Status and Trends Water Monitoring: Nontarget Analysis

Summary: Nontarget analysis (NTA) is an advanced analytical approach that can be used to identify potentially concerning contaminants that are novel or unexpected. This powerful tool based on mass spectrometry provides a broad, open-ended view of thousands of synthetic and naturally derived chemicals simultaneously, in contrast to more typical targeted analysis of known contaminants. The RMP has implemented NTA special studies to inform decision-making and is now incorporating this approach into Status and Trends monitoring.

This document summarizes a proposed study design for NTA in Bay water, for discussion at the RMP's Emerging Contaminants Workgroup meeting in April 2024. We propose dry and wet season sample collection in summer 2025 and water year 2026, respectively. At the UC Davis laboratory of Dr. Tom Young, all samples will be subjected to two types of NTA, allowing a more comprehensive assessment of the presence of both polar and nonpolar contaminants. Contaminants will be tentatively identified via matching to available spectral libraries; confirmation is an option for compounds with commercially available reference standards considered high priority. To direct nontarget compound identification toward bioactive compounds, samples will also be subjected to five in vitro bioassays probing receptors not yet examined in Bay matrices. In addition, NTA data will be explored relative to characteristics including geography, seasonality, and available S&T target contaminant data to provide additional insights on contaminant occurrence and the potential influence of sources and pathways.

Deliverables will include a spreadsheet of tentatively identified compounds, a draft manuscript to be submitted to a peer-reviewed journal, publicly accessible baseline spectra for future NTA comparisons, and archived extracts of water samples.

Estimated Cost:	$383,250 (Status & Trends); lower cost if scope is reduced
Oversight Group:	ECWG
Proposed by:	Rebecca Sutton (SFEI) and Tom Young (UC Davis)
Time Sensitive:	Yes, 2025 dry season cruise will include the most extensive list of analytical parameters to aid in interpreting nontarget data

PROPOSED DELIVERABLES AND TIMELINE

Deliverable	Due Date
Task 1. Develop sampling plan	June 2025
Task 2. Collect dry season samples	August 2025
Task 3. Collect wet season samples	March 2026
Task 4. Laboratory analysis	August 2026
Task 5. Assemble available target data for comparison	September 2026
Task 6. Draft manuscript	February 2027
Task 7. Presentation to ECWG meeting	April 2027
Task 8. Final manuscript (for submission)	June 2027

Background

In 2020, the Regional Monitoring Program for Water Quality in San Francisco Bay (RMP) undertook a review of the Status and Trends (S&T) Program, a vital component of the RMP. The goals of the S&T Program are to measure contaminant concentrations in San Francisco Bay to determine if any contaminants are at levels of concern in some or all regions of the Bay, and if concentrations are changing over time. The S&T Program has predominantly focused on legacy contaminants included on the USEPA Priority Pollutant List since its inception in 1993. However, many contaminants of emerging concern (CECs) have now been identified as present in the Bay at concentrations that may impact Bay aquatic life, leading to a major management focus on CECs in the region. Regulatory agencies are interested in having robust CECs data for the Bay so emerging issues can be detected before they cause substantial and potentially persistent effects on Bay water quality, ecosystems, and beneficial uses.

A new addition to the S&T platform, nontarget analysis (NTA) will be performed as part of the S&T Program at least every ten years to identify new CECs in Bay water. NTA uses high-resolution mass spectrometry to identify the presence of a broad suite of contaminants. It is a primarily qualitative tool that can be used to assess contaminant occurrence and geographic distribution in the Bay, and may also inform our understanding of the influence of contaminant pathways. NTA can rely on either liquid chromatography (LC), best suited for identifying polar compounds, or gas chromatography (GC), best suited for identifying nonpolar compounds. For a more comprehensive screening for the presence of unanticipated contaminants, applying both LC- and GC-based methods is useful. Prior RMP special studies using NTA have resulted in detection of new CECs of potential concern (e.g., tire-derived contaminants, ethoxylated surfactants) that merited follow-up targeted monitoring, and have informed overall S&T monitoring priorities.

