Answers to Referee #3

We appreciate so much all your comments and constructive questions that make us to clarify and improve the quality of the paper. Thanks.

General comments

1) The measurements are very difficult to connect to air pollution and air pollution impacts due to the complexity and of the transfer function between the atmosphere and bioaccumulation on the plant. Although the derivation or measurement of the transfer function is clearly beyond the scope of the current paper, the authors should offer some recommendation of future work that is needed to allow such measurements to better address atmospheric pollutant concentrations or the impacts of atmospheric pollutants.

This comment reminds us how far is still our understanding of how the amount of air pollutants accumulated by our biomonitor is precisely linked to their concentrations in the atmosphere. It is not to justify, but we are not alone on this. The now voluminous literature on air pollution biomonitoring has only a few attempts on that direction with different degree of complexity and success. They usually deal with single air pollutants, which simplify many matters compared to multipollutant studies. Complexities arise mainly from the many processes continuously adding (to) and subtracting air pollutants from a biomonitor. An air-to-biomonitor pollution transfer function should contain as many terms as the number of relevant identified processes, which must be somehow quantified. Not an easy task.

As we describe in the manuscript, our main objective was much more modest. This was our first glance at a quite complex region in terms of air pollution. We basically looked for spatial distribution patterns because they are indicative of differential exposition to particular pollutants. Once we found them, we are now in a better position to look for other aspects, such as the one you mention, and also check for possible deleterious effects on the biomonitor. We can also work with only one or two selected pollutants to be able of adding more detail; for instance, toxic metals or PAH showing high signal for particular sources.

Of course, we want to find how to infer air pollution levels from the concentrations in our biomonitor. This practical purpose implies research approaches like the one you mention. We are now planning for measuring air vs. biomonitor correlations by measuring pollutants at selected sites (both for *in situ* and transplanted biomonitors) and dry and wet deposition with co-located active or passive monitors, as well as site meteorological measurements. Chemical differentiation by particle size would be very helpful, since this size greatly determines dispersion and deposition.

2. Given the importance of dust in the accumulation of metals on or within the plant, the authors need to address the assumption that the accumulation of pollutants is bioaccumulation and not just surface deposition. Would similar results for metals or PAH have been obtain for the crustal material if an inert object shaped as the plant was deployed in the location of the plant?

We used the bioaccumulation concept in its widest sense; i.e., the amount of each pollutant remaining in and on our plant at the moment of being sampled. The surface of this plant is adapted to retain particles deposited or impacted on it, as well as to absorb water. It uses the minerals dissolved from these particles as nutrients. When there is particular interest on separating adsorbed and absorbed materials, usually when looking at biological effects from pollutants, the samples are washed with water or acidic solutions (for metals), or solvents (for PAH). The amount of pollutants on the surface is usually positively correlated with the amount incorporated in the tissues.

We would expect some similarity (degree of positive correlation) between the amounts of pollutants retained by a biomonitor and its inert surrogate. However, some differences could be expected from the active retention of some metals by the biomonitor, especially when they are metals biologically useful for the plant.

3) It is not clear for me from the paper if the uptake of C and N are dominated from the uptake of carbon dioxide and nitrogen fixation, which is suggested in the text but not necessarily supported by the data. If this is the case for either of these components, than the enrichment of C and N isotopes are not really of interest to the air pollution community and are more relevant to the global biogeochemical cycle.

Most, if not all, constitutive plant C comes from $CO_2$ fixation by photosynthesis. It is not so for N, which is incorporated by different processes, including N fixation by bacteria and fertilization by useful forms of N, which may come as particles or dissolved in water. We did not even try to separate these components. Because of the biological features of our plant, summarized in the introduction, we assumed all its accumulated materials as derived from air sources.

As explained in the introduction, the C and N isotopes help in identifying some air, water and soil pollution sources. The interest of the air pollution community on this subject, at least part of this community, is clearly indicated by the increasing number of reports on this issue in environmental journals. Sorry, we cannot say to which extent is this more or less relevant to the global biogeochemical cycle.

4. What quality control was done to assure that the observed bioaccumulation was not due to natural differences in the plants, which correlated with spatial location, or natural ecosystem differences. In terms of the soil, it appears that some of the soil accumulation is associated with local dust that could be independent of air pollution. This issue needs to be better discussed as the implication that all differences were due to local pollution seems inappropriate.

Natural plants, unfortunately, are not standardized materials for air monitoring. This imposes taking some provisions to assure sampling fairness among sites. We did not reach the fine point of having plants with the same measured ability to capture air pollutants at each site. As described in the sampling section, we took a couple of minimal provisions to cover this aspect: 1) We used the same plant species at all sites, which leaves apart major biological differences between species that may affect their

ability to trap pollutants, and 2) We used only leaves of similar age (time of exposition) for final measurements. In addition, this plant grows on the crown surface of mesquite trees; i.e., completely exposed to the surrounding air, and the natural ecosystem (vegetation) is fairly similar and opened throughout the study area.

There is, apparently, a conceptual implication in the second part of your comment. It seems that you do not consider as air pollutants those materials entering into the air from soils sources, whereas we do. You may or may not agree, but the concept of natural sources of air pollution is quite well established. Nothing to teach you, but soil is one of them. Not to mention the heavily polluted agriculture soils at our study region, which would classify somewhere in between natural and anthropogenic sources.

According to our interpretation of the factor analysis results, the regional crustal/soil sources are the most important factors behind the observed distribution of most metals. About 75% of the "explained" variance was associated to them. The rest appeared associated to industrial and agriculture activities. This does not preclude potential inputs from sources located outside our study area, but we were unable to capture that aspect. This may be possibly part of that 26% "unexplained" variance, but this would need further research.

