



# Mid-infrared spectroscopy of earthworm bodies to investigate their species belonging and their relationship with the soil they inhabit

P. Trao, C. Rumpel, Y. Capowiez, Pascal Jouquet, C. Pelosi, Anne Pando,  
Hanane Aroui Boukbida, Nicolas Bottinelli

## ► To cite this version:

P. Trao, C. Rumpel, Y. Capowiez, Pascal Jouquet, C. Pelosi, et al.. Mid-infrared spectroscopy of earthworm bodies to investigate their species belonging and their relationship with the soil they inhabit. *Applied Soil Ecology*, 2021, 162, 103894 [6 p.]. 10.1016/j.apsoil.2021.103894 . hal-03159318

**HAL Id: hal-03159318**

**<https://hal.science/hal-03159318>**

Submitted on 13 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

**Mid-infrared spectroscopy of earthworm bodies to investigate their species belonging and their relationship with the soil they inhabit**

Thao Pham<sup>a</sup>, Cornelia Rumpel<sup>a</sup>, Yvan Capowiez<sup>b</sup>, Pascal Jouquet<sup>a</sup>, Céline Pelosi<sup>b</sup>, Anne Pando<sup>a</sup>, Hanane Aroui Boukbida<sup>a</sup>, Nicolas Bottinelli<sup>a, c \*</sup>

<sup>a</sup> IRD, CNRS, Sorbonne Université, UPEC, INRAE, Institut d'écologie et des sciences de l'environnement, IESS-Paris, FEST team, F-93143 Bondy, France

<sup>b</sup> INRAE, UMR 'EMMAH', Site Agroparc, 84914 Avignon cedex 09, France

<sup>c</sup> Soils and Fertilizers Research Institute (SFRI), Duc Thang, Bac Tu Liem, Hanoi, Vietnam

Corresponding author: Nicolas Bottinelli

E-mail: [nicolas.bottinelli@ird.fr](mailto:nicolas.bottinelli@ird.fr)

## Abstract

Interactions of earthworm species with their environment and in particular the soil they inhabit is poorly understood. In this study, we aimed to use mid infrared absorption (MIR) spectroscopy of earthworm bodies in combination with discriminant analysis of principal components (DAPC), in order to investigate if their biogeochemical composition is determined by their species belonging or the soil they inhabit. Our conceptual approach included the comparison of MIR spectra of four earthworm species (*Lumbricus terrestris*, *Allolobophora chlorotica*, *Aporrectodea caliginosa* and *Aporrectodea icterica*) sampled from four different agricultural fields in France. Moreover, we tested the influence of the habitat on the earthworm MIR spectra by exchanging individuals and soils in an incubation experiment.

MIR spectra varied according to species and sampling site. The MIR spectra of earthworm bodies also changed during the incubation experiment, and for *L. terrestris*, *A. chlorotica* and *A. icterica* depended on the properties of the soil ingested. This suggested that the food source in different habitats impacted the biogeochemical composition of earthworm bodies. DAPC allowed for the identification of earthworm species ranging from 64 to 100% of accuracy (mean 78%). We thus conclude that MIR spectroscopy may be a suitable approach to identify earthworm species, but that their signature may also allow to investigate their ecology as determined by the soil they inhabit.

**Key words:** feeding behavior, ecological category, lumbricid species

## 1. Introduction

Earthworms play an important role in the functioning of the soil system due to their bioturbation activity and the production of biostructures with specific physical, chemical and biological properties (Bottinelli et al., 2015; Van Groenigen et al., 2019). Their impact on soil functioning is variable and often species-specific even within a single ecological category (Capowiez et al., 2015; Vos et al., 2019). Thus, the identification of soil-earthworm interactions and in particular the earthworms' feeding behavior at the species level remains a crucial step for a detailed understanding of their impacts on soil functions. However, the identification of earthworms using the conventional morpho-anatomical and even modern barcoding methods do not allow for any biogeochemical insights in the interaction of earthworms with their soil habitat.