Using NTA within a Status and Trends framework indicates a commitment to consistent monitoring and analysis, providing a valuable dataset for detecting changes over time, including the appearance of new contaminants. This study design is suggested to provide initial baseline data, and can serve as a pilot for future NTA implementation. Leveraging the existing S&T platform for sample collection allows comparisons of NTA findings to available quantitative targeted data during the wet and dry season, which may provide clues as to sources and transport processes associated with tentatively identified compounds. The qualitative information provided by NTA can also inform later targeted monitoring efforts, as in previous special studies. While initial identification of unknown compounds is limited by the compound-specific spectra available in spectral libraries, in the future retrospective analysis can be applied to identify unknowns via matching to spectra of more recently characterized compounds.

The proposed study design includes an opportunity to use NTA with measures of biological activity using bioassays to help identify and prioritize contaminants that are potentially toxic. This approach is similar to the effect-directed analysis that resulted in the discovery of 6PPD-quinone as a causal agent of coho mortality.[1]

Study Objectives and Applicable RMP Management Questions

Table 1. Study objectives and questions relevant to the RMP ECWG management questions.

Management Question	Study Objective	Example Information Application
1) Which CECs have the potential to adversely impact beneficial uses in San Francisco Bay?	Identify Bay water contaminants not yet characterized by targeted monitoring efforts and assess correlations with target analytes and bioassay results. Evaluate future monitoring needs and toxicity data gaps.	Have previous targeted monitoring efforts focused on contaminants and classes posing the highest relative risk to the Bay? Which newly identified contaminants merit further monitoring?
2) What are the sources, pathways, loadings, and processes leading to the presence of individual CECs or groups of CECs in the Bay?	Comparison of dry and wet season Bay water contaminants observed to evaluate the influence of different pathways (wastewater, stormwater).	Are there seasonal differences in the observations that suggest the influence of different pathways?
3) What are the physical, chemical, and biological processes that may affect the transport and fate of individual CECs or groups of CECs in the Bay?	Evaluate nontarget data alongside target data for evidence of transformation products. Evaluate nontarget data alongside bioassay results for potential biological activity.	Do nontarget results indicate the presence of transformation products not currently evaluated via target analyses? Do bioassays suggest any identified compounds might pose risks to wildlife?
4) Have levels of individual CECs or groups of CECs changed over time in the Bay or pathways? What are potential drivers contributing to change?	Establish baseline data for comparison with future Status and Trends nontarget analysis.	Does regular application of nontarget analysis reveal changes over time?
5) Are the concentrations of individual CECs or groups of CECs predicted to increase or decrease in the future?	N/A	N/A
6) What are the effects of management actions?	N/A	N/A

Approach

Bay Water Sampling

Collection of Bay water samples will be coordinated with the RMP S&T dry season water monitoring cruise in the summer of 2025 and the wet season monitoring activities in the winter of 2025-26. All samples will be grab samples of Bay water.

During the dry season water cruise, 22 sites will be sampled, a combination of six fixed stations and 16 random stations across all five Bay segments, along with two field duplicates and two field blanks. Wet season sampling currently consists of 16 overall samples, with eight at near-field sites and six at deep Bay stations. The near-field sites include three in-Bay stations near stormwater inputs (two storm events) plus one station that is also influenced by wastewater input. Four deep Bay sites will be sampled within three weeks of the storms sampled at the near-field locations. Two additional field blanks and two field duplicates will be collected as part of wet season monitoring. Overall, 46 samples will be collected and transported to UC Davis, where they will be extracted within 24 hours of delivery to the laboratory following procedures outlined below.

