Eboigbe et al. Response to reviewers:

We note our responses are in blue and we use the notations RxCx to define a specifically numbered comment (C) relating to a specifically numbered reviewer (R). RxARx refers to a specifically numbered Author Response (AR) that relates to a reviewer comment.

REVIEWER 1 (R1):

R1C1: The article "Mercury contamination in staple crops impacted by Artisanal Small-scale Gold Mining (ASGM): Stable Hg isotopes demonstrate dominance of atmospheric uptake pathway for Hg in crops" examines Hg in soil, crop, and atmosphere in the vicinity of ASGM operations. Authors measure THg and MeHg concentrations, as well as isotopic measurements of Hg. The paper is well written and data is clearly presented and discussed.

General comments:

The sampling design is well-structured, with multiple environmental sample types. The samples themselves are very valuable, as ASGM sites are understudied in the context of the global Hg cycle. The drawback of the study is that the number of samples per environmental sample type (soil, air, crop) is quite low, as the authors themselves state. Nonetheless, the authors used multiple analytical approaches to make the best use of these samples and the study is valuable for readers of Biogeosciences and researchers in the field, with minor corrections needed.

R1AR1: We appreciate the kind sentiments of the reviewer and also their understanding of the logistical, political, and ethical challenges of undertaking work of this nature. We appreciate the reviewer's understanding of the efforts made to get the most out of the samples that we were able to obtain from the mining and community partners.

Specific comments:

R1C2: Section 2.5.2. In the best practices for the analysis of Hg isotopes (e.g., as outlined in Blum et al., 2017: https://doi.org/10.2138/rmg.2017.82.17), the importance of using Tl internal standard is well explained. The authors do not report using Tl internal standard for their Hg isotope analysis. What is the reasoning behind not using it? In published syntheses of Hg isotopic work, studies conducted without the use of Tl internal standard are often excluded from data analysis.

R1AR2: We appreciate this comment and understand that it can raise concerns to some reviewers. From a strict scientific point of view, Blum and Johnson (2017) recommend use of Tl, but do not rigorously compare to the case of 'not using Tl'. From a more practical point of view, the Blum lab (and the reviewer's JSI lab) uses a Nu Instruments MC-ICPMS, which is an instrument known to be undergo large shifts in mass bias during a 24h session. Use of Tl helps correcting mass bias on the Nu and it is used in all Nu labs. We did our Hg isotope analyses in the Sonke lab on a Neptune MC-ICPMS, well known for its high stability of mass bias. Consequently, Tl is of little use on Neptune instruments. Note that both Blum and Sonke labs recommend sample matrix cleanup before analysis, which also helps avoiding matrix-induced mass bias and this was performed on all samples in this current study. Finally, on a Nu machine, the 12 Faraday cups allow measuring all Hg and Tl isotopes simultaneously. On many Neptunes, the 9 cups are physically restricted in their movement and do not allow collection of 203Tl, 204Hg, and 205Tl simultaneously; the Sonke lab therefore privileges measurement of 204Hg over Tl isotopes. To our knowledge, we are not aware of the exclusion of Sonke lab data, or other Neptune data (analysed without Tl) from review or synthesis works. The numerous high profile studies from both

labs attest to the data quality; see for example Jiskra et al. (2021) which is published in Nature and uses the same methods without Tl-mass bias correction.

Refs: Jiskra, M., Heimbürger-Boavida, L.E., Desgranges, M.M., Petrova, M.V., Dufour, A., Ferreira-Araujo, B., Masbou, J., Chmeleff, J., Thyssen, M., Point, D. and Sonke, J.E., 2021. Mercury stable isotopes constrain atmospheric sources to the ocean. *Nature*, *597*(7878), pp.678-682. https://doi.org/10.1038/s41586-021-03859-8

R1C3: Line 297: The d202Hg value of 0.29 ± 0.98 ‰ for Farm 1 does not indicate low variability as the authors state, ~1 ‰ SD is quite high for Hg isotopes. Please rephrase this discussion accordingly.

R1AR3: We will rephrase this sentence to the following:

"All soil samples exhibited **relatively small (compared to other contaminated soils)** variation in δ^{202}Hg (PS: 0.29 ± 0.98 ‰; Farm1 -0.26 ± 0.43 ‰) and Δ^{199}Hg MIF signal (PS: -0.09 ± 0.12 ‰; Farm 1 -0.07 ± 0.03 ‰) **(Grigg et al., 2018; McLagan et al., 2022; Vaňková et al., 2024)**."

Soils are a notoriously heterogeneous matrix. Thus, we do maintain that the changes, even the MDF, are small compared to what are found at contaminated sites (e.g., McLagan et al., 2022: δ^{202}Hg range ≈ 3-4 ‰; Grigg et al., 2018: range ≈ 1‰; Vankova et al., 2024: range ≈ 1.5‰) and these references have been added to the sentence.