A vast body of the literature compiled over the past two decades has established that infrared spectroscopy combined with calibration models can offer a time-efficient and cost-effective method of quantifying various soil properties (Viscarra Rossel et al., 2006), e.g., organic carbon, nitrogen and soil texture. More recently, infrared spectroscopy has proved to be a very useful and efficient tool for the identification of soil organisms (Bellon-Maurel and McBratney, 2011), e.g., insects (Johnson, 2020; Jouquet et al., 2014) and microorganisms (Wenning and Scherer, 2013), or for estimating nematode abundance in soil (Barthès et al., 2011). The latter works provided evidence that the differences of the spectral composition among individuals may be related to the chemical composition of their bodies as determined by their species belonging. Among the methods, mid-infrared (MIR) spectroscopy might be the most suitable for the quantification of soil organisms, since many absorption bands of organic molecules are within this spectral range. Conversely, the visible and near-infrared regions, are limited to considerably weaker overtones (Bellon-Maurel and McBratney, 2011).

The aim of the present study was to use MIR spectra of earthworm tissues for the identification of earthworm species sampled in different environments and to investigate if the spectral information is related to their habitat. If a relationship with the habitat was evidenced, this consequently would indicate a link between food sources and the biogeochemical composition of earthworm bodies. To this end, we compared the spectral signature of the body mass of four common Lumbricid species sampled in contrasting soil habitats at four different sites in France. We investigated if the spectral signature of earthworm bodies differed locally (two adjacent sites) and/or regionally under contrasting land use. In addition, we investigated the relation with their soil habitat by exchanging earthworms and soils during a laboratory incubation. Since earthworms not only live in but also feed on the surrounding soil, we hypothesized that MIR as a phenotype method may also be useful to elucidate their feeding behavior in contrasting environments. Moreover, we hypothesized that species belonging of earthworms living in different soil types may be identified using MIR spectra and that the influence of species belonging would be stronger than that of soil habitat.

## **2. Materials and methods**

### *2.1. Earthworm sampling*

A total of 161 adults or sub-adults of four Lumbricid species (37 *Lumbricus terrestris*, 43 *Allobophora chlorotica*, 40 *Aporrectodea caliginosa* and 41 *Aporrectodea icterica*) were collected in four study fields in France: two grasslands at Versailles (V1 (48°48'29.6"N 2°05'26.4"E), V2 (48°48'19.5"N 2°05'05.9"E)) in the region 'Ile de France', and two orchards at Avignon (A1 (43°54'54.0"N 4°53'02.4"E), A2 (43°53'09.6"N 4°47'56.4"E)) in the region 'Provence-Alpes-Côte-d'Azur' i.e. at about 600 km distance. At Avignon, soils under the

orchards are Alluviosols of the Durance valley and they have been managed without pesticide input for 10 years. Soils at Versailles are Luvisols of silt-loamy texture and have been under herbal vegetation since 20 years with regular mowing. All four soils showed contrasting physicochemical properties, even locally (Table 1).

It is noteworthy that *A. chlorotica* sampled at Versailles were the green morph, whereas those sampled at Avignon were the leucotypic morph (Bouché, 1972). Individuals were washed with distilled water and kept for four days in the dark at 12°C on moistened filter paper in Petri dishes to completely empty their guts. Filter papers were cleaned and re-moistened every 12 h. After gut clearance, the earthworms were frozen for 24 h, and the posterior part of their body was removed (last 5 to 10 segments) because it often contains some remaining soil that could bias the analysis. The earthworms were then dried in an oven at 60°C for two days and manually ground into powder.

## 2.2 Incubation experiment

To further investigate the relationship between earthworms and their habitat we carried out an incubation experiment with eight individuals of each species from one sampling location, a soil under fallow sampled at Versailles (V2). Individuals were incubated in this soil or in an orchard soil sampled at Avignon (A2). Both soils were collected from the 0-25 cm depth, air-dried, crushed and sieved to 2 mm and put in microcosms consisting of PVC boxes (500 cm<sup>3</sup>). The soil in each of the boxes was watered with distilled water to reach 60% of the water holding capacity and manually compacted to reach a bulk density of 1.15 g cm<sup>-3</sup>. One earthworm individual was added to each microcosm. In total, 64 microcosms were set up: 4 earthworm species × 2 soils × 8 replicates. The microcosms were kept at 15°C in the dark for 60 days, but the soil was renewed

after 30 days. At the end of the incubation, the earthworms were collected, and samples were prepared as explained in section 2.1 (i.e., washed, kept on moistened filter paper for 4 days, frozen for 24 h, dried and manually ground).