Nontarget Analysis and Interpretation

The Young laboratory at UC Davis has developed and validated general-purpose broad-scope screening strategies based on integrated application of liquid and gas chromatography with quadrupole time-of-flight mass spectrometry (LC-QTOF/MS and GC-QTOF/MS) capable of analyzing diverse environmental sample types.[2-8] Aqueous-phase compounds are concentrated using solid phase extraction cartridges, and the cartridges are eluted. Extracts are split and solvent exchanged to produce appropriate samples for LC- and GC- analysis (or for bioassay analysis) and are further concentrated as necessary to facilitate analysis. Each extract is spiked with internal standard; the LC mixture contains 12 labeled internal standards to support retention time alignment[9] and assessment of matrix suppression effects. Compounds in the GC-fractions are analyzed using an Agilent 7200B GC-QTOF/MS with electron ionization (EI) and retention indexing is conducted using a standard normal alkane series (C_8-C_{32}). The LC-fractions are analyzed on an Agilent 6530 LC-QTOF/MS in both positive and negative electrospray ionization (ESI+/ESI-) modes. Raw data from all GC- and LC-QTOF/MS experiments are converted from vendor format to analysis base file format for further processing (Reifycs Analysis Base File Converter v. 4.0.0). All data are subsequently deconvoluted and aligned using MS-DIAL (v. 4.90).[10]

A primary advantage of this workflow is the ability to handle both GC-EI and LC-ESI data using similar workflows. Deconvoluted and aligned features in each data set are tentatively identified by searching against relevant databases (e.g., NIST17 for GC-EI data and vendor and online databases including MassBank of North America for LC-ESI data). Extensive parameter selection experiments and workflow performance evaluations have been performed for these workflows.[2-6] When bioassay and/or toxicity

data are available for comparison, samples with the highest abundances of compounds correlated with sample bioactivity are reanalyzed to obtain targeted MS/MS data for compound identification using multiple in silico fragmentation approaches including MS-FINDER[11] and Sirius CSI-FingerID.[12] Annotations of GC-EI compounds are cross-checked against results from Agilent Unknowns Analysis (v. 10). Compounds prioritized using these methods are subsequently obtained commercially (when possible) and the identities are confirmed.

Extracts from all water samples and associated quality assurance samples will be stored frozen (-20 °C) for a minimum of three years beyond the completion of the project to facilitate subsequent confirmation and/or retrospective analyses (e.g., identification of new contaminants of concern, identification of causes of toxicity). All chromatography and mass spectrometry results from the study will be made accessible to SFEI in a format and location to be decided at the end of the study period.

Bioassays

Past work in the Young laboratory (and work by others) has shown the power of coupling in vitro and in vivo assays with nontarget analysis to direct nontarget compound identification toward bioactive compounds using bioassay results.[4-5] In this work we propose to focus on a different set of receptor bioassays than used in previous work on SFEI pilot projects on estrogenicity[13], which revealed minimal activity. The suite of bioassays to be used includes those designed to probe thyroid receptor (TR), retinoid-X receptor (RXR), glucocorticoid receptor (GR), and progesterone receptor (PR) activities. There is evidence that each of these assays can be responsive to the types of compounds found in treated wastewater and ambient water samples.[14-17] Although many bioassays use mammalian cells, the endpoints they test are highly conserved across vertebrates. The five bioassays proposed for use include:

1. *Xenopus laevis* tadpole induced metamorphosis assay. Tests for thyroid hormone and retinoid-X receptor (RXR) signaling in vivo, in a well characterized aquatic vertebrate system. Uses an integrated transgenic reporter gene as well as morphological analysis of multiple target tissues (head and jaw remodeling, brain expansion, tail resorption),
2. Thyroid hormone responsive mammalian (rat) pituitary cell line with an integrated thyroid hormone responsive reporter gene,
3. A mammalian RXR only reporter transfection assay, to rule in or out compounds affecting TR activity through its RXR partner,
4. Cell based bioassay for glucocorticoid receptor activity,
5. Cell based bioassay for progesterone receptor activity.

