Response letter

Respond to Anonymous Referee #1

General comment/Overview

The manuscript by Diao et al. investigates different fractionation processes caused by H-exchange with organic material, sublimation and evaporation processes during cryogenic vacuum distillation of plant samples. The manuscript is certainly of high interest to scientists working with CVD and provides valuable insights into different fractionation processes. However, I feel that the study has certain shortcomings, which prevent the certainty with which some conclusions were made. I also feel that the description of the experiments + results and discussion could be improved to enhance the clarity of the manuscript. Please find my major and minor comments below.

Response: Thank you for your comments. We have revised the manuscript according to your comments. Please find the point-to-point response below.

Specific comments

Major

(1) L. 53-58: The authors conclude that there are two critical H-exchange steps: the first one being during rehydration, the second one being during CVD. Yet, according to the authors only the second one is relevant/"of interest" (L. 56), because in a natural set-up, samples are not rehydrated. However, if you think about that plants are also de- and rehydrating naturally, e.g. over the course of a season, this step and a possible H-exchange might also occur under natural conditions and should be taken into account, when comparing source water with plant water. The question is of course, if this exchange is negligible under natural conditions under high plant water fluxes.

Response: Thank you for pointing this out. We agree that studying "de- and rehydration" in natural conditions could be a highly interesting topic. However, we can only speculate whether this is true or not, as we cannot provide evidence for these phenomena from our experimental study as we only work with ex situ wood samples and different compounds in our rehydration experiments. In our opinion, the influence of natural de- and rehydration on the difference between plant and source water should be assessed using in situ tree xylem water measurements. We added a comment to the discussion (Line 244–247).

(2) Drying procedure of sample material: why did you dry your material only at 60°C and for 24h? I know for organic materials it is common to use 60-70°C. However, I feel that for such an experiment a completely dry sample (105°C) would have been of immense importance. At least it should have been dried for 48h. This has to be addressed somehow and could potentially compromise your results. Same issue in L. 128.

Response: Thank you for pointing this out. In our study, the drying procedure of sample material was only for experiment 1. Because the dried samples in the experiment 2 were rehydrated with excess of water for isotopic equilibration and no dried materials were used in the experiment 3 and 4.

For the stem and twig pieces, the original materials were first dried at 60°C for 24 h in order to cut them into pieces and then for weighing. Then, 200 mg of pieces were weighed in Exetainers and dried again at 60°C for 24 h. In total, the sample were actually dried for 48 h as suggested by the reviewer. For the powdered materials, they have already been dried and stored in a dry environment prior to the experiment. Those samples were used as isotope reference standards for δ^2H analysis in our laboratory (Schuler et al. 2022). So, the powdered materials were further dried once in the Exetainers at 60°C for 24 h. We added some sentences to clarified this. See lines: 106–107; 112; 118; 122–126.

We also would like to note that it is very difficult to have completely dry samples, even if at 105°C for 48 h. That's because moisture in the air will be absorbed within seconds by the dry samples when the samples are taken out from the oven and when opening the cap for injecting the reference water.

To address potential problems regarding our sample drying procedure, we conducted a test using ca. 200 mg stem pieces, stem powder and caffeine at two different drying procedures (60°C, 48 h; 105°C, 48 h). The samples were weighed every 12 h. After 48 h, we opened the cap of Exetainer for 5 s in the lab to simulate the procedure of injecting the reference water. Then the cap closed and samples were weighted again. We added a description of this new test at lines: 132–139. The results are shown in the table below.

