Responses to Prof. Yu-Ju Chen

Comments: *Thank you for submitting your manuscript to Analytical Chemistry. We look forward to working with you to begin the publication process.*

Unfortunately, the submitted version of the manuscript greatly exceeds our journal's page limit guidelines. Experience teaches us that lengthy manuscripts often involve peer review that is slowed and less detailed. The length guidelines (http://pubs.acs.org/page/ancham/submission/authors.html) state a maximum length of eight journal pages for Articles. Based on counting 1000 words per journal page, single-column figures or tables as 250 words and double-column figure or table as 500 words, your paper is about 11.6 pages. Please keep the essential details of the Experimental Section in the main body of the manuscript. To shorten your article to eight pages, consider placing the more detailed experimental protocols, development of equations, and additional tables / figures in the Supporting Information section.

Responses: Many thanks for your consideration and we have revised out MS according to your suggestion. The text in the MS is approximately 6169 (without title page and abstract), with 4 figures (Figure 1, 2.74*5 inches, ~300 words; Figure 2, 4.01*7 inches, ~440 words; Figure 3, 3.35*7 inches, 372 words; Figure 4, 4.09*7 inches 455words) and 1 table (1.85*3.24 inches, ~102 words), equally to 1669 words (if correct). Now the length of the MS is approximately 7839 words, following the requirement of the journal.

Comments: *In addition, please address the following technical issues:*

1) If the manuscript is accompanied by a Supporting Information for Publication file or files, a brief description of the supplementary material is required in the manuscript. The appropriate format is:

Supporting Information: XXXX(Brief statement in non-sentence format listing the contents of the material supplied as Supporting Information.)

Examples of sufficient descriptions: "Supporting Information: 1H NMR spectra for all compounds" or "Additional experimental details, materials, and methods, including photographs of experimental setup".

Please be as kind as to improve the description in the Supporting Information paragraph in your manuscript instead of listing the figure/table legends.

Responses: Many thanks for your comments, and we have revised it in our MS as following:

"**Supporting Information:**

Supporting Information 1: supplemental methods, figures, and tables. Supporting Information 2: output files of 2H-SIAM(1.0) for the recovery study of 7 contaminants (UPLC-ESI-HRMS). Supporting Information 3: output files of 2H-SIAM(1.0) for the soil pyrene degradation study (GC-EI-HRMS)."

Comments: *2) Please provide a TOC graphic only on the last separate page of your manuscript file, label it as "For Table of Contents Only" and resize it to fit in an area 3.25 inches by 1.75 inches (approx. 8.25 cm by 4.45 cm).*

Responses: Many thanks for your comments, and we have provided it according to your suggestion.

Comments: *3) Please provide a point-by-point response to the previous reviewers' comments*

on es-2021-079298 upon revision submission.

I am requesting a minor revision of your manuscript to meet our guidelines. Please submit your revision on your ACS Paragon Plus homepage by 09-Mar-2022.

Responses: Many thanks for your consideration, and we have revised our MS according to the comments from es-2021-079298 as following.

Responses to the comments from es-2021-079298.

Reviewer: 1

Comments: *This article by Chen et al describes the application of 2H-labeled stable isotope assisted metabolomics (2H-SIAM) for the screening of potential transformation products of environmental contaminants. The authors seem to have done a lot of work, including synthesis, detailed identification including MS and NMR and developing an algorithm complete with GUI that is available on GitHub. While it appears that an impressive amount of work has been done on a topic that is of interest to ES&T readers, there are a number of improvements that should be made to the article and supporting information.*

Responses: Many thanks for your consideration, and we have revised our MS according to your suggestion.

Comments: *The choice of PAHs to be analysed with ESI-MS is quite strange, given that ESI is notoriously poor for ionising PAHs (this fact is acknowledged by the authors, nonetheless it is challenging to produce convincing results if the compounds do not ionise well).*

Responses: Many thanks for your comments.

Firstly, the accuracy of the ^{2}H-SIAM pipeline has been verified by tracking 7 typical environmental contaminants. All the 7 isotopologue pairs were correctly tracked from thousands of noises, proving the performance of ^{2}H-SIAM pipeline.

Secondly, the retention time of isotopologue pairs of PAHs is different, providing different situation against ^{13}C-SIAM. If we choose other examples, for instance, sulfamethoxazole, RT of the isotopologue pairs is same. It could be analyzed by scan-based algorithm, e.g., MetExtract II, which will cover up merits of ^{2}H-SIAM(1.0).

Thirdly, we indeed understand some relative lower identification confidence level of the TPs may affect our confidence to the results from non-targeted analysis. However, personally, without purification or synthesis putative TPs as references, there is no uncrossable gap between identification level 2 and level 5. The difference may just be due to their location in transformation pathway. TPs in the upstream of transformation pathway may have better signals and with higher identification confidence level. TPs in the downstream of transformation pathway may have very weak signals. That is because the natural mass of TPs in the downstream of transformation pathway are always small. With weak MS signals, they could only provide lower identification confidence.

It should be note that, identification confidence level does not mean their biological or environmental significance. A typical example is the study accomplished by Edward P. Kolodziej in 2018 [1]. They use non-targeted screening to identify possible reasons that lead to the death of *Oncorhynchus kisutch*. Finally, they focused on 57 possible compounds from 103 features, and published their study in EST. Among these 103 features, there are bicyclic amines, but they did not know 6PPD-quinione. Until 2021, they purify and synthesis 6PPD-quinione and confirm the reason, leading to the death of *Oncorhynchus kisutch*, is bicyclic amines,

6PPD-quinione. The finding is then published in Science in 2021 [2].

Thus, the non-targeted approaches are just the beginning of the study, not the end. Moreover, as we emphasized in the end of our MS:

"it should be noted that non-targeted analysis by MS is blind to stereochemistry and often regiochemistry [3]. Isomers of the TPs cannot be excluded, and exact annotations should be done by comparing with references [4, 5].".

Thus, before purification, NMR structure analysis, and synthesis of the reference, they are all putative TPs. Further biological or environmental significance should be evaluated by purification of extracts, or synthesized reference.

Comments: *While the code is openly available, there is limited interlinking between code, documentation and test data, such that it is not possible to use this easily and see how it works. I would strongly suggest that the authors consider adding a README file to describe the repository, the software and the contents (and provide some use documentation – or point the users to where this is, since this appears to perhaps be on Dryad instead of with the code?). The authors could consider renaming the GitHub repository to e.g. "2H-SIAM" instead of "code".*

Responses: Many thanks for your valuable comments. We have changed our data storage from Dryad to Mendeley Data, which provides better service. Additionally, we also renamed our repository in GitHub according to your comments. We provided a README.md file in GitHub, providing our contact information and a link to Mendeley Data for example data, output files, software, and details to use of 2H-SIAM(1.0). We have rewritten our README.pdf for Mendeley Data, providing more details and a step-by-step protocol to reproduce ^{2}H-SIAM pipeline by example data, MZmine2 and ^{2}H-SIAM(1.0). You could find our revision as following:

"The code of ^{2}H-SIAM (V1.0) was published in https://github.com/kechen1984/2H-SIAM, which provide a link from Mendeley Data, depositing example data of the soil pyrene degradation study, software, and details for use of ^{2}H-SIAM(1.0)."