Collectively, these assays will probe the samples for a wide range of potential compounds that could trigger toxic effects within the San Francisco Bay ecosystem and would support the nontarget analysis by focusing identification efforts on confirming suspect identities (or identifying unmatched features) that are most strongly related to in vitro or in vivo effects in the sample extracts.

Budget

Table 2. Budget (full scope; lower cost if scope is reduced)

Expense	Estimated Hours	Estimated Cost
Labor		
Study Design	24	4,680
Sample Collection	24	3,600
Data Technical Services	24	3,840
Analysis and Reporting	160	38,400
Subcontracts		
Tom Young, UC Davis		323,730
Direct Costs		
Equipment		2,000
Travel		2,000
Shipping		2,000
Open Access Publication Fee		3,000
Grand Total		**383,250**

Budget Justification

SFEI Labor
Labor hours are estimated for SFEI staff to manage the project, develop the study design, support sample collection, analyze data, present findings, and assist with preparation of a manuscript to be submitted to a peer-reviewed journal. Costs for sample collection are minimized through leveraging RMP S&T dry and wet season monitoring.

Data and Technical Services
Nontarget data are not subject to typical data management processes. A small data services budget is included to facilitate prompt transfer of target and ancillary data to the analytical partner to aid in interpreting data.

Analytical Costs (UC Davis)
The analytical contract covers a total of 46 samples at a per sample cost of $3,400 (includes independent analyses using LC-ESI+, LC-ESI- and GC-EI), and full suspect screening of data from each platform. Sample extracts would also be screened using five separate bioassays (*Xenopus laevis* tadpole induced metamorphosis assay, TH

responsive mammalian cell line with an integrated TH responsive reporter gene, a mammalian RXR only reporter transfection assay, and in vitro cellular bioassays for glucocorticoid and progesterone receptor agonism) at a cost of $75,400. The fixed cost for data analysis, interpretation, and lead authorship on a scientific manuscript is $62,500. Including the negotiated indirect rate of 10%, the total subcontract with UC Davis covering the above work is $323,730. If the scope of the study design is reduced, this contract would be reduced as well.

Reporting

The primary deliverable will be a draft manuscript to be submitted to a peer-reviewed journal (draft February 2027, final June 2027). Feedback from a presentation of findings to the ECWG can be incorporated into the manuscript. In addition, we will produce a spreadsheet of tentatively identified compounds, and publicly accessible baseline spectra for future NTA comparisons.

References

1. Tian, Z.; Zhao, H.; Peter, K. T.; Gonzalez, M.; Wetzel, J.; Wu, C.; Hu, X.; Prat, J.; Mudrock, E.; Hettinger, R.; Cortina, A. E.; Biswas, R. G.; Kock, F. V. C.; Soong, R.; Jenne, A.; Du, B.; Hou, F.; He, H.; Lundeen, R.; Gilbreath, A.; Sutton, R.; Scholz, N. L.; Davis, J. W.; Dodd, M. C.; Simpson, A.; McIntyre, J. K.; Kolodziej, E. P. A Ubiquitous Tire Rubber–Derived Chemical Induces Acute Mortality in Coho Salmon. *Science* **2021**, *371* (6525), 185–189. https://doi.org/10.1126/science.abd6951.
2. Moschet, C.; Lew, B.M.; Hasenbein, S.; Anumol, T.; Young, T.M. "LC- and GC-QTOF-MS as Complementary Tools for a Comprehensive Micropollutant Analysis in Aquatic Systems," *Environmental Science and Technology*, **2017**, 51: 1553-1561 (10.1021/acs.est.6b05352).
3. Moschet, C.; Anumol, T.; Lew, B.M.; Bennett, D.H.; Young, T.M. "Household dust as a repository of chemical accumulation: New insights from a comprehensive high-resolution mass spectrometry study," *Environmental Science & Technology*, **2018**, 52, 2878-2887 (10.1021/acs.est.7b05767).
4. Black, G.P.; Anumol, T.; Young, T.M. "Analyzing a broader spectrum of endocrine active organic contaminants in sewage sludge with High Resolution LC-QTOF-MS suspect screening and QSAR toxicity prediction," *Environmental Science: Processes and Impacts*, **2019**, 21, 1099-1114 (10.1039/C9EM00144A).
5. Black, G.P.; He, G.; Denison, M.S.; Young, T.M. "Using Estrogenic Activity and Non-Targeted Chemical Analysis to Identify Contaminants in Sewage Sludge," *Environmental Science & Technology*, **2021**, 55, 6729-6739 (10.1021/acs.est.0c07846).
6. Young, T.M.; Black, G.P.; Wong, L.; Bloszies, C.S.; Fiehn, O.; He, G.; Denison, M.S.; Vogel, C.F.A.; Durbin-Johnson, B. "Identifying Toxicologically Significant Compounds in Urban Wildfire Ash using in Vitro Bioassays and High Resolution Mass Spectrometry," *Environmental Science & Technology*, **2021**, 55, 3657-3667 (10.1021/acs.est.0c06712).
7. Alaimo, C.P.; Li, Y.; Green, P.G.; Kleeman, M.J.; Young, T.M. "Diversity of Carbonyl Compounds in Biogas and Natural Gas Revealed Using High-Resolution Mass Spectrometry and Non-Target Analysis," *Environmental Science & Technology*, **2021**, 55:12809−12817 (10.1021/acs.est.1c01646).