### *2.3. MIR spectra acquisition and pre-processing*

Ground earthworm samples were scanned in diffuse reflectance mode with a spectrometer (FTIR 660, Agilent ex-Varian) in the medium infrared spectral range from 4000 to 400  $\text{cm}^{-1}$  with a KBr beamsplitter and a silicon detector. Only spectra from 3700 to 424  $\text{cm}^{-1}$  were selected for data analysis in order to disregard the noise at the two edges of the spectral range. The diffuse reflectance measurements were made at 2  $\text{cm}^{-1}$  intervals and converted to absorbance. The raw spectra were subjected to pre-processing, including Savitzky-Golay smoothing (fourth-order polynomial filter with a 33-point window) and normalization by the standard normal variate. We then removed bands associated with  $\text{CO}_2$  interference (2500-2300  $\text{cm}^{-1}$ ). Data pre-processing was performed with the packages “spectacles” and “signal” in R software.

### *2.4. Statistical analysis*

We analysed our data by principal component analysis in order to discriminate earthworm species from contrasting sites. As this approach did not allow for any differentiation, we used discriminant analysis of principal components (DAPC) to assess any changes in the MIR spectra between sampling sites and after laboratory incubation. DAPC is a multivariate method, which relies on data transformation using principal component analysis as a prior step to discriminant analysis, which ensures that variables submitted to discriminant analysis are uncorrelated, and that their number is less than that of analyzed individuals (Jombart et al., 2010). Unlike principal

components analysis, which summarizes the overall variability among soil spectra (both between-group and within-group variability), DAPC summarizes the spectral differentiation between groups, while overlooking within-group variation. A 10-fold cross validation repeated 100 times was used to decide how many principal components to retain. DAPC was also used as a model to test the capacity of MIR data to identify earthworm species. To this end, the whole dataset representing 161 individuals was split using species and locality (A1, A2, V1, V2 and laboratory) as stratification into two parts, of which 115 individuals (75%) were selected for model construction and the remaining 46 individuals (25%) were used to assess the prediction accuracy. The overall accuracy (number of correct predictions/total number of validation data) of the model and the accuracy for each species were computed with the confusion matrix of the validation dataset. All statistical calculations and plots were carried out using R software (version 3.5.2) using “ade4”, “splitstackshape”, “ade4genet” and “ggplot2” packages.

### **3. Results**

#### *3.1 MIR spectra of earthworm species sampled in contrasting pedoclimatic environments*

The average MIR spectra of earthworm bodies from each site showed clear separation between the regions Provence-Alpes-Côte-d’Azur (Avignon 1 and 2) and Ile de France (Versailles 1 and 2) in the regions 3000-2500 and  $< 1000\text{ cm}^{-1}$  (Fig. 1a). As a result, the discriminant analyses of principal components (DAPC) separated earthworms according to the two different regions Ile de France and Provence-Alpes-Côte-d’Azur’ along the first discriminant function (LD1), which explained 96% of the variability. In contrast, the second discriminant function (LD2) revealed only slight local differences of earthworms’ spectral signatures for orchard soils sampled at Avignon.



159

### 160 3.2. MIR spectra of earthworms sampled after the incubation experiment

161 The DAPC of the earthworms' MIR spectra before and after incubation in the laboratory (Fig. 2)  
162 showed an important impact of the incubation conditions on the MIR spectra, particularly along  
163 the first discriminant function (LD1), which explained more than 70% of the variability for the  
164 four species. Indeed, the MIR signature changed during the incubation experiment regardless the  
165 soil type used (Avignon or Versailles). The MIR signatures after incubation were similar for *A.*  
166 *caliginosa*, while they were significantly different for *L. terrestris*, *A. chlorotica* and *A. icterica*  
167 when incubated in the two contrasting soils.

168

### 169 3.3. Automatic identification of earthworm species

170 The average MIR spectra for each species had similar peaks with slight difference of  
171 absorbances mainly in the regions 3500-3000 and  $< 1800\text{ cm}^{-1}$  (Fig. 3a). The DAPC based on the  
172 training dataset separated all the species (Fig. 3b). The first discriminant function (LD1)  
173 explaining 53% of the variability, opposed *L. terrestris* to *A. caliginosa* and *A. icterica*, with *A.*  
174 *chlorotica* having intermediate values. The first discriminant function (LD1) was mainly linked  
175 to the bands 3514, 1745, 1147, 1028, 980 and  $717\text{ cm}^{-1}$  (Fig. 3c). The second discriminate  
176 function (LD2) (explaining 28% of the variability) opposed *A. icterica* to *A. chlorotica*, with *A.*  
177 *caliginosa* having intermediate values. The second discriminant function (LD2) was mainly  
178 linked to the bands 3271, 1680, 1169 and  $1041\text{ cm}^{-1}$ . The overall accuracy (78%) obtained in the  
179 confusion matrix from comparison of the predicted and observed samples indicated a moderate  
180 classification accuracy (Fig. 4). The accuracy was lowest for *A. icterica* (67%) and *A. caliginosa*  
181 (64%) and highest for *A. chlorotica* (83%) and *L. terrestris* (100%).