Comments: *In the article itself, the authors describe what 2H-SIAM (the GUI) does and how they consider it is better than other approaches, but do not demonstrate this clearly with scientific results.*

Responses: Many thanks for your comments. ^{2}H-SIAM(1.0) is developed basing on previous efforts e.g., X13CMS and Credentialing features. Properties of ^{2}H-SIAM(1.0) and other algorithms were summarized in Table S1.

1. ^{2}H-SIAM(1.0) accept a features list with .csv format, enabling its compatible with other data processing algorithms, e.g., XCMS, MS-DIAL and MZmine2.
2. ^{2}H-SIAM(1.0) has triple filters which shows better performance than only 1 filter. It has been proved in our data (Table 1) and discussion.
3. ^{2}H-SIAM(1.0) is compatible with other isotope labels by providing mass differences between natural atom and isotope atom (Table S2) and discussion.
4. ^{2}H-SIAM(1.0) is compatible with other MS platform, such as GC-EI-HRMS that we used.
5. ^{2}H-SIAM(1.0) is features-based algorithms, which is better than scan-based algorithms when deal with isotopologue pairs eluted with different RT. The theory and reason have been carefully illustrated in our MS (Figure 2 a, d and e).
6. ^{2}H-SIAM(1.0) equipped with GUI interface (Figure S2, GitHub, Mendeley Data).
7. Comparison to the other SIAM study has been briefly provided [6-8].

Comments: *I have a number of additional comments, split into major, minor, proofing and supporting information below. The article needs additional proof reading; some examples from the beginning of the article have been given in the minor comments but this is not a comprehensive listing. Line numbers are taken from the author document (not the system PDF).*

Responses: We highly appreciate your helpful, insightful, and careful work. We have thoroughly revised our MS according to your comments. With your help, we believe that the quality of the MS is much better than before.

Comments: *Major:*

L79-80 (and other occurrences): the authors refer often to the cost (cost of 13C compounds, lower cost of 2H compounds) without presenting any concrete numbers or comparisons except for one brief comment once. Please back up these claims with some details?

Responses: Many thanks to your comments, and we have revised our MS according to your comments by providing following introduction:

"However, the extremely high cost of ^{13}C-labeled contaminants (e.g., ^{13}C-labeled persistent organic pollutants) limits the application of ^{13}C-SIAM in the environment. As listed in Cambridge Isotope Laboratories (www.isotope.com), 0.12 mg of pyrene-^{13}C3 costs $955, which is about 15000 times more expensive than pyrene-*d*10. Then, it is only available to carry out 13C-SIAM in a very small microcosm, such as a few grams of soil. The ^{13}C-SIAM study, in a planted soil system or a 500-liter bioreactor, is unacceptable due to the high cost of ^{13}C-labeled contaminants."

Comments: *L191: https://doi.org/10.5061/dryad.612jm6446 is not available - thus it is not possible to assess the "raw data, algorithms and details for use". As mentioned above, it would be highly desirable to have a README for the code in GitHub to point users towards this resource if/when it exists, so that the code can be used.*

L193-5: at the temporary download link, data is available. At the Dryad site, no mention is made of the authors or documentation (a download starts automatically – perhaps this could be made optional, as if the download contains all raw data, this would be huge – it is impossible for me to check with current internet connection). Further down, the documentation is available and this appears well written and describes 2H-SIAM, but I find it confusing that this is deposited on Zenodo but available through Dryad, whereas Zenodo has features to provide author listing and descriptions beyond what is currently shown on the Dryad site (note: it seems Dryad has an agreement with Zenodo to host some Dryad contents). The "How_to_Use_2H-SIAM(V1.0).pdf" document does not contain information about authors or contact details either, nor is this obvious in the GitHub repository or anywhere else; how are users meant to contact the authors? The authors should consider annotating these resources with their information and details a bit better, so that they can gain the full credit for their extensive efforts!

https://datadryad.org/stash/share/ebaZDjzyoq-dSst3xfg6hYY0yI7CGqVk4SotU9DkZcs

I do note that these comments are intended to help, since the authors have obviously gone to extensive efforts for the availability of their code and data and are to be commended on this (it is, indeed, not easy to connect up everything together and the onus is partially on the community and journals to better enable this and make it easier for authors to do so!).

Responses: We highly appreciate your understanding and careful work. This is out first time to try to provide algorithm and codes, and we indeed face many challenges. Now, we have changed our data storage, rewritten our README.pdf for Mendeley Data, providing more details and a step-by-step protocol to reproduce ^{2}H-SIAM pipeline by example data, MZmine2 and

^{2}H-SIAM(1.0). Additionally, Mendeley Data provide service to download an indicated document as you required.

Comments: *L198-207: Some of this is repetition of the introduction and is better placed in introduction, not results.*

Responses: Many thanks to your comments and we have deleted it to make the MS more concise.

Comments: *L208-215 (and other locations): The authors are discussing results that have not yet been clearly presented to the readers. Please back up all claims and discussions with cross referencing to figures/tables/text describing what you claim.*

Responses: Many thanks to your comments and we have revised it according to your suggestions. As at the very beginning, I did not want to talk about money, and I think readers may check by themselves. Now, we have added the description in the introduction about the cost of ^{13}C-labels and ^{2}H-labels as following:

"However, the extremely high cost of ^{13}C-labeled contaminants (e.g., ^{13}C-labeled persistent organic pollutants) limits the application of ^{13}C-SIAM in the environmental studies. As listed in Cambridge Isotope Laboratories (www.isotope.com), 0.12 mg of pyrene-^{13}C3 costs $955, which is about 15000 times more expensive than pyrene-*d*10. Then, it is only available to carry out ^{13}C-SIAM in a very small microcosm, such as a few grams of soil. The ^{13}C-SIAM study, in a planted soil system or a 500-liter bioreactor, is unacceptable due to the high cost of ^{13}C-labeled contaminants."

Comments: *L220-228: This again seems to be more motivation and reason for writing the 2H-SIAM algorithm (and thus would belong more in the introduction) – or are the authors trying to perform a comparison? If the latter, then more details would be needed for a comparison, which are not given. It seems better suited for the introduction in the current state, and I would suggest Table 1 goes into the supporting information, to leave the authors more space to present their own results and algorithms.*

L229-236: again this seems to be more background than results?

Responses: Thanks a lot for your valuable comments and we have revised it according to your suggestion. We have moved this description into Introduction section. Additionally, according to your comments, the comparison of algorithms has been moved to supporting information.

