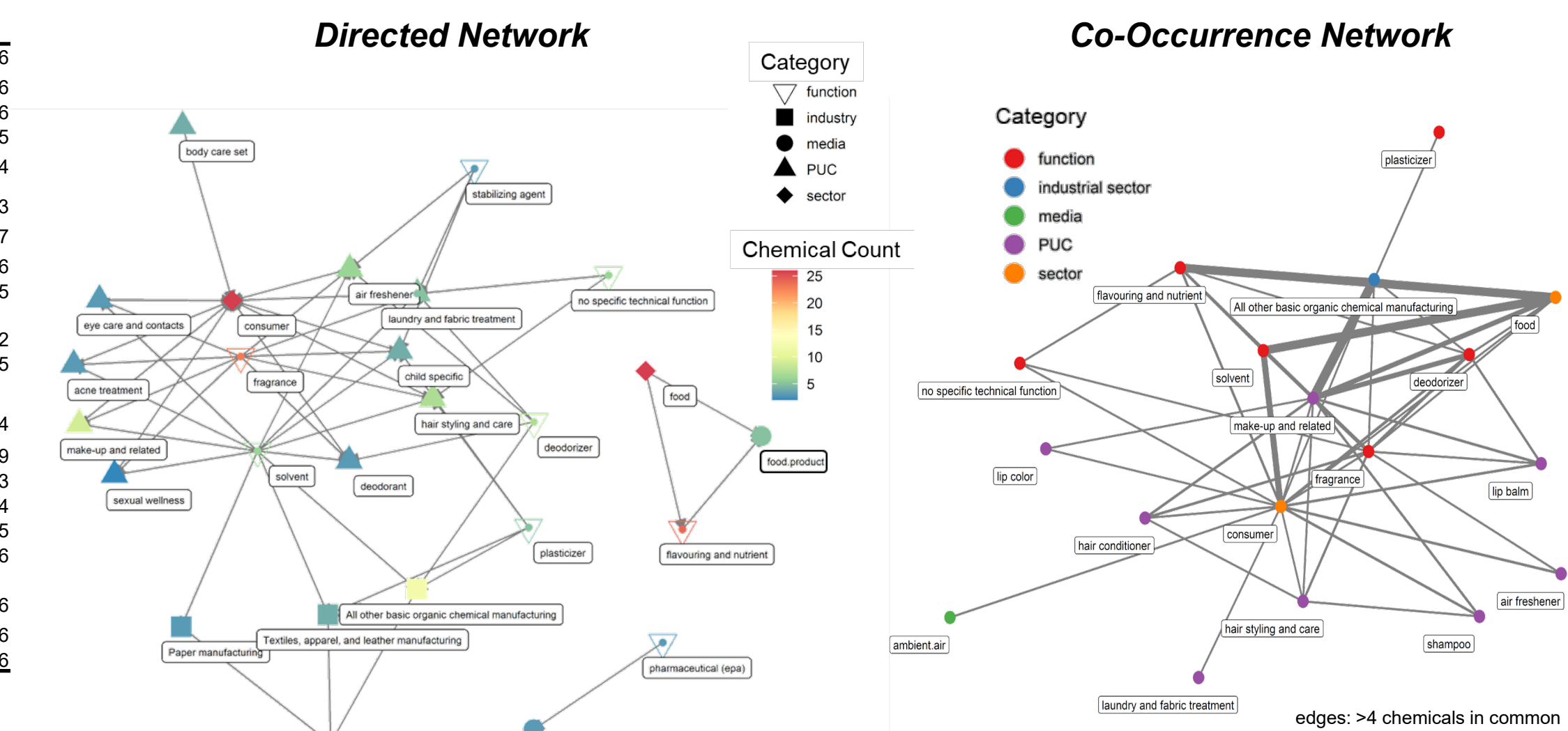

Case Studies

- ExpoPrint enrichment and associated network visualizations were used to examine two case study sets of chemicals tentatively identified in NTA
  - Chemicals tentatively identified in women but not in men in an EPA ORD analysis of pooled serum samples[6] (63 chemicals).
  - Chemicals tentatively identified in an EPA ORD analysis of finished tap water collected in California (327 chemicals, data courtesy of Seth Newton and John Sloop).

Results

Case Study 1: Chemicals Tentatively Identified in Pooled Serum from Women

| Category	ExpoPrint	Odds Ratio	p	Number of Chemicals
Broad Sector	consumer	6.77	<0.001	26
Broad Sector	food	3.52	<0.001	26
Function Category	no specific technical function	6.24	0.001	6
Function Category	plasticizer	5.54	0.003	5
Function Category	flavouring and nutrient	5.33	<0.001	24
Function Category	fragrance	4.35	<0.001	23
Function Category	deodorizer	3.96	0.004	7
Function Category	solvent	3.76	0.009	6
General Category	raw materials	14.45	<0.001	5
	All other basic organic			
Industrial Category	chemical manufacturing	3.98	0.004	12
Measured in Media	food.product	6.12	0.003	5
Product Family	body care set	8.14	0.003	4
Product Family	make-up and related	6.25	0.001	9
Product Type	body scrub	7.37	0.014	3
Product Type	eye liner	5.80	0.012	4
Product Type	lip color	4.69	0.014	5
Product Type	shampoo	4.55	0.012	6
Broad Sector	consumer	6.77	<0.001	26
Broad Sector	food	3.52	<0.001	26
Function Category	no specific technical function	6.24	0.001	6

The 20 most highly enriched ExpoPrints for a list of chemicals uniquely identified in pooled serum from women.