8. Hattaway, M.E.; Alaimo, C.; Wong, L.; Teerlink, J.; Young, T.M. "Spatial and Temporal Variability of Micropollutants Within a Wastewater Catchment System," *Environmental Science: Processes and Impacts*, **2024**, 26, 357 - 367.
9. Aalizadeh, R.; Alygizakis, N.; Schymanski, E.; Krauss, M.; Schulze, T.; Ibáñez, M.; McEachran, A.D.; Chao, A.; Williams, A.J.; Gago-Ferrero, P.; Covaci, A.; Moschet, C.; Young, T.M.; Hollender, J.; Slobodnik, J.; Thomaidis, N. "Development and Application of Liquid Chromatographic Retention Time Indices in HRMS-based Suspect and Non-target Screening," *Analytical Chemistry*, **2021**, 93:11601-11611 (10.1021/acs.analchem.1c02348).
10. 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. Nature methods 2015, 12 (6), 523-526.
11. Tsugawa, H.; Kind, T.; Nakabayashi, R.; Yukihira, D.; Tanaka, W.; Cajka, T.; Saito, K.; Fiehn, O.; Arita, M., Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. Analytical Chemistry 2016, 88 (16), 7946-7958.
12. Dührkop, K., Fleischauer, M., Ludwig, M. *et al.* SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat Methods* **16**, 299–302 (2019). https://doi.org/10.1038/s41592-019-0344-8.
13. Denslow, N.; Kroll, K.; Mehinto, A.; Maruya, K. 2018. Estrogen Receptor In Vitro Assay Linkage Studies. SFEI Contribution No. 888. San Francisco Estuary Institute: Richmond, CA. https://www.sfei.org/documents/estrogen-receptor-vitro-assay-linkage-studies
14. Escher, B. I., et al. (2014). "Benchmarking Organic Micropollutants in Wastewater, Recycled Water and Drinking Water with In Vitro Bioassays." Environmental Science & Technology 48(3): 1940-1956.
15. Leusch, F. D. L., et al. (2018). "Comparison of in vitro and in vivo bioassays to measure thyroid hormone disrupting activity in water extracts." Chemosphere **191**: 868-875.
16. Chen, Q., et al. (2016). "Glucocorticoid activity detected by in vivo zebrafish assay and in vitro glucocorticoid receptor bioassay at environmental relevant concentrations." Chemosphere 144: 1162-1169.
17. Šauer, P., et al. (2018). "Two synthetic progestins and natural progesterone are responsible for most of the progestagenic activities in municipal wastewater treatment plant effluents in the Czech and Slovak republics." Water Research 137: 64-71.