Comments: *L242-3: "We propose a 1:3 and 3:1 ratio" – upon the basis of which results? Please demonstrate the reason for choices with actual data.*

Responses: Thanks a lot for your comments. The higher ratio, the better performance to remove noise. However, higher ratio of the mixtures may result in the loss of positive signals as they were diluted. We have revised our MS as following:

"We propose a 1:3 and 3:1 ratio ($Mix_{1:3}$ and $Mix_{3:1}$ samples), keeping an appropriate equilibrium between accuracy and sensitivity for tracking isotopologue pairs. That is because a high mix ratio will dilute extracts, leading to the loss of positive signals."

According to your comments, here we provided a simple reason for the choices. However, we do not want to discuss too much here due to the limits of the journal pages.

Comments: *L245-251: still few results, just more context?*

Responses: Many thanks for your comments. Here we outline the usage of ^{2}H-SIAM pipeline, and it is necessary as many readers prefer to read directly from the section of Figures, Result and Discussion.

Comments: *L261-265: please name the chemicals used, or point the readers to a table where they are listed.*

Responses: Many thanks for your comments and we have provided these details in Methods.

Comments: *L270: Indeed PAHs are very poorly ionisable in ESI – how were they ionized at all?*

Responses: Yes, PAHs seems ionized in a different way, comparing to IEM model (small ion ejection from a charged nanodroplet), and uncovering the intricacies of the ESI process has proven to be surprisingly difficult and remains an active area of research [4]. Maybe the π bond in PAHs restrain the combination of H^{+}. The engineer from Thermo QE know the phenomenon, but do not know why as well.

Comments: *Up to line 293: Again no details of comparisons, just text describing how it's better, with no evidence of the comparison to back up the claims.*

Responses: Thanks a lot for your comments and we have moved most of this discussion into the introduction section.

Comments: *L323: Here there is a justification why the choice of pyrene – because it is extensively covered in KEGG. However it's poorly ionisable so creating a "challenging situation" to verify is on the one hand good and laudable, but on the other hand, to prove initially that the software works as expected, shouldn't a better example be chosen (e.g. there are many detailed pharmaceutical and pesticide degradation pathways) to establish that the software works, before a challenging example is chosen?*

Responses: Many thanks for your comments. As we introduce above, the accuracy has been verified by tracking 7 typical environmental contaminants from soil extract. Then we use an example (pyrene) to show the performance of ^{2}H-SIAM(1.0) with challenge situation with poor ionization and different retention time. If we choose some other examples, for instance, sulfamethoxazole, RT of the isotopologue pairs is same. It could be analyzed by scan-based algorithm, e.g., MetExtract II, which will cover up merits of ^{2}H-SIAM(1.0).

Comments: *L320: KEGG provides abundant information – but this information is not obvious to the readers in the results presented. It is not clear which TPs are verified via KEGG (or other methods).*

Responses: Many thanks for your comments. We use MZmine2 to match against KEGG and PubChem database, and KEGG generally only provide accurate mass of the TPs. We have firstly annotated 5 of them as TPs of pyrene and they were marked as red in Figure 4. We have the link from KEGG here, https://www.kegg.jp/pathway/map00624+C14335. But for conciseness, we do not put it in the MS, as people interested in TPs of PAHs will find it by themself. Further identification by MS^{2} is accomplished by match against *in silico* spectrum predicted by CFMID and SMILE. In many cases, laboratorial MS^{2} spectrum is unreachable for TPs of contaminant. Additionally, we provide more details about how to analysis data for the example data in our revised Method section as following.

"The tracked features from HRMS were recorded as MS^1 precursors, and they were annotated against KEGG and PubChem database (by MZmine2), or manually. Then, the MS^1 precursors from UPLC-ESI-HRMS analysis were further fragmented in quadrupole mode with a 0.4 m/z isolation window and nominal collision energy of 40, and MS^2 fragmentations were scanned with orbitrap at a resolution of 140000. The MS^2 fragments from UPLC-ESI-HRMS or fragments from GC-EI-HRMS were matched against the mass spectrum predicted by CFM-ID 3.0 for further identification".

Comments: *L325-333: The triple quantitative filters appear to help in terms of reducing the number of features, but it is not clear what each of these three filters are – nor is it clear whether true TP features are remaining and only non-TP features are being removed, or whether true TP signals are also being removed.*

Responses: Many thanks for your valuable comments. We have revised our Method according to your comments to make it clearer. The accuracy of ^{2}H-SIAM(1.0) has been verified by tracking 7 typical environmental contaminants from thousands of features from soil extract. We have revised our Method as following:

"Algorithm ^{2}H-SIAM(1.0)

When a features list (.csv format) is imported into ^{2}H-SIAM(V1.0), it starts with the calculation of means of replicates. Then, the features list is inspected within an indicated duration of RT for pairs of two features as potential isotopologue pairs; the potential natural compounds are denoted as M, and the potential isotope-labeled compounds are denoted as M'. The mean intensities of the features M and M' from samples $Mix_{1:3}$ and $Mix_{3:1}$ are denoted as $M_{Mix1:3}$, $M_{Mix3:1}$, $M'_{Mix1:3}$ and $M'_{Mix3:1}$ respectively. They are used for the calculation of the following 3 actual ratios (f_n):

$$f_1=\frac{M_{Mix1:3}}{M'_{Mix1:3}},\ f_2=\frac{M_{Mix3:1}}{M'_{Mix3:1}},\ f_3=\frac{M'_{Mix1:3}}{M'_{Mix3:1}}$$

users defined parameters R_n and $Tol._n$ are used to set triple filters (F_1, F_2 and F_3) for TPs as following:

Filter n, F_n: $R_n \times Tol._n < f_n < R_n \times Tol._n^{-1}$

where, R_n is theoretic ratios for F_n, and $Tol._n$ is tolerances for R_n, $0 < Tol._n < 1$.

When f_1 follows the requirement of F_1, a pair of two features (isotopologue pair) is tracked by the algorithm and further filters F_2 and F_3 will be sequentially evaluated. The features passing triple filters contain features as potential TPs, deserving further annotation and identification."

For the question "*whether true TP signals are also being removed.*", we could confirm that, **if the signals of TPs follow the filters with indicated tolerance, they will be correctly tracked**. That do not mean that we could track every signal from TPs. **The factors affecting MS signals** and **the setting of the tolerance for filters** will affect whether filters could correctively track TPs.

That is because, firstly, signals from ESI varied drastically. As we know, variation of ESI source is much bigger than EI source, and generally technique repeat for the measurement from ESI is acceptable when RSD < 30%; meanwhile for EI source, commonly, it is smaller than 2%. Secondly, for the TPs located in the downstream of transformation pathway, signals from these TPs are always weak. That is because the natural mass of TPs in the downstream of transformation pathway are always small, which will significantly affect their signals detected by MS. In that case, their signals may vary drastically and finally cannot be tracked by our filters with indicated tolerance. That is the reason we noted in the discussion

"The key for efficiently tracking TPs is to acquire a high-quality features list, and the optimized parameters for the selected algorithm are critically important.".