Material	T (°C)	The percentage of moisture removed (%)				Water vapour absorbed in 5s (%)
		12h	24h	36h	48h	
Twig piece	60	5.55±0.06 a	5.60±0.05 b	5.74±0.05 c	5.84±0.06 c	0.72±0.11
	105	8.68±0.21 a	8.89±0.12 a	8.91±0.16 a	8.89±0.08 a	1.80±0.08
Stem powder	60	5.05±0.04 a	5.21±0.04 b	5.31±0.07 c	5.33±0.05 c	0.72±0.06
	105	8.28±0.19 a	8.32±0.25 a	8.28±0.18 a	8.28±0.10 a	1.70±0.13
Caffeine	60	0.44±0.03 a	0.41±0.09 a	0.43±0.04 a	0.43±0.04 a	0.67±0.04
	105	1.50±0.15 a	1.59±0.11 a	1.62±0.09 a	1.58±0.04 a	1.53±0.09

Note: Different letters in each row indicate a statistically significance difference at $p<0.05$ (T-test)

The results show that all the tested materials were dried to a constant weight after 36 h and 12 h at 60°C and 105°C, respectively. However, more moisture was removed at 105°C, especially for twig pieces and stem powder. Noticeably, after opening the cap for 5 s, the samples which were dried at 105°C absorbed more lab water vapour than the samples which were dried at 60°C. These results suggest that a complete drying is very difficult, regardless of whether the samples were dried at 60°C or 105°C. However, the difference of 1-2 % in removed moisture after reabsorption between 60°C and 105°C should not affect the isotopic results of our study, because strongly depleted reference water was used to amplify the differences. It should also be noted, that the higher uncertainty at lower amounts of water is observed across all our experiments, again showing that the remaining moisture in the dried material is not a major driver of our overall results. We have added this table to the supplement, see the new Table S2.

Corresponding texts were added at lines: 132–139; 255–264.

(3) Over-/undersaturation of sample material: So if some water amounts were not able to fully saturate the material and others oversaturated the material, I would be really careful with the conclusion drawn from this experiment, as it does not reflect "real" plant samples. If relative water content is not an issue, why did you not choose different weights of your samples material to avoid over-/undersaturation?
Response: First, the main objective of Experiment 1 was to investigate the phenomena of H-exchange of different plant material and compounds across a range of water to biomass ratios. It was not intended to reflect a natural gradient of relative water content. Second, given that the isotopic variation along the gradient were similar across material and reference water (without any organics), led us conclude that AWA rather than RWC is important. We thus designed experiment 2 to demonstrate this point. We added text indicating that we intended to generate a range of water/biomass ratios (Line 120–121). As a general consideration, we would also like the to mention that controlled experiments, as we did, always have some advantages (like testing for certain effects), but also disadvantages (like not necessarily being representative for "real" samples).

(4) In general, there are a lot of assumptions in the material and methods sections, such as "Thus, by the end of the rehydration, the isotope ratios of water in the small stem segments are assumed equal to the isotope ratio of the reference water after rehydration (δref after rehyd) and not to be equal to the original reference water (δref)." (L. 137-139). Can you be sure about this? This should be considered in the discussion.
Response: Good point. Indeed, we cannot be 100% sure whether the exchangeable H in the different material was in full isotopic equilibrium with the water, as we can't determine the isotope ratio of the exchangeable H in the material after rehydration. However, at least some percentage of the exchangeable H will exchange with the water under the chosen experimental conditions (25°C), particularly those that are not linked via hydrogen bridges (Sepall and Mason, 1961). As indicated by experiment 1, the input of hydrogen derived from organic material on the reference water can be significant at lower water amounts. We thus note that the original isotope ratio of the reference water has changed and that the isotope ratio of the reference water after exchange reflect a better reference. Several sentences have been added at the end of the second paragraph of section 3.2 in the Discussion to address this issue (Line 296–303).