Refs: McLagan, D.S., Schwab, L., Wiederhold, J.G., Chen, L., Pietrucha, J., Kraemer, S.M. and Biester, H., 2022. Demystifying mercury geochemistry in contaminated soil–groundwater systems with complementary mercury stable isotope, concentration, and speciation analyses. *Environmental Science: Processes & Impacts*, *24*(9), pp.1406-1429. https://doi.org/10.1039/D1EM00368B

Grigg, A.R., Kretzschmar, R., Gilli, R.S. and Wiederhold, J.G., 2018. Mercury isotope signatures of digests and sequential extracts from industrially contaminated soils and sediments. *Science of the Total Environment*, *636*, pp.1344-1354. https://doi.org/10.1016/j.scitotenv.2018.04.261

Vaňková, M., Vieira, A.M.D., Ettler, V., Vaněk, A., Trubač, J., Penížek, V. and Mihaljevič, M., 2024. Tracing anthropogenic mercury in soils from Fe–Hg mining/smelting area: Isotopic and speciation insights. *Chemosphere*, *357*, p.142038. https://doi.org/10.1016/j.chemosphere.2024.142038

R1C4: Lines 319-324: This paragraph seems out of place, fitting more to the introduction part into the justification for the chosen experimental design/methods (or somewhere else, but not here).

R1AR4: We believe this is the appropriate place in the manuscript because it provides direct comparison to another study of Hg in crops, specifically cassava, that examined inner peel, outer peel, and flesh. We use this study to highlight that this would be an advisable method of improving upon the work we have done (in future studies) to more conclusively identify if root epidermis/cortex does provide an effective barrier to Hg uptake and translocation to above ground tissue. We highlight the fit here by stating that the paragraph prior (lines 303-316) is specifically discussing the different uptake pathways (soil-to-root vs air-to-foliage).

R1C5: Lines 397-398: The subtraction of MDF for the soil-to-shallow roots (from Yuan et al 2022) is explained in supplementary material section S4. But until carefully reading the supplementary section, this subtraction is quite unclear to the reader, disturbing the reading flow. The authors should add a succinct explanation for this subtraction in the main text in lines 397-398.

R1AR5: This is a good suggestion, and we will make adjustments to more clearly define the two-end-member mixing model. Nonetheless, rather than adding this to the discussion, we deem it more appropriate to include this in the methods section (Section 2.6). Lines 227-229 will be updated to the following:

"A two-endmember mixing model was used to quantitatively determine source pathways for Hg in internal crop tissues according to Equation S4.1. The δ202Hg values of foliage for each crop are used as the first endmember: air-foliage uptake pathway. Similarly, the soil-root endmember must be the δ^{202}Hg signature of Hg immediately after uptake to the roots. Hence, the mean δ202Hg value for Farm1 soils minus the soil-to-shallow roots (roots above 150cm) MDF (ε^{202}Hg: -0.35) taken from Yuan et al. (2022) is used. Details of this two-endmember mixing model and the data used in the derivation of the endmembers are provided in Section S4."

R1C6: Lines 221-225, 392-395: The PTD analysis for Hg is not very robust. Therefore, authors should note that the conclusions they draw from these analyses are speculative (also in the discussion of the results, lines 392-395). Additionally, the term "speciation" is used too loosely, as speciation is, by definition, qualitative or quantitative measurement of chemical species, while in this case, there is no information about what the measured species are. Please replace "speciation" with "analysis" or "PTD analysis" in places where referred to PTD analysis, throughout the text.

R1AR6: We agree that PTD analyses are not highly robust, which is why we consider these analyses "quantitative and complementary" (lines 223-224). We also agree speciation is not the best terminology and we will change the sub-heading to "2.5.3 Hg speciation/fractionation analyses" and references to "speciation" to "speciation/fractionation" and "PTD speciation analyses" to "PTD analyses" as suggested.

R1C7: Lines 484-488: Would it make sense for future research to include some livestock near ASGM farms? Hg isotopic signatures in the livestock tissues could tell a very interesting story too. It could be worth mentioning in the text.

R1AR7: We concur will add the following statement at the end of the section the reviewer refers to:

"Adding Hg stable isotope analyses to any future work around Hg in livestock in ASGM areas could provide valuable insight into the biogeochemical processes involved."

R1C8: Section S3 and Figure S8: Authors mention certain reference standards were used for PTD analysis. Where can the reader see these desorption results for reference standards? The point of using these standards is to see if the desorption peaks of samples overlap with some desorption peaks of Hg standards. Why are the standards not shown then?

R1AR8: The standards we referred to here are from the three referenced studies [Biester and Scholz (1996), Mashyanov et al. (2017), and McLagan et al. (2022)]. We totally agree with the reviewer that when trying to identify specific fractions or species, it is best to cross reference to these standards. However, we stress that we deem these analyses "quantitative and complementary". The purpose of including the data presented in Section 8 (and referred to in lines 392-395) was to highlight the dual peaks observed in peanut and maize roots. This provides complementary data that support the notion of two Hg pools in the roots: one from direct uptake from soils (like in the epidermis/cortext) and another translocated to the roots from the air-foliage uptake pathway. These dual peaks that are clearly seen for peanut and maize roots were not

observed in other plant tissues (Figure S8.1). We don't deem it necessary to compare to standards as these release peaks fall into the range of 190-300°C where many species/fractions are released (McLagan et al., 2022) and it would not be possible to accurately match these peaks to any specific species/fractions.

R1C8: Table S7.3: It would be clearer if authors wrote "Peanut soil 1, Peanut soil 2, …" to make it clear that these are soil analyses and not crop analyses.

R1AR8: All samples in Table S7.3 will include soil in the sample labels in the revised version of the SI. The same will also be done for Table S7.2

Technical corrections:

R1C9: Line 31: "significantly high" should be "higher"

R1AR9: We will change this to "significantly higher" as this is a statistically tested difference.

R1C10: Line 40: "soil derived" should be "soil-derived"

R1AR10: This will be corrected.