The setting of **tolerance for filters is the choice of "efficiency" or "accuracy".** Whether isotopologue pairs could be tracked by filters depends on the tolerance you set up. When you set up a strict filter, you may get fewer positive signals, but you could remove more noise. If you set up a loose filter, you may get all of the positive signals inside, but you have to face much more noise.

The theory for the 1st filter for all SIAM is simple and same, e.g., ^{2}H-SIAM(1.0), X13CMS and MetExtract II, and they just **compare intensities** from **natural** and **isotope-labeled signals with the indicated ratio (e.g., 1:3) and indicated tolerance** to track isotopologue pairs and to remove noise.

For ^{2}H-SIAM(1.0) and Credentialing features, they have 3 filters, and the 2nd filter also compare intensities from natural and isotope-labeled signals but with different ratio (e.g., 3:1). The 3rd filter compares signals from **ONLY isotope-labeled signals** from two mixed samples ($Mix_{1:3}$ and $Mix_{3:1}$, theoretical ratio is 3). As illustrated in our MS:

"Additionally, since signals from the same amount of natural and ^{2}H-labeled contaminants are not always equivalent, F_1 and F_2 constitute relaxed filters with a lower Tol. of 0.3, and F_3 constitute a stricter filter with a higher Tol. of 0.5.".

Comments: *L335-6: it is not clear how the TP annotation is being performed – nor what the authors consider a Level 3 vs 5 (see later comments). How many of these TPs were ground truth – i.e. confirmed at Level 1 or are verified TPs under these conditions? This is still a lot of speculation (if only Level 3 or lower).*

L345: Please indicate early on in the manuscript what the authors consider to be Level 2, 3, 4, 5 – although they have cited the 2014 paper, the levels (and accompanying interpretation) do not seem to match.

L349: Level 4 implies molecular formula only … but the authors claim this turns into Level 3 because of matching CFM-ID fragmentation – which implies a structure is known. Please see above comment, this interpretation does not seem to fit the Level scheme used, especially since credentialing was not explicitly mentioned?

Responses: Many thanks for your careful work. We indeed had some mistake due to careless copy and paste, and finally submitted wrong table and figure in the MS. We have corrected them in our MS. Additionally, after thinking over our study, we carefully and slightly revised the standard for identification level, and the reason and the revised standard for identification level are as following:

"Identification Confidence

Compared to common non-targeted annotation, the tracked features from ^{2}H-SIAM coeluted with their isotopologues, confirming its origination from substrates and providing basic structure information. Additionally, the number of ^{2}H-labels provides potential structure information. Thus, we annotated TPs with identification confidence levels proposed by Schymanski et al. but with little modification [9].

Level 1 was achieved by matching the RT and MS^2 with reference standards; Level 2a was achieved by matching the number of possible ^{2}H-labels and at least two major fragment ions with MS^2 libraries (from *in silico* or references spectrum), or achieved by matching at least three major fragment ions with MS^2 libraries; Level 2b was achieved by matching the number of possible ^{2}H-labels and one major fragment ions with MS^2 libraries, or achieved by matching at least two major fragment ions with MS^2 libraries; Level 3 was achieved by matching the number of possible ^{2}H-labels and natural isotope pattern, or achieved by matching one major fragment ions with MS^2

libraries; Level 4 was achieved by matching the number of possible ^{2}H-labels, or natural isotope pattern; Level 5 was achieved by matching the exact mass."

"credentialing" refers to the feature with its isotope labeled isotopologue. Indeed, our discussion is not clear, and we have revised it as:

"Additionally, isotope labels are important for tracking possible TPs in the environment because the environment matrix may contain TPs naturally.".

Comments: *L337-343: The authors are comparing their method with other studies, but not under completely comparable conditions. They should be running the other software on their own data to make these claims, not comparing different software being used to process different datasets and claiming that different number prove that theirs is better.*

Responses: Many thanks for your valuable comments. Indeed, comparing to the data existed seems much easier and fair than what we have done with our own data. However, we choose 3 experiments to verify the performance of ^{2}H-SIAM due to the following reasons:

1. There are very little SIAM study in the environmental study. To our knowledge, just two studies were carried out by SIAM [7, 8]. None of them shared raw data.

2. We need to carry out ^{2}H-SIAM to study PAHs, as their isotopologue pairs elute with different RT. This will contribute to illustrate the merit of ^{2}H-SIAM(1.0).

3. Furthermore, as we mentioned above, the theory for filters is easy, and "The key for efficiently tracking TPs is to acquire a high-quality features list, and the optimized parameters for the selected algorithm are critically important.". The quality of the features list decides the quality of the result of SIAM.

Many SIAM algorithms, e.g., NTFD, HiTIME and MetExtract II, process raw data will their embedded data processing algorithms. They are much different than ours, and we use MZmine2 to get the features list. In these cases, it is unable to fairly compare them with our pipeline, as we will definitely obtain different features lists by different algorithms.

For other algorithms, e.g., X13CMS, Credentialing features, they use XCMS to process raw data, which indeed is compatible for our pipeline. But some problems still exist. For example, Credentialing features do not support the study of ^{2}H-labels, and we could not use data from ^{2}H-labels. Additionally, X13CMS just use 1 filter. Thus, the advantage of our ^{2}H-SIAM(1.0) over it is clear.

Due to these reasons, if we compare algorithms with the same data, the manuscript will be very fragmented. We must face different situation, use different parameters, and to illustrate one by one. The obtained comparison may be still unfair. Thus, we just, briefly and arbitrarily, compared our study to 3 previous studies with simple description [6-8].

Comments: *Conclusions: While the authors have provided an interesting demonstration, more work could be done to validate and ground truth this. Hopefully the provided comments will help the authors revise this.*

Responses: Many thanks for your valuable comments.

Comments: *Figure 2: Please indicate in the caption whether d is the deuterated and e the undeuterated forms? Please also clarify somewhere the use of M and M' notation? This is quite hard to read.*

Responses: Many thanks for your valuable comments, and we have revised it according to your suggestion.

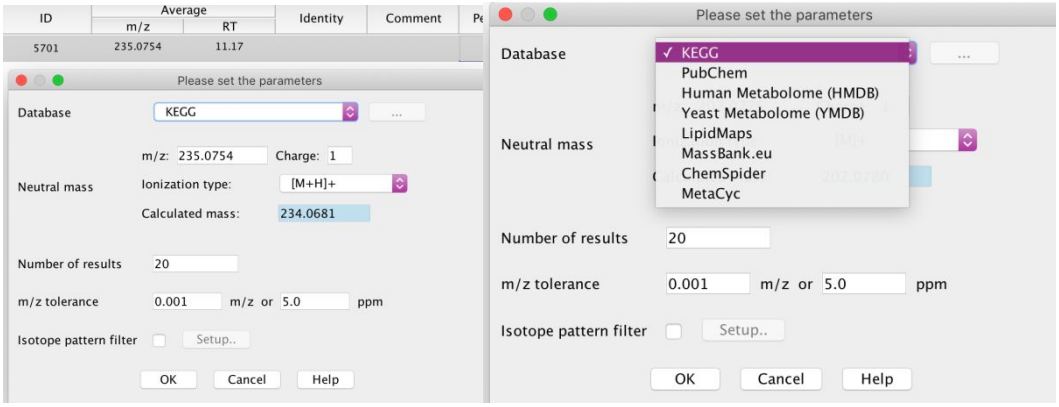
Comments: *Figures in general: the use of dark green and dark blue makes it very difficult to tell between the peaks in the figures, perhaps reconsider the colour choice to make these easier to distinguish?*

Responses: Many thanks for your careful work. Over the figures, we use dark green to denote natural compounds and use dark blue to denote isotope labeled compounds. Collocation of dark green and dark blue has been widely used, and I like them very much.