(5) Experiment 3: I am afraid I don't really get the whole experiment. From what I understand you were only interested in the effects after the water has been extracted from the sample, thus the freezing in liquid nitrogen. But why do you then write "before the extraction started"? I guess you left the water in the liquid nitrogen for a certain amount to simulate an extraction? Why not also freeze the reference water at -20°C and extract it the way as in experiments 1 and 2? This would also allow a statement of the effect of heating the water and potential evaporation effects before freezing the water. There is certainly some clarification needed for this experiment.
Response: Sorry for causing this misunderstanding. In experiment 3, the reference

water was also frozen, simply in the U-shaped water collection tube with liquid nitrogen, not in the Exetainer in a freezer at -20°C (as in Experiment 1 and 2). So, before the extraction started the water was frozen, then sample tubes and the collection tubes were put in the 80°C water bath and liquid nitrogen, respectively, and vacuum applied. That's why we wrote "before the extraction started". We have revised section 2.3 to clarify those unclear descriptions (Line 165–167).

With respect to "freeze the reference water at -20°C and extract it the way as in experiments 1 and 2", we have already done that experiment. This is actually our "pure reference water extraction" in Fig 1b. A description of this can be found in the last paragraph of section 2.1: "As a control, the experiment was repeated without any material by adding only the range of reference water into the vial." (Line 129–130). So, in experiment 3, we went in-depth to see what happened during the sublimation of the frozen extracted water.

(6) As four experiments were conducted, it is sometimes really confusing for the reader to follow the argumentation, as you always have to go back to the methods to see, which experiment exactly the authors are talking about. Potentially, Fig. S1 could be moved into the manuscript, but in a clearer manner with a clearer description.

Response: We have thoroughly revised Fig. S1 to make the individual experiment more understandable. As suggested, we moved Fig. S1 into the manuscript, see the new Fig. 1.

(7) Figure 1-3: There are some undiscussed effects in Figure 1 and 2:

Fig. 1: why do powder materials drop below the reference line at high water amounts? This remains unexplained

Response: We have discussed the difference in Δ^2H between pieces and powdered materials when AWA > 600 µl in third paragraph of section 3.1 (Line 247–255). A possible explanation is that the isotopic compositions in the exchangeable H of pieces and powdered samples were different.

Fig. 2: for H, your dH is negative for water samples larger than 400 microl, for 18O, they are still clearly above the reference line. This should certainly be discussed.

Response: The difference between the Δ and the reference line might be explained by analytical uncertainties in δ_{ref}, causing offsets in Δ to the reference line. We therefore added the standard deviation of the reference line in the original Fig. 1 and Fig. 3 to express this uncertainty (now they are Fig. 2 and Fig.4). In these figures, all the δ values were referenced against one single δ_{ref} value (an average value from repeated measurements of the reference water). The standard deviation of the reference line was not added in the original Fig. 2, because each point in this figure has its own reference value (i.e., $\delta_{ref\ after\ rehyd}$); the standard deviation of the reference line was not added in the original Fig. 4 because δ_{CVD_ave} was taken as a reference.

Fig. 3: for H in your sublimation and evaporation test, dH is steadily decreasing with absolute water amount. You argue that this is analytical uncertainty, however it is a quite

clear pattern, which remains undiscussed. Also, the fit for dH in the sublimation experiment is somewhat off and not well representative of the data.

Response: Sorry for causing this misunderstanding. We agree that the Δ^2H is steadily decreasing with absolute water amount and that this pattern is very clear. The analytical uncertainty argument was made on the negative values in Δ^2H at AWA > 600 µl, not on the overall decreasing pattern. See the text: "We suppose that the negative Δ^2H and $\Delta^{18}O$ values for the AWA > 600 µl were purely caused by the analytical uncertainty, because no incomplete extraction could occur given that the reference water was added directly into the collection tube." (Line 333–335)

To be consistent, we used the same model fits for the results from all the experiments. The reason for the fit for Δ^2H in the sublimation experiment not well representative of the data is probably that the Δ^2H of samples at 100 and 200 µl were somewhat shifted up a little bit compared with the fitted line.

(8) Recommendation to extract more than 600 microl: From the results, I cannot clearly see, why the authors recommend at least 600 microl of water for extraction. In Figure 2 you clearly see that at 600 microl there is still quite a substantial offset for 2H and 18O, in Fig. 3 this is also the case for the evaporation experiment and 2H. In Fig. 4 there is a negative offset at 600 microl for 18O. This should be discussed in more detail.