5) The use of factor analysis to help understand sources of the metals need to re-evaluated to assure that the interpretation of the results is appropriate. The use physical mechanisms of deposition (both wet and dry) are very complex processes that are dependent on particle size. To this end, the factors may not necessarily reflect sources and may reflect factors impacting deposition with or without source differentiation.

We do agree with these comments, particularly because of the small size of our data set. Since our sample size was heavily determined by time and funds available, we would like to see whether new, very different results would appear after increasing the number of sampling sites. The spatial distribution of most pollutants was pretty well defined (we only showed a few of them). We expect no major pattern changes by adding new sampling sites. However, some of them (e.g., Cd, Pb) may be better defined if we added sites.

Could you be more specific about the need of reevaluating our factor analysis? As you may guess, we have tried a number of preliminary FAs: for metals or PAHs alone, lowering or increasing the number of variables, different rotations, and so on. Although we got more explained variance in some solutions, we decided to present the one in our manuscript after considering several aspects of our data, as explained in the statistical section.

**Specific comments**

1) Page 5815

-What QA/QC was done to demonstrate that the extraction and ICP-OES analysis method was accurate for V, Ba, Pb, Sr, Ni, Ca, Cr, Sb, Mo, and Cd. The text indicates that QA/QC was checked for other light elements that optical ICP is well suited but some of these other elements are more difficult by ICP.

From your listed metals, we refer in the methods digestion (extraction) recovery values for Ba and Ca. For the rest of elements in your list, we have not a really technical explanation. Upon our analyst's request, we had to make a decision to reduce the list of elements ICP analyzed for recovery purposes. So, we decided to go only for a few elements with certified values for our reference material for metals (pine needles, cited in the text). Since V, Pb, Sr, Ni, Cr, Sb and Mo lack certified values, they were left apart and assumed to be fully extracted by our procedure.

-In addition, what was done for blank subtraction to address field and laboratory contamination? Were any checks performed? What was done to validate that the extraction quantitatively extracted these metals?

A blank solution with the same digestion acids was processed equally to the samples and the concentrations of the respective positive elements discounted from the sample concentration. Only checks for repeatability were done, as explained in the methods. Extractions were evaluated by percent recovery as described in the methods and the comment above.

2) Page 5817 – What QA/QC was done for PAH analysis. Were any checks performed? What was done to validate that the extraction quantitatively extracted these PAH.

- In this case, we fully validated our HPLC measuring procedure before analyzing the samples, and recorded it internally at our institution (IMP-153004-06-01, documentation available in Spanish). Validation parameters included repeatability, linearity, accuracy, reproducibility, detection and quantification limits (partially described in our methods). Each sample batch included: system and extraction blanks, a spiked sample and duplicates, and a batch specific calibration curve.

-Checks: batch calibration curves were continuously verified with control charts.

-Because no reference material for PAH in *Tillandsia* was available, for extraction validation we used the added standard system: Cromosorb resin was used as matrix surrogate; it was added with different concentrations of known standard PAH mixture and processed as the real samples (extracted, cleaned, concentrated and analyzed). Recovery was 98% on average.

3) Page 5825 – The connection of high Ca and the limestone areas is not clear. Are the authors suggestion that the Ca is from the mines or that the soil in these areas are just different and high in Ca. If the authors are suggesting that the high Ca is from the mines, how is this determined?

The areas where high Ca concentrations were observed are characterized by sedimentary rocks soils (limestone); thus, a number of quarries (open sky mines) and cement plants (shown by a letter C on Figure 6c) are located here. As a result of this, samples taken at these areas were highly influenced by both mine activities and dust resuspension.

4) Page 5827 – The interpretation of factor 3 is confusing. Are these elements lumped together because they are from the same source or because the deposition process that lead to their accumulation on or in the plant are similar? Do the authors believe that these are all from cement facilities, multiple sources in the near proximity of the cement facilities, and atmospheric parameters that lead to higher uptake of the pollutants from an atmospheric mix that is from a wide range of sources.

Elements in F3 are lumped together because they had a high degree of spatial correlation; i.e., roughly, their concentration in the biomonitor increased or decreased at the same sites. Interestingly, Ca, typically a soil/crustal element, correlated better with elements from fossil fuel combustion (F3) than with those from crustal/soil origin. Actually, it also appeared in the crustal/soil factor (F1), but with lower loading (0.49). The high correlation of Ca with V, Ni, chrysene and pyrene at Mezquital Valley is clearly an anthropogenic effect: increased Ca emissions from limestone quarries and cement plants, which use heavy amounts fossil fuels. So far, we cannot answer your last questions in your comment; all your options may be partially true. Factor 3 still includes contributions from the refinery, the electricity plant and the cement plants among other smaller local sources. Thus, we need to do something else to be able of discriminating among those specific sources by biomonitoring techniques.

5) The use of isotope ratios is confusing in the factor analysis. Is this a reflection of location that is represented by isotopic ratios?

We do not think so. In this case, isotopes were just other chemical variables more or less correlated with the rest. Not surprisingly, the $\delta^{13}C$ appeared with negative loading at factor 3, meaning that this ratio decreases when fossil fuel combustion emissions increase. This is discussed in the text (see also figures 4 and 9). The $\delta^{15}N$ was quite different due to inputs from other sources, apparently as (or more) important than fossil fuel combustion, basically the local agriculture with wastewater. In this case of N we need further work to be more specific about what industrial and agriculture N compounds are determining the observed regional pattern.