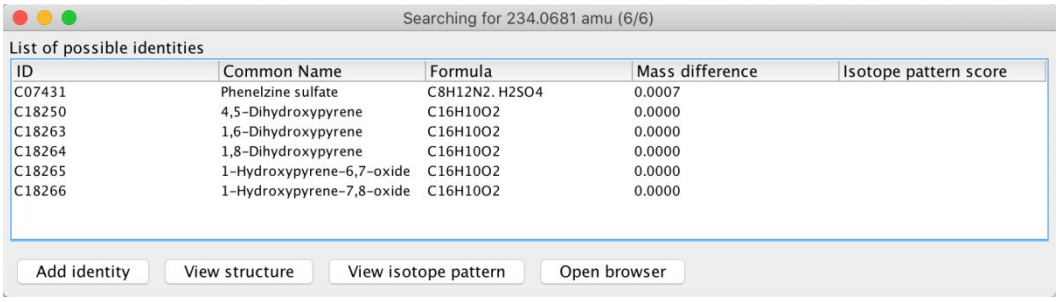
Comments: *Figure 4b: It is not clear what criteria are being used to assign the levels, since all of these have structures, but some are Level 5, some Level 4, some Level 3, some Level 2 … and surely the middle is Level 1. What reaction has been used to form the structure bottom right (M225)? Which of these TPs have been found in KEGG, which in other resources, where did the other structures come from?*

Responses: Many thanks to your careful work. As introduced above, we indeed made some mistake in the annotation of identification level. We have corrected all of them. Additionally, we have provided more details about how to annotate them in Method as we pasted above.

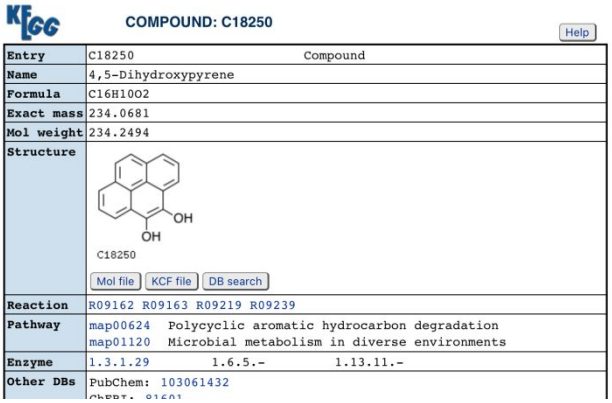
Briefly, the tracked features by ^{2}H-SIAM(1.0) were firstly matched against KEGG and PubChem database or manually. It is the function of MZmine2 as following:



then we obtained briefly as following:



We could generally obtain formula and potential name. Then we need to check or deduce whether the name (structure) follows the TPs of pyrene.



M_{235}, *d*10, L3

Due to the formation of -OH will lead to the loss of ^{2}H-labels, and we obtained M'$_{235\text{-}d10}$ with 10 ^{2}H-labels, the formation of two oxide is much more reasonable.

Then, if possible, we will use CFMID to check the fragment from LC-MS or GC-MS. It should be noted that mass spectrometry is hard to identify the location of -O- ring.

The annotation needs many tools, website, and knowledge, and manual work is necessary.

For the question, "*What reaction has been used to form the structure bottom right (M225)?* ".

It is hard to say now, as finding clear pathway needs huge work. For example, we just check positive model here, we may loss some TPs of pyrene which should be examined in negative model [8]. Our latest work has tracked more than 20 TPs of pyrene in positive model. In the future, maybe we could track more than 30 TPs of pyrene by the combination of positive and negative model. In that moment, we will be clearer about the formation of M225.

Additionally, M225 is a very interesting product as it is a mineralization product of pyrene. We are ready to carry out further study. I could paste little latest information here, organized in the other work.

"M_{225} may transform to phenanthro[4,5-bcd]thiophene, **which was reported and enriched in lobsters closing to a closed coal-coking plant** [10, 11]. Thus, phenanthro[4,5-bcd]thiophene may be originated from the mineralization of pyrene in the environment by biotic and abiotic factors."

Thus, even though we did not have MS^2 of M_{225}, and its identification level is only 3 (by our modified standard), we will try to purify or synthesis it in the future to understand its environmental significance.

Comments: *Figure 4 caption, L628: please include the proper reference (superscript number, as done for the references in the main text) rather than this text?*

Responses: Many thanks to your careful work, and we have revised the caption according to your comments.

Comments: *Table 1: as mentioned above, perhaps this is better placed in the SI, to allow more space for original results? Have the authors considered the PAVE algorithm?*

Responses: Many thanks to your valuable comments, and we have moved it into SI according to your comments. Additionally, PAVE is designed for special work, e.g., growing cells with different isotopic conditions, e.g., unlabeled, 15N, 13C, and 15C+13N. Additionally, theoretically, ^{2}H-SIAM(1.0) could also track features labeled by ^{15}N, ^{13}C, and their combination by QE with enough resolution and accuracy. The function of annotation could be accomplished by other algorithm, e.g., CFMID4.0 and SIRIUS4.

Minor comments:

Comments: *L69: The databases PubChem, MassBank, KEGG and METLIN are all very different (and contain very different information), none are cited directly, nor is it quite clear what information the authors wish to retrieve from these respective databases?*

Responses: Many thanks to your comments. They are all powerful in finding TPs, some of them provide MS^1, and some of them provide MS^2 spectrum. I have deleted these sentences to focus on our topic.

Comments: *L80-82: There have been some efforts on HDX (hydrogen-deuterium exchange) on environmental contaminants that is not cited by the authors; perhaps the authors could discuss in a little more detail the pros and cons of HDX vs SIAM? What is the cost of 2H-SIAM vs e.g. HDX and which would be preferable in which circumstances? Would complementary information be revealed?*

Responses: Many thanks for your valuable comments and we have provided more details in our revised MS about their cost as following:

"As 100 grams of D_2O (99.9%, Sigma-Aldrich) cost only ~$300, and catalysts could be reused, the possible H/D-exchange reaction offers various affordable ^{2}H-labeled contaminants. Thus, it enables us to carry out ^{2}H-SIAM in a complex environmental matrix, for instance, an ecotron with several cubic meters of reconstituted aquatic ecosystem."