Response: We agree that the exact level of 600 µl is somewhat arbitrary, but it was chosen because (i) the discrepancy among Δ values at AWA > 600 µl were much smaller than those < 600 µl; (ii) despite there are still offsets when AWA > 600 µl, the offsets were relatively small and within the range of systematic error of the isotope ratio of the reference water (i.e., standard deviations are now given as grey bars). Therefore, extracting a minimum amount of 600 µl of water could reduce large isotopic enrichments. We have added some sentences to discuss this in more detail: See lines: 231–232; 430–431.

(9) Plant material: almost all experiments were only conducted on Larix decidua. This should be included in the discussion, as the results could be completely different for other woody species and especially for herbaceous plants.

Response: Given that hydrogen and oxygen fractionations were observed at lower amounts across all experiments, we believe that our main conclusion is not necessarily related to the physical and chemical composition of the sample. However, we agree that this could play a role and that more systematic studies testing differences among woody and herbaceous plant species should be useful for ruling out further issues. We have included this in the discussion, see lines 358– 364.

Minor

(1) L. 16: Would be good to mention what the tested organic materials exactly were. "organic materials" could be anything

Response: The sentence was revised to "…using isotopically depleted water, water at natural isotopic abundance, woody materials with exchangeable H, and organic materials without exchangeable H (cellulose triacetate and caffeine)" (Line 16–17).

(2) L.32-35: This is certainly questionable these days. I would be more careful with this statement

Response: Agreed. We have used the word "generally" in this sentence to show the conservativeness of this statement (Line 33–36).

(3) L. 35: replace faithfully. Do you mean actually?

Response: Done. "faithfully" was replaced by "actually".

(4) L. 47-49: This sentence is hard to read, please consider rephrasing it.

Response: The sentence was revised to "In woody tissue, oxygen-bound H atoms can exchange with those in surrounding liquid water and water vapour" (Line 48–50).

(5) L. 53: delete "exactly"

Response: Done.

(6) L. 56: Consider deleting the first "the H-exchange of interest". "Only the latter is of interest, because…"

Response: We didn't delete the first "H-exchange of interest", see our response to the major comment #1. The sentence has been revised to "Only the latter is of interest, because it is the H-exchange process that theoretically affects the isotopic composition of CVD extracted water from actual plant samples."

(7) L. 67-69: please consider rephrasing this sentence, hard to read

Response: The sentence was revised to "If this were to be true, what really matters would not be the RWC, but rather the absolute amount of water being extracted…" (Line 69–70).

(8) L. 80: check for consistency of "vapour" or "vapor" in the whole manuscript (e.g. L. 71)

Response: We replaced "vapor" by "vapour" in the whole manuscript.

(9) L. 100: please give the exact values for 2H and 18O

Response: Done. $\delta^{2}H$: -465.9‰, $\delta^{18}O$: -174.2‰.

(10) L. 103-110: why did you choose Larix decidua? Your conclusions could only be valid for this one species due to wood anatomy, etc.

Response: Twig pieces (in Experiment 1) and fresh twig segments (in Experiment 4) of *Larix decidua* were used because we intended to use materials from plants with the same genus. For instance, stem pieces, stem powder, and cellulose were all derived from a *Larix sibirica* in Siberia and thus from the same material and origin (Schuler *et al.,* 2022). While wood anatomy varies among species, the chemical composition/structure of cellulose in wood is relatively conserved. Also, given that hydrogen and oxygen fractionations at lower amounts were observed across all

experiments, we believe that our main conclusion is not necessarily related to species-specific difference in the physical and chemical composition of the sample. See also our response to the major comment #9.

(11) L. 117: I think you should justify why you used exactly 200 mg
Response: The amount was chosen because of practical consideration (e.g., size of vials used in extraction line) and because it therefore generally approximates samples sizes in other experiments. We added: "The amount of material was chosen because of practical considerations and to generate a range of water/biomass ratios" (Line 120–121).

