

Dear Carlos H. Vergara,

According to our records, you are the corresponding author of the following studies:

- Carlos\_H\_Vergara\_Coffea\_arabica\_Mexico\_2004 (Study ID): 16 site/s. Raw data were collected from Rader et al., PNAS 2016; 113, 1.

Please confirm that (i) you are the correct corresponding author of the listed studies, and (ii) we have your permission to re-use your data, and (iii) list all co-author/co-owners affiliations and acknowledgements/funding information in your 'Data\_ownership' (excel) file.

The above datasets have been processed in line with OBServ guidelines (see 'General information about OBServ data processing' at the end of this file). In order to check the correctness of the results obtained from the processing stage, we would like to share them with you, along with a report and some queries. To do so, we have created one folder for each study. There you can find the following items:

- One folder that contains the raw data that was sent to us.
- One 'field\_level\_data' (csv) file with the processed data.
- One 'insect sampling' (csv) file, if sampling information was provided to us.
- One 'Explanation column names' (xlsx) file with auxiliary information to interpret the data in 'field\_level\_data' and 'insect\_sampling' files.
- One 'Summary report and queries' (pdf) file generated from your corresponding 'field\_level\_data' file.
- One folder with background information on data file conversion. It contains at least one R script, whose name is the study ID. This file was used to process the raw data and generate the previous csv files. Thus, if needed, you can go through all the decisions that were taken to process your file. In case you want to run the script notice that paths and working directories should be updated to your relative local paths. Besides, if your R script needs additional input files or post-processing (such as 'Table\_organism\_guild\_META.csv', 'taxon\_table.csv', or 'add\_taxon\_constraint\_column\_DAINESE\_RADER\_OTHER.R'), such files have been also included in the 'background information on data file conversion' folder.

Please review at least the 'field\_level\_data' and the 'Summary\_report\_and\_queries' files, answer the queries in the latter and send your answers/comments back to us, along with your 'Data\_ownership' (excel) file.

If you have any questions, please do not hesitate to contact us.

Best regards,

Alfonso Allen-Perkins and  
the OBServ Team

=====  
General information about OBServ data processing  
=====

If several censuses were conducted in a given field/orchard, we recorded the total sampling effort that was made, that is, the total sampling time [in minutes] and the total sampled area [in m<sup>2</sup>].

- The values of richness, abundance and visitation rates for a given site are obtained by aggregating the records of insects that were observed during the total sampling time. Consequently, in this database richness, abundance and visitation rates do not reflect the mean value of the respective surveys or rounds in each site, but the total one.
- Pan-traps data were not taken into account to estimate the values of richness and abundance, respectively, in each site.

- When possible, visitation rates were only derived from timed observations to a given number of flowers, and their units were set to [visits per 100 flowers and hour].
- Richness data were not recorded if the percentage of identified species (or morphospecies) was lower than or (approximately) equal to 75%.