Case Study 2: Chemicals Tentatively Identified in Drinking Water

Category	ExpoPrint	Odds Ratio	p	Number of Chemicals
Broad Sector	consumer	3.06	<0.001	15
Function Category	monomers	32.66	<0.001	3
Function Category	processing aids not otherwise specified	3.99	0.006	6
General Category	cleaning and safety	4.83	0.003	6
Industrial Category	Furniture and related product manufacturing	27.87	<0.001	3
Industrial Category	Plastic material and resin manufacturing	7.35	<0.001	17
Industrial Category	Adhesive manufacturing	6.33	<0.001	10
Industrial Category	Paint and coating manufacturing	6.00	<0.001	15
Industrial Category	Construction	5.54	0.004	5
Industrial Category	Synthetic rubber manufacturing	5.18	0.006	5
Measured in Media	drinking.water	11.79	<0.001	9
Measured in Media	indoor.dust	10.92	<0.001	6
Measured in Media	wildlife.birds.	8.53	0.001	5
Measured in Media	wildlife..terrestrial.vertebrates.	8.29	0.001	5
Measured in Media	wastewater..influent.effluent.	8.26	<0.001	13
Measured in Media	human.blood...whole.serum.plasma.	7.72	<0.001	7
Measured in Media	wildlife..aquatic.vertebrates.mammals.	7.55	0.002	5
Product Family	acne treatment	5.11	0.005	5
Product Family	engine maintenance	4.36	0.003	7
Product Type	auto fluids and additives	4.76	0.002	7

The 20 most highly enriched ExpoPrints for a list of chemicals tentatively identified in drinking water.
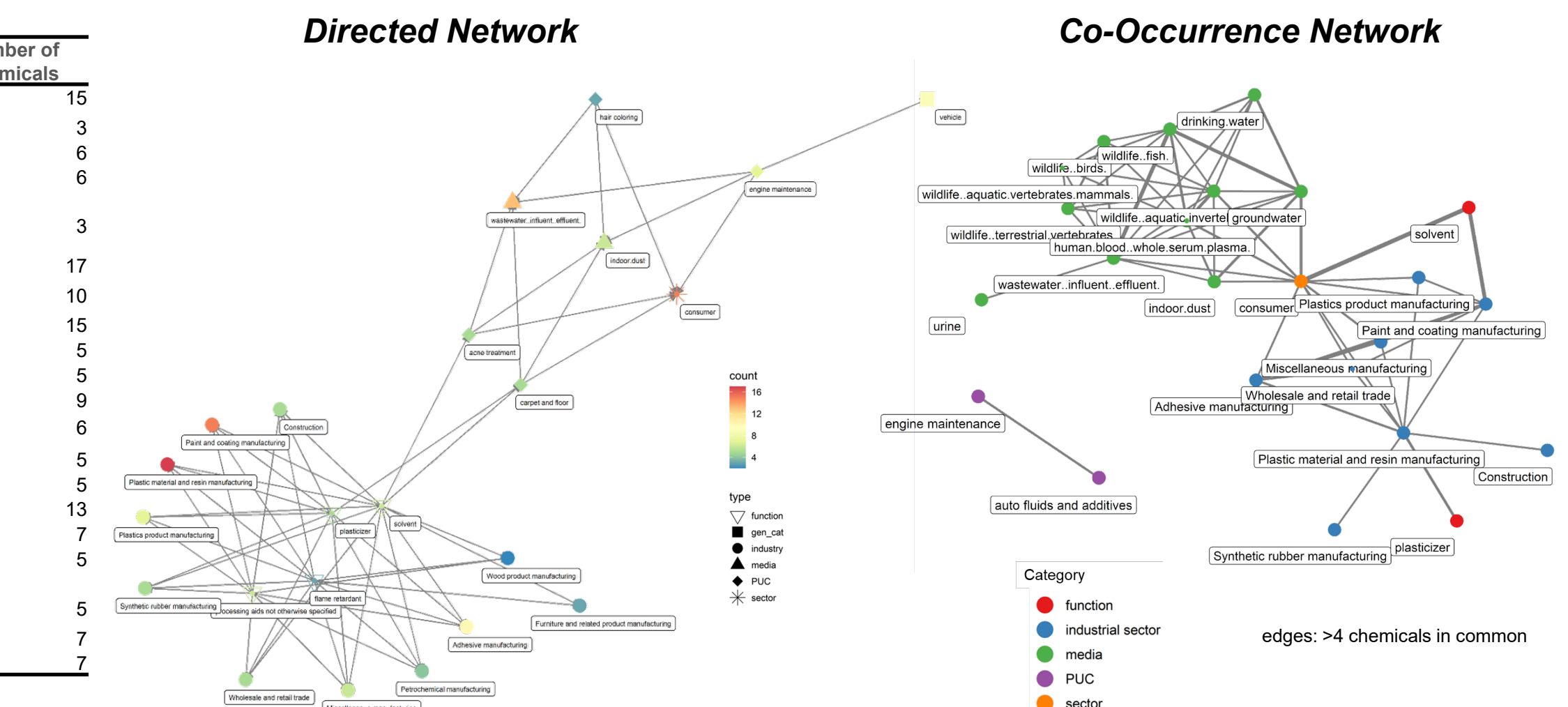
Conclusions

- The ExpoPrint methods have potential for elucidating sources and exposure pathways associated with chemicals identified in NTA.

- These methods will also be useful in NTA workflows for identifying the source and structure of unknown compounds that co-occur with known chemicals.

- Understanding the origin of exposure sources for chemicals sharing common mechanisms of toxicity will potentially inform mitigation priorities and risk management strategies.

References