(12) L. 136-137: This is something you suspect, but do not know for sure
Response: Yes. We assumed that a 24 h rehydration is enough to reach an equilibration between water and sample. This assumption was made based on the experimental design of Chen et al. (2020), who rehydrated their samples for 24 h at 25 Celsius. A discussion about this issue has been added in the second paragraph of section 3.2 (Line 296–303). See also our responses to the major comment #4.

(13) L. 157: 2h is kind of an unrealistic timeframe for CVD, but I guess for investigating the evaporation effect it is okay
Response: Yes, the 2h was just for investigating the evaporation effect. We believe that the magnitude of evaporative fractionation depends on how long evaporation is allowed and on the temperature in the CVD extraction, but the water amount effect on the evaporative fractionation should be evident in the CVD extractions of our study.

(14) L. 162-171: from the description it is not clear why you needed this experiment
Response: We have added a sentence at the beginning of this paragraph: "In order to test whether the CVD-related isotope fractionation, which were observed with labelled water, can also be seen at natural isotope abundance…" (Line 179–180).

(15) L. 173: Did you use any material to cover your samples, e.g. glass wool, to avoid particles to be drawn into your vacuum pump?
Response: We have added "The sample tubes were blocked by PP fiber filters (Nozzle protection filter, Socorex Isba SA, Ecublens, Switzerland) to avoid particles being drawn into the U-tubes with the extracted water or the vacuum pump" (Line 198–199).

(16) L. 205: should be Fig 1a & b
Response: Done.

(17) L. 220-221: "However, if this was true"
Response: Done.

(18) Fig. 1: reference line for d18O is missing
Response: Done. It has been revised.

(19) L. 234: causes?
Response: Done.

(20) L. 243-244: I don't fully understand how you can exclude that there is a dynamic exchange during the extraction. You do not know for sure, what H is bound in your samples, although you allowed it to incubate with water of known isotopic composition. Also, the conditions during extraction are also different. I would be careful with such conclusions.
Response: Sorry for the unclearness of this paragraph. We were trying to say that the dynamic H exchange during extraction is not the main reason causing the offsets in Δ values in our study and the negative offsets in Δ values as shown in Chen *et al.*, 2020. Our conclusion is based (i) on experiment 1 showing that the actual H-exchange effect results in a strongly positive but not negative offsets; (ii) on experiment 2 denoting that Δ patterns were similar to material without exchangeable H (e.g. reference water in experiment 1). We responded to this in major comment #4 and the minor comment #12, a short discussion on this has been added for clarity. See lines: 280–286.

(21) L. 250: The trend for 18O is very similar to the one for 2H. If you remove the outlier at d18O = 16 and RWC = 36, this will also be a linear trend. However, the relationship is still weak ($R^2 = 0.28$) and your data points are strongly scattered
Response: Good observation. We've tried removing that outlier in $\Delta^{18}O$; however, the relationship is still weak ($R^2 = 0.29$). We therefore kept the original data. We revised this sentence to "… we found a weak, positive linear trend in …".

(22) L. 254: Here I disagree. You should use a linear mixed effect model and treat stem segment size as random effect, if stem segment size is not an explanatory variable.
Response: Thank you for pointing this out. We have tried to use a linear mixed effect model with stem segment size (7 levels) as a random effect: lmer(Δ~RWC+(1+RWC|Stem segment size),data = data). The results show that the relationship between Δ and RWC was still not significant. However, because of the low amount of replicates (n = 3) per stem segment size, we didn't include these results into the manuscript. Besides, we tested the relationship between Δ and RWC using a linear mixed effect model with AWA as a random effect (two levels: AWA < 400 µl and AWA > 400 µl; i.e., the red and blue data points in the original Fig. 2). The results still showed a nonsignificant relationship. The results were put in the supplement. See Table S3.