## 4. Discussion

### 4.1. Importance of species belonging and pedoclimatic environment for MIR spectral signature of earthworms from regionally and locally differentiated sampling sites

MIR spectra were recorded for earthworms sampled from four different soils in two French regions (Provence-Alpes-Côte-d'Azur and Ile de France) and thus give information about local and regional influences on the composition of their body tissues. At the regional scale, MIR of earthworms could clearly be differentiated, while at the local scale (Avignon and Versailles), the differences between the two local sampling sites were much less important. These differences among regional and local signatures of earthworm bodies cannot be caused by soil particles remaining in the earthworms' gut because they were emptied during the 4 days of starving, and only the anterior part of the body was used for the analyses. The changes in the MIR spectra may therefore be attributed to the composition of their habitat since soil properties and vegetation were different between fields and orchards. Similar results were found for MIR spectra of termite colonies, which showed a link between the physiological states of termite populations, and their habitat (environment and their mound properties) (Jouquet et al., 2018). The findings are also in line with observations made for the soil microbial biomass, which changes its composition and properties depending on the food source in their environment (Kramer and Gleixner, 2008; Rumpel et al., 2001). It has been shown that the MIR signals as biochemical fingerprints of bacterial strains have a high potential to elucidate pathogen host interaction (Novais et al., 2019). We therefore hypothesized that this may be also the case for earthworm-soil habitat interactions and investigated if the biogeochemical signature of the earthworm bodies may be influenced by the soil they inhabit.

#### 4.2 Impact of soil habitat on earthworm body composition

In order to test our hypothesis, we carried out a laboratory incubation experiment, where earthworms were incubated in the soil they usually inhabit (earthworms from V2 in soil V2) as well as a second soil sampled in another region under contrasting land use (earthworms V2 in soil A2). During the 60-day incubation, the MIR spectra of all species changed, with some differences among species. This rapid biogeochemical change due to the incubation conditions could probably be explained by the stress caused by the experiment (e.g., manipulation, size and depth of the boxes, and absence of food supply, natural gradients for soil moisture and daily variation in temperature). Behavioral changes in earthworms may be related to their response to a change in the soil environment, e.g., soil moisture, temperature (Kretzschmar, 1991; Uvarov et al., 2011) and/or food supply (Cook and Linden, 1996). This behavioral change thus influenced the biogeochemical signature of earthworm bodies. The MIR signatures of *L. terrestris*, *A. chlorotica* and *A. icterica* species were additionally influenced by the soil they inhabited. In analogy with the changing MIR spectral composition of microorganisms, which depend on their host environment, we suggest, that the food source in form of soil organic matter and minerals may be different in the two soils with contrasting physico-chemical conditions. MIR spectroscopy may therefore be a promising tool to complement genetic analysis by a phenotyping approach allowing to assess biogeochemical interactions between earthworms and their habitat.

#### 4.3. Identification of earthworms via MIR spectroscopy

DAPC allowed for the automatic identification of earthworm species with an overall accuracy of 78%. Although this result seems elevated at first glance, it remains lower than the accuracy obtained from mitochondrial DNA barcoding and other methods of DNA analysis (Decaëns et al., 2013). This may be explained with the phenotypic nature of the MIR signature, which is influenced by the earthworms' habitat in addition to species belonging. Interestingly, in the DAPC plan, the four species were somewhat located according to their ecological category classification (Bouché, 1972). Indeed, the epi-anecic *L. terrestris* was separated from the two endogeic species (*A. caliginosa* and *A. icterica*) along the first discriminant function (LD1), with the intermediate category *A. chlorotica* located between them. This result confirms that the feeding behavior of earthworms, in terms of amount and quality of ingested soil and organic matter may affect the MIR signature of earthworms. MIR might thus be a suitable method to investigate their ecology.