Comments: *L134 and many occurrences: "xzXML" is used incorrectly instead of "mzXML" – please fix all occurrences of "xzXML" to "mzXML" (this also includes figures and potentially tables, plus SI).*

Responses: Many thanks for your careful work and we have corrected them according to your comments.

Comments: *L140: Please provide the direct link to the Dryad dataset where mentioned (e.g. as a reference, if not an inline URL) so that readers can find this data.*

Responses: Many thanks for your careful work and we have use Mendeley Data instead.

Comments: *L146: Adjust the equation to mention "height" not "heigh"*

Responses: Many thanks for your careful work and we have corrected them according to your comments.

Comments: *L158: "watered twice a month" – how much? What moisture content?*

Responses: Many thanks for your careful work and we water them with 1 mL water per gram soil twice a month. Providing little water will accelerate the transformation of pyrene. As that will not affect the result of the study, we do not take it seriously.

Comments: *L200: If reference standards are available, non-target screening can also enable Level 1 identification (matching RT, MS and MS/MS to the reference standard).*

Responses: Many thanks for your comments. I think non-target screening refers to annotate features without references.

Comments: *L218: Fig S14 – please have figures and SI figures in order?*

L326: (again – more occurrences) please check the order of figures

Responses: Many thanks for your valuable comments and we have adjusted the sequence of

figures and tables according to your suggestion.

Comments: *L362: tentatively (not temporally)?*

Responses: Many thanks for your careful work and we have corrected it according to your comments.

Comments: *L386: What is "PBCM0639661"? This is not a valid PubChem identifier but rather the external identifier provided by the depositor (which appears to be a legacy contributor). Perhaps using the CID 616438 is better? https://pubchem.ncbi.nlm.nih.gov/compound/616438*

Responses: Many thanks for your careful work and we revised it by providing the full name instead.

Comments: *Proofing comments:*

L22: illustrates

L23: "prospects of applying … in environmental studies. We demonstrate that …"

L25: "easy synthesis of…". "A pyrene…"

L30: "for the first time"

L31: "over other SIAM pipelines"

L36: "a powerful tool for" (not in)

L40: "The 2H-SIAM(V1.0) algorithm …" … to perform SIAM in environmental studies"

L41: enables

L43: studies.

L84: environmental studies.

L110: detailed information is …

L117-8: "rest of the experimental procedure …"

Responses: Many thanks for your careful work and we have corrected them according to your comments.

Comments: *Supporting Information comments:*

L55-56: (and other locations): please use the formal citation style (superscript numbers) as done in the main text, rather than these text based inline references.

L88: as for main text, please replace all occurrences of xzXML with mzXML

L93-95: This sounds like free advertising for mzMine2 – please try to keep your comments objective, all of these approaches have various pros and cons and suit different audiences.

Responses: Many thanks for your careful work and we have corrected them according to your comments.

Comments: *L97-99: I'm afraid this doesn't make sense to me without additional context, what is ADAP (mentioned several times, not just here), what are all these data processing steps referring to? Is this in mzMine2 or elsewhere?*

Responses: Many thanks for your careful work and we have revised it to make it clear. Indeed, ADAP is one of the algorithms embedded in MZmine2 for peaks/features picking and chromatogram deconvolution.

Comments: *SI in general: a lot of parameters are presented, but it is not always clear if these are the same between various conditions, or whether they are varying and why – perhaps this*

could be made more clear?

Responses: Thanks a lot for your valuable comments. I also want to use the same parameters for all of the task. That may lead to a bad performance of ^{2}H-SIAM as list in the new Table 1, as we introduce above:

"The key for efficiently tracking TPs is to acquire a high-quality features list, and the optimized parameters for the selected algorithm are critically important.".

Additionally, for different contaminants, we may choose different solvents, e.g., the choose of MeCN and MeOH. That will also lead to the difference for setting parameters. We try to make everything clear for my readers, which may complicate the description. Anyway, we will try out best to make it clearer than before.

Comments: *L180: 40 V – note that Orbitrap instruments report "nominal collision energy" or NCE, rather than an exact voltage.*

Figure S1: correct xzXML to mzXML, Varies to Various

Responses: Thanks a lot for your valuable comments and we have revised them according to your comments.

Comments: *Figure S3: What are these three filters? Give them keywords instead of just Filter 1, Filter 2, Filter 3?*

Responses: Many thanks for your careful work. We have removed this figure and composed a new table (Table 1) to show how ^{2}H-SIAM discard noise and its performance:

Table 1 Performance of ^{2}H-SIAM(1.0) in this study.

Features	Experiments		
	7 Contaminants	Pyrene Degradation	
	UPLC-ESI-HRMS	GC-EI-HRMS	UPLC-ESI-HRMS
Total	7704	21955	4376
Paired	1720	2548	658
Filter 1	891	1122	318
Filter 2	249	301	81
Filter 3	130	162	52
Annotated	7	5	14

Comments: *Figure S4: Very difficult to see the difference between dark blue and dark green (applies to many figures).*

Responses: Many thanks for your valuable comments and we have revised it by using much thinner line to make them clearer.

Comments: *Figure S5: 162 is still a lot of features, how many of these are real?*

Responses: Many thanks for your valuable comments. As introduced in our MS, we annotated 5 of them among 162 tracked features. The figure has been removed to make the description clearer.

Comments: *Figure S7: These ratios do not appear particularly consistent with the 3:1 / 1:3 ratios?*

Responses: Yes, that's true, and that is the reason why we need to set appropriate tolerance for filters.

Comments: *Figure S9: This is missing some of the comparison plots of the other figures*

above? Is there an overview of the "Ms" somewhere? Perhaps provide a cross reference?

Responses: Many thanks for your comments. These signals are weak, and we could not get their MS^2 from UPLC-ESI-HRMS. Further identification will be done after extraction and purification in our future work.

Comments: *Table S1: If a name is given for M219, why is it Level 5? Ditto for Level 4s, these all have names? Some Level 3s are marked with possible isomers, but some aren't – this is difficult to interpret according to the original level scheme. Same applies to Table S2.*

Responses: Many thanks for your careful work. We indeed made a mistake to paste wrong figures and tables for identification levels. We have corrected them in our revised MS. Additionally, we could annotate every feature, and the differences are their identification levels, referring to their possible structure.

Comments: *Table S3: This would be better provided as a CSV file with sufficient column headers, rather than the non-intuitive merged information presented in the "tentative match" column. The information for the Filter columns is likewise very difficult to understand and interpret, despite the note at the bottom.*

Table S4 and S5: see comment for Table S3.

Responses: Many thanks for your valuable comments and we have revised it according to your comments. We have provided 2 more supporting information for them.

Comments: *All in all a commendable effort by the authors, hopefully these comments help present the information in a way that the results are more clear to the interested readership.*

Responses: Thanks a lot for your kindly and selfless comments. I have read them carefully and I would like to express my heartfelt gratitude for your work on our MS. I have carefully revised our MS according to your comments, and the valuable comments are very enlightened for us to improve the quality of the MS and our future research work.