(23) L. 273: what experiments? I fear at this point the reader has already forgotten, which one experiment 3 was
Response: We added a sentence to remind the reader of experiment 3 as follows: "In this experiment, different amounts of reference water were either injected directly into the collection tubes then subject to CVD extraction, or subject to controlled evaporative conditions in a climate chamber for 2h" (Line 328–329).

(24) L. 276: suggest? Suppose? Instead of think
Response: "think" has been replaced by "suppose that".

(25) L. 301-303: I think this is true for all your experiments, especially because they were mainly conducted on Larix
Response: We agree. We have added a sentence to clarify that even though we conducted experiments mainly on *Larix*, we expect our findings to have more general applicability. See our response to the major comment #9.

(26) L. 317: I think -7 and +10 ‰ are quite substantial variations
Response: Good point. We have checked the accuracy and precision of the isotope analyses for all experiments and found that the precision of quality control standard of the batch with the extracted tap water (previous Fig. 4a, blue dots) was low, explaining these substantial variations. Therefore, we repeated the "tap water" experiment and the analysis of the samples. The results are now given in the new Fig. 5, showing that Δ^2H of pure tap water also followed the inversely proportional pattern as found in other experiments.

(27) L. 321: please elaborate further on why the pattern was evident for 18O, but not 2H
Response: We don't really know why the pattern was evident for 18O but not 2H, we can only speculate that this was caused because of a H-exchange between twig water and organic material, which have an isotopic similarity. We added "While we found no clear explanation of the absence of the AWA-dependency for Δ^2H in this experiment, we can only speculate that the expected pattern was hidden by processes shaping hydrogen but not oxygen isotopic variations such as H-exchange effects between water vapour and extracted water with similar isotopic compositions" (Line 408–411).

(28) L. 322-327: this paragraph is controversial and confusing. You kind of invalidate your own results.
Response: Thank you for pointing this out. We were trying to say that the dampened pattern obtained from the tap water extraction compared to the reference water extraction was not because they had different initial isotope composition (this paragraph), but more likely because of the vapour exchange with the lab air (the next paragraph). The first sentence of this paragraph made this paragraph confusing, but we were not intending to invalidate our own results. Therefore, we have revised this paragraph to improve its clearness and the connection with the next paragraph. See lines: 385–405.

(29) L. 328-331: You could also use a different argumentation: the tap water and twig water is closer to the laboratory water vapour than the isotopically depleted water you used and thus, in these conditions of your laboratory, the initial signature of your extracted water determines the magnitude of isotopic fractionation
Response: Thanks for this suggestion. We have added: "It is therefore possible that the

difference in Δ between reference water and tap water extractions is caused by the exchange of extracted water with the water vapour in the laboratory. The isotope ratios of the laboratory water vapour are closer to that of the tap water and twig water, but differ greatly compared to the depleted reference water. Therefore, in these conditions, the initial isotopic signature of the extracted water determines the magnitude of the isotope signature of the extracted water after the exchange with laboratory water vapour and therefore the observed Δ values." See lines: 395–401.

(30) L. 334: …but there might be a significant correlation in the range of 45 – 53%, as you only have two samples at 57%. Please check this. This could change your conclusion.
Response: After removing the two samples at 57%, the relationship between Δ^2H and RWC was still not significant ($p = 0.2$), and the relationship between $\Delta^{18}O$ and RWC was significant ($p = 0.03$), but still weak ($R^2 = 0.21$). Given that we have no good reason to remove those two samples and that the results were not significantly changed by removing them, we kept the original data. Thus, the results still support that RWC is not a reliable factor for correcting isotopic offset caused by CVD.

(31) Fig. 4: please write Larix decidua in italic
Response: Done.

(32) L. 350: check the references in the brackets, there seem to be a few extra brackets.
Response: Done. The extra brackets of years in those references were removed.

(33) Fig. S1 is lacking a clearer description of the experiments
Response: We have improved the original Fig. S1 and moved it into the manuscript. It is now the new Fig. 1 See also our response to the major comment #6.

Respond to Anonymous Referee #2

General comments

In the manuscript 'On uncertainties in plant water isotopic composition following extraction by cryogenic vacuum distillation' by Haoyu Diao et al. the authors investigated the biases caused by CVD extraction of plant samples. In different experiments they tested the influence of H-exchange effects, absolute water amount, and evaporation and sublimation enrichments on the isotopic composition (d18-O, d2-H) of the extracted water. The manuscript is well structured and nicely written. The topic fits well to the scope of the journal and appears to be of interest for isotope hydrologists. However, the discussion could be more comprehensive to make things clearer.

Most of my comments and remarks match those of Referee 1, who raised many important points already.

Response: Thank you for your comments. We have revised the manuscript according to your and the comments of Reviewer 1, particularly the discussion, as suggested. Please find the point-to-point response below.

Specific comments

(1) L. 128 ff: I suspect that drying at 60°C for 24h might eventually be too low and/ or too short. Was this procedure tested in advance? Did you check e.g. by weighing the samples after another 24h whether there was still some weight loss or not? Additionally, the 'appropriate' procedure could also be different for the twigs and the small segments. If your stem segments were not completely dry, the residual water will of course be strongly enriched, which will affect your results.

Response: Thank you for pointing this out. As explained in the response to reviewer 1, in our study, the drying procedure of sample material was only for experiment 1. Because the dried samples in the experiment 2 were rehydrated with excess of water for isotopic equilibration and no dried materials were used in the experiment 3 and 4.

For the stem and twig pieces, the original materials were first dried at 60°C for 24 h in order to cut them into pieces and then for weighing. Then, 200 mg of pieces were weighed in Exetainers and dried again at 60°C for 24 h. In total, the sample were actually dried for 48 h as suggested by the reviewer. For the powdered materials, they have already been dried and stored in a dry environment prior to the experiment. Those samples were used as isotope reference standards for δ^2H analysis in our laboratory (Schuler et al. 2022). So, the powdered materials were further dried once in the Exetainers at 60°C for 24 h. We added some sentences to clarified this. See lines: 106–107; 112; 118; 122–126.

We also would like to note that it is very difficult to have completely dry samples, even if at 105°C for 48 h. That's because moisture in the air will be absorbed within seconds by the dry samples when the samples are taken out from the oven and when opening the cap for injecting the reference water.

To address potential problems regarding our sample drying procedure, we conducted a test using ca. 200 mg stem pieces, stem powder and caffeine at two different drying procedures (60°C, 48 h; 105°C, 48 h). The samples were weighed every

12 h. After 48 h, we opened the cap of Exetainer for 5 s in the lab to simulate the procedure of injecting the reference water. Then the cap closed and samples were weighted again. We added a description of this new test at lines: 132–139. The results are shown in the table below.

Material	T (°C)	The percentage of moisture removed (%)				Water vapour absorbed in 5s (%)
		12h	24h	36h	48h	
Twig piece	60	5.55±0.06 a	5.60±0.05 b	5.74±0.05 c	5.84±0.06 c	0.72±0.11
	105	8.68±0.21 a	8.89±0.12 a	8.91±0.16 a	8.89±0.08 a	1.80±0.08
Stem powder	60	5.05±0.04 a	5.21±0.04 b	5.31±0.07 c	5.33±0.05 c	0.72±0.06
	105	8.28±0.19 a	8.32±0.25 a	8.28±0.18 a	8.28±0.10 a	1.70±0.13
Caffeine	60	0.44±0.03 a	0.41±0.09 a	0.43±0.04 a	0.43±0.04 a	0.67±0.04
	105	1.50±0.15 a	1.59±0.11 a	1.62±0.09 a	1.58±0.04 a	1.53±0.09

Note: Different letters in each row indicate a statistically significance difference at $p<0.05$ (T-test)