Reviewer: 2

Comments: *Overall the paper is properly written and organized and the paper has for sure a value.*

Responses: Many thanks for your valuable comments.

Comments: *However, I think it fails in the presentation of the topic, background and framework: authors are mixing some concepts related to metabolomics, leading to some misunderstandings and confusion.*

Metabolomics approaches deal with the study of endogenous metabolites of an organism whose levels are altered due to an external stressor.

However, authors use the term "metabolomics" to refer to the non-target analysis of transformation products of chemical exogenous compounds (e.g. pyrene) in environmental samples (soil), not a a living organism.

They say for instance:

Line 65: "Stable Isotope Assisted Metabolomics (SIAM) is the most convenient and reliable approach for screening TPs from noise features 11-13, facilitating TPs screening with high confidence 14."

In none of these references 11-14 the screening of TPs of xenobiotic in environmental samples addressed but the examination of endogenous metabolites responding to stressors

Responses: Many thanks for your careful comments.

Firstly, as you mentioned above, we did not use the term of "*TPs of xenobiotic*", and we used "TPs". **We believe that the TPs include the concept of endogenous and xenobiotic metabolites.** Metabolites of sugar, organic acid, amino acid, and lipid are the TPs of these molecules. Moreover, TPs also include the conjunction among them, for instance, Glu + Leu => Glu-Leu. We just use the sentence you referred to describe how SIAM be used in biology.

Secondly, the transformation products of contaminants in the environment are the result of biological and non-biological factors. Most of them were recognized as the result of activities of microbes. In this sense, they are the metabolites of contaminants. Thus, it should be appropriate to use the word "Metabolomics" to describe them.

Thirdly, there are 2 examples, using Stable Isotope-Assisted Metabolomics (SIAM) to study TPs of contaminants in environment or biological samples.

1. Tian, Z. Y.; Vila, J.; Yu, M. A.; Bodnar, W.; Aitken, M. D., Tracing the Biotransformation of Polycyclic Aromatic Hydrocarbons in Contaminated Soil Using **Stable Isotope-Assisted Metabolomics**. Environmental Science & Technology Letters 2018, 5, (2), 103-109.

2. Cheng, Z. P.; Sun, H. W.; Sidhu, H. S.; Sy, N. D.; Wang, X. R.; Gan, J., Conjugation of Di-n-butyl Phthalate **Metabolites** in Arabidopsis thaliana and Potential Deconjugation in Human Microsomes. Environ Sci Technol 2021, 55, (4), 2381-2391.

So, our definition is correct and accepted by scientific world.

Comments: *Line 57-59*

"It has shown its advantages in understanding the fates of organic contaminants in environmental study, including understanding their enrichment pathways and the discovery of TPs 9, 10"

Neither reference 9 nor 10 are nor refer to metabolomics studies

Responses: Many thanks for your comments. The following is the old paragraph:

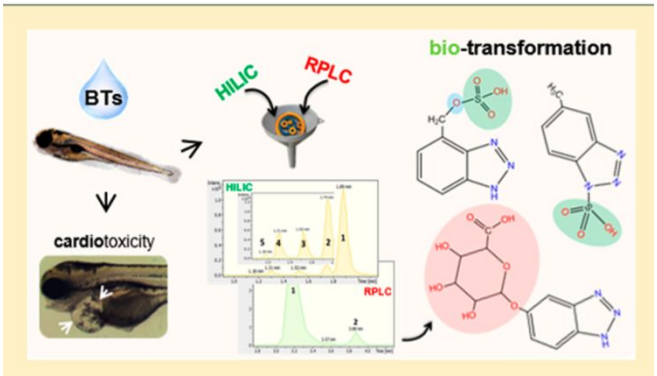
"Over the past decade, the growth, evolution, and accessibility of UPLC-ESI-HRMS (Ultra Performance Liquid Chromatography - Electrospray Ionization - High-Resolution Mass Spectrometry) has witnessed the progress of untargeted metabolomics and its application in the scientific world, which proposes a valuable option to the issue [6-8]. It has shown its advantages in understanding the fates of organic contaminants in environmental study, including understanding their enrichment pathways and the discovery of TPs [9, 10]."

In this paragraph, we describe the use of **HRMS and untargeted metabolomics**. The citation is as following:

9.Damalas, D. E.; Bletsou, A. A.; Agalou, A.; Beis, D.; Thomaidis, N. S., Assessment of the Acute Toxicity, Uptake and Biotransformation Potential of Benzotriazoles in Zebrafish (Danio rerio) Larvae Combining HILIC- with RPLC-HRMS for High-Throughput Identification. Environ Sci Technol 2018, 52, (10), 6023-6031.

10. Escher, B. I.; Stapleton, H. M.; Schymanski, E. L., Tracking complex mixtures of chemicals in our changing environment. Science 2020, 367, (6476), 388-+.

For citation 9, the following is its figure abstract; citation 10 is a REVIEW.



These citations refer to "understanding their enrichment pathways and the discovery of TPs". They are cited properly.

Indeed, Stable Isotope Assisted Metabolomics generally focus on the study of metabolites of isotope labeled sugars, lipids and so on, occurring in living organisms. But we try to use these approaches in the environmental studies as references *Environmental Science & Technology Letters 2018, 5, (2), 103-109* and *Environ Sci Technol 2021, 55, (4), 2381-2391*.

In summary, to avoid misunderstanding in the future, we have revised and shorten our MS as follwoing:

"Anthropogenic organics in the earth generate countless transformation products (TPs) by biological and non-biological factors. Their fates in the environment is of increasing concern, because they may constitute highly mutagenic and carcinogenic properties [12, 13]. Over the past decade, the growth, evolution, and accessibility of High Resolution Mass Spectrometry (HRMS) has witnessed the progress of non-targeted analysis (NTA), also referred to as "non-target screening", " untargeted metabolomics", and "untargeted screening", among several other related terms [5, 14, 15]. Additionally, Stable Isotope Assisted Metabolomics (SIAM) enables global tracking of isotopic labels from parent compounds in non-targeted metabolomics in living organisms [16-19]. Numbers of algorithms and software contribute to their raw data process, statistical analysis, and annotation [20-23]."

References for responces:

1. Peter, K. T.; Tian, Z. Y.; Wu, C.; Lin, P.; White, S.; Du, B. W.; McIntyre, J. K.; Scholz, N. L.; Kolodziej, E. P., Using High-Resolution Mass Spectrometry to Identify Organic Contaminants Linked to Urban Stormwater Mortality Syndrome in Coho Salmon. *Environ Sci Technol* **2018,** *52*, (18), 10317-10327.

2. Tian, Z. Y.; Zhao, H. Q.; Peter, K. T.; Gonzalez, M.; Wetzel, J.; Wu, C.; Hu, X. M.; Prat, J.; Mudrock, E.; Hettinger, R.; Cortina, A. E.; Biswas, R. G.; Kock, F. V. C.; Soong, R.; Jenne, A.; Du, B. W.; 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.