The results show that all the tested materials were dried to a constant weight after 36 h and 12 h at 60°C and 105°C, respectively. However, more moisture was removed at 105°C, especially for twig pieces and stem powder. Noticeably, after opening the cap for 5 s, the samples which were dried at 105°C absorbed more lab water vapour than the samples which were dried at 60°C. These results suggest that a complete drying is very difficult, regardless of whether the samples were dried at 60°C or 105°C. However, the difference of 1-2 % in removed moisture after reabsorption between 60°C and 105°C should not affect the isotopic results of our study, because strongly depleted reference water was used to amplify the differences. It should also be noted, that the higher uncertainty at lower amounts of water is observed across all our experiments, again showing that the remaining moisture in the dried material is not a major driver of our overall results. We have added this table to the supplement, see the new Table S2. Corresponding texts were added at lines: 132–139; 255–264.

(2) I'm not surprised that the differences in the isotope data are dependent on AWA rather than on RWA (e.g. L. 224, L.252). Possible reasons could be

- very small water droplets are exposed to ambient air during collection (L. 180-181)
- for small samples the sample volume might be too small relative to the size of the extraction line i.e. the volume which is filled with water vapour. Can you estimate the size (volume) of your system?

and should be considered in the discussion.

Response: We agree with you regarding the two possible reasons of the observed dependence on AWA rather than on RWC. We discussed these two points in the second paragraph of section 3.3. See "… a bi-directional exchange between water droplets and … water vapour of the laboratory …" and "… the smaller the water droplets, the larger the ratio of water vapour volume to water droplets volume in the water collection tube …" (Line 343–345; 349–350).

As suggested, we estimated the volume of the whole system (including all tubes

and lines) to about 3500 cm^3. We suppose that the two processes mentioned above could mostly occur in the U-shaped water collection tube (Figure S1) after it is detached from the extraction line. Therefore, we added the volume of the U-shaped water collection tube (78 ml) to the description of Figure S1, illustrating the CVD-extraction line. We have also mentioned that the system volume may have an effect on the isotope fractionation during CVD extraction at the end of section 3.3 (Line 362–364).

(3) L. 322: You assume that the effect of small amounts is different, depending on the isotope ratio? At least for isotope values in the range of your study it should not be in a measurable range.
Response: Sorry for causing this misunderstanding. We were trying to say that the dampened pattern obtained from the tap water extraction compared to the reference water extraction was not because they had different initial isotope composition (this paragraph), but more likely because of the vapour exchange with the lab air (the next paragraph). The first sentence of this paragraph made this paragraph confusing, but we were not intending to invalidate our own results. Therefore, we have revised this paragraph to improve its clearness and the connection with the next paragraph. See lines: 385–405.

(4) Figure 4: How can the huge differences in SD between D18-O and D2-H (Fig. 4a and c) be explained?
Response: We have checked the accuracy and precision of the isotope analyses for all experiments and found that the precision of quality control standard of the batch with the extracted tap water (previous Fig. 4a, blue dots) was low, explaining these substantial variations. Therefore, we repeated the "tap water" experiment and the analysis of the samples. The results are now given in the new Fig. 5, showing that Δ^2H of pure tap water also followed the inversely proportional pattern as found in other experiments.

Technical corrections
(1) Throughout the manuscript: isotope fractionation should almost everywhere be singular not plural, please check.
Response: Done.

(2) L. 71: 'liquid-vapour' instead of 'liquid-vapor'
Response: Done.

(3) L. 89: and L. 263: 'induced' instead of 'introduced'?
Response: Done. We replaced "introduced" by "induced."

(4) L. 310/ 311: '…allowed for investigating…' instead of '…allowed us to investigate…'
Response: Done.

(5) L. 340: …mean values ± 1 SD like in Fig. 3?
Response: Yes. Done.

(6) Figure S1: Typo in legend of Experiment 2 ‘Abosolute water amount’
Response: Done. This typo has been fixed.