3. Tripathi, A.; Vazquez-Baeza, Y.; Gauglitz, J. M.; Wang, M. X.; Duhrkop, K.; Nothias-Esposito, M.; Acharya, D. D.; Ernst, M.; van der Hooft, J. J. J.; Zhu, Q. Y.; McDonald, D.; Brejnrod, A. D.; Gonzalez, A.; Handelsman, J.; Fleischauer, M.; Ludwig, M.; Bocker, S.; Nothias, L. F.; Knight, R.; Dorrestein, P. C., Chemically informed analyses of metabolomics mass spectrometry data with Qemistree. *Nat Chem Biol* **2021,** *17*, (2), 146-+.

4. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., Unraveling the Mechanism of

Electrospray Ionization. *Anal Chem* **2013,** *85*, (1), 2-9.
5. Sindelar, M.; Patti, G. J., Chemical Discovery in the Era of Metabolomics. *J Am Chem Soc* **2020,** *142*, (20), 9097-9105.
6. Mahieu, N. G.; Huang, X.; Chen, Y., Jr.; Patti, G. J., Credentialing Features: A Platform to Benchmark and Optimize Untargeted Metabolomic Methods. *Anal Chem* **2014,** *86*, (19), 9583-9589.
7. Cheng, Z. P.; Sun, H. W.; Sidhu, H. S.; Sy, N. D.; Wang, X. R.; Gan, J., Conjugation of Di-n-butyl Phthalate Metabolites in Arabidopsis thaliana and Potential Deconjugation in Human Microsomes. *Environ Sci Technol* **2021,** *55*, (4), 2381-2391.
8. Tian, Z. Y.; Vila, J.; Yu, M. A.; Bodnar, W.; Aitken, M. D., Tracing the Biotransformation of Polycyclic Aromatic Hydrocarbons in Contaminated Soil Using Stable Isotope-Assisted Metabolomics. *Environmental Science & Technology Letters* **2018,** *5*, (2), 103-109.
9. Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J., Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol* **2014,** *48*, (4), 2097-2098.
10. King, T. L.; Uthe, J. F.; Musial, C. J., Polycyclic aromatic hydrocarbons in the digestive glands of the American lobster, Homarus americanus, captured in the proximity of a coal-coking plant. *B Environ Contam Tox* **1993,** *50*, (6), 907-914.
11. Hale, R. C.; Aneiro, K. M., Determination of coal tar and creosote constituents in the aquatic environment. *Journal Of Chromatorgraphy A* **1997,** *774*, (1-2), 79-95.
12. Brussaard, C. P. D.; Peperzak, L.; Beggah, S.; Wick, L. Y.; Wuerz, B.; Weber, J.; Arey, J. S.; van der Burg, B.; Jonas, A.; Huisman, J.; van der Meer, J. R., Immediate ecotoxicological effects of short-lived oil spills on marine biota. *Nature Communications* **2016,** *7*, 11206.
13. Gonzalez-Gaya, B.; Martinez-Varela, A.; Vila-Costa, M.; Casal, P.; Cerro-Galvez, E.; Berrojalbiz, N.; Lundin, D.; Vidal, M.; Mompean, C.; Bode, A.; Jimenez, B.; Dachs, J., Biodegradation as an important sink of aromatic hydrocarbons in the oceans. *Nat Geosci* **2019,** *12*, (2), 119–125.
14. Perez de Souza, L.; Alseekh, S.; Scossa, F.; Fernie, A. R., Ultra-high-performance liquid chromatography high-resolution mass spectrometry variants for metabolomics research. *Nat Methods* **2021**.
15. Place, B. J.; Ulrich, E. M.; Challis, J. K.; Chao, A.; Du, B.; Favela, K.; Feng, Y.-L.; Fisher, C. M.; Gardinali, P.; Hood, A.; Knolhoff, A. M.; McEachran, A. D.; Nason, S. L.; Newton, S. R.; Ng, B.; Nuñez, J.; Peter, K. T.; Phillips, A. L.; Quinete, N.; Renslow, R.; Sobus, J. R.; Sussman, E. M.; Warth, B.; Wickramasekara, S.; Williams, A. J., An Introduction to the Benchmarking and Publications for Non-Targeted Analysis Working Group. *Anal Chem* **2021,** *93*, (49), 16289-16296.
16. Hootman, K. C.; Trezzi, J.-P.; Kraemer, L.; Burwell, L. S.; Dong, X.; Guertin, K. A.; Jaeger, C.; Stover, P. J.; Hiller, K.; Cassano, P. A., Erythritol is a pentose-phosphate pathway metabolite and associated with adiposity gain in young adults. *P Natl Acad Sci Usa* **2017,** *114*, (21), E4233-E4240.
17. Chen, Y. J.; Mahieu, N. G.; Huang, X. J.; Singh, M.; Crawford, P. A.; Johnson, S. L.; Gross, R. W.; Schaefer, J.; Patti, G. J., Lactate metabolism is associated with mammalian mitochondria. *Nat Chem Biol* **2016,** *12*, (11), 937-943.
18. Mueller, D.; Heinzle, E., Stable isotope-assisted metabolomics to detect metabolic flux changes in mammalian cell cultures. *Curr Opin Biotech* **2013,** *24*, (1), 54-59.

19. Creek, D. J.; Chokkathukalam, A.; Jankevics, A.; Burgess, K. E. V.; Breitling, R.; Barrett, M. P., Stable isotope-assisted metabolomics for network-wide metabolic pathway elucidation. *Anal Chem* **2012,** *84*, (20), 8442-8447.
20. Llufrio, E. M.; Cho, K.; Patti, G. J., Systems-level analysis of isotopic labeling in untargeted metabolomic data by X-13 CMS. *Nat Protoc* **2019,** *14*, (7), 1970-1990.
21. Forsberg, E. M.; Huan, T.; Rinehart, D.; Benton, H. P.; 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.
22. Huan, T.; Forsberg, E. M.; Rinehart, D.; Johnson, C. H.; Ivanisevic, J.; Benton, H. P.; Fang, M. L.; Aisporna, A.; Hilmers, B.; Poole, F. L.; Thorgersen, M. P.; Adams, M. W. W.; Krantz, G.; Fields, M. W.; Robbins, P. D.; Niedernhofer, L. J.; Ideker, T.; Majumder, E. L.; Wall, J. D.; Rattray, N. J. W.; Goodacre, R.; Lairson, L. L.; Siuzdak, G., Systems biology guided by XCMS Online metabolomics. *Nat Methods* **2017,** *14*, (5), 461-462.
23. Zhu, Z. J.; Schultz, A. W.; Wang, J. H.; Johnson, C. H.; Yannone, S. M.; Patti, G. J.; Siuzdak, G., Liquid chromatography quadrupole time-of-flight mass spectrometry characterization of metabolites guided by the METLIN database. *Nat Protoc* **2013,** *8*, (3), 451-460.