We conclude that MIR spectral investigations are complementary to molecular methods, as they allow to take into account effects of the soil habitat on earthworm biogeochemical composition in addition to the identification of species belonging.

## Acknowledgements

This project was financially supported by the Federation Ile-de-France for Research on the Environment (FIRE) and the University of Science and Technology of Hanoi (USTH) through the Mater 2 internship grant awarded to Thao Pham. We would like to thank the people who helped in the field, Solène Orrière and Véronique Etievant.

## References

250 Barthès, B.G., Brunet, D., Rabary, B., Ba, O., Villenave, C., 2011. Near infrared reflectance  
 251 spectroscopy (NIRS) could be used for characterization of soil nematode community. *Soil*  
 252 *Biol. Biochem.* 43, 1649–1659. <https://doi.org/10.1016/j.soilbio.2011.03.023>  
 253 Bellon-Maurel, V., McBratney, A., 2011. Near-infrared (NIR) and mid-infrared (MIR)  
 254 spectroscopic techniques for assessing the amount of carbon stock in soils - Critical review  
 255 and research perspectives. *Soil Biol. Biochem.* 43, 1398–1410.  
 256 <https://doi.org/https://doi.org/10.1016/j.soilbio.2011.02.019>  
 257 Bottinelli, N., Jouquet, P., Capowiez, Y., Podwojewski, P., Grimaldi, M., Peng, X., 2015. Why is  
 258 the influence of soil macrofauna on soil structure only considered by soil ecologists? *Soil*  
 259 *Tillage Res.* 146, 118–124. <https://doi.org/10.1016/j.still.2014.01.007>  
 260 Bouché, M.B., 1972. *Lombriciens de France: écologie et systématique*. Institut national de la  
 261 recherche agronomique Paris.  
 262 Capowiez, Y., Bottinelli, N., Sammartino, S., Michel, E., Jouquet, P., 2015. Morphological and  
 263 functional characterisation of the burrow systems of six earthworm species (Lumbricidae).  
 264 *Biol. Fertil. Soils* 51, 869–877. <https://doi.org/10.1007/s00374-015-1036-x>  
 265 Cook, S.M.F., Linden, D.R., 1996. Effect of food type and placement on earthworm  
 266 (*Aporrectodea tuberculata*) burrowing and soil turnover. *Biol. Fertil. Soils* 21, 201–206.  
 267 <https://doi.org/10.1007/BF00335936>  
 268 Decaëns, T., Porco, D., Rougerie, R., Brown, G.G., James, S.W., 2013. Potential of DNA  
 269 barcoding for earthworm research in taxonomy and ecology. *Appl. Soil Ecol.* 65, 35–42.  
 270 <https://doi.org/https://doi.org/10.1016/j.apsoil.2013.01.001>  
 271 Johnson, J., 2020. Near-infrared spectroscopy (NIRS) for taxonomic entomology: A brief  
 272 review. *J. Appl. Entomol.* 144, 241–250. <https://doi.org/https://doi.org/10.1111/jen.12732>

273 Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a  
 274 new method for the analysis of genetically structured populations. *BMC Genet.* 11, 94.

275 Jouquet, P., Capowiez, Y., Bottinelli, N., Traoré, S., 2014. Potential of near infrared reflectance  
 276 spectroscopy (NIRS) for identifying termite species. *Eur. J. Soil Biol.* 60, 49–52.  
 277 <https://doi.org/https://doi.org/10.1016/j.ejsobi.2013.11.004>

278 Jouquet, P., Pando, A., Aroui, H., Harit, A., Capowiez, Y., Bottinelli, N., 2018. Evidence from  
 279 Mid-Infrared Spectroscopy (MIRS) that the biochemical fingerprints of *Odontotermes*  
 280 *obesus* colonies change according to their geographical location and age. *Insectes Soc.* 65,  
 281 77–84. <https://doi.org/10.1007/s00040-017-0589-0>

282 Kramer, C., Gleixner, G., 2008. Soil organic matter in soil depth profiles: distinct carbon  
 283 preferences of microbial groups during carbon transformation. *Soil Biol. Biochem.* 40, 425–  
 284 433. <https://doi.org/https://doi.org/10.1016/j.soilbio.2007.09.016>

285 Kretzschmar, A., 1991. Burrowing ability of the earthworm *Aporrectodea longa* limited by soil  
 286 compaction and water potential. *Biol. Fertil. Soils* 11, 48–51.  
 287 <https://doi.org/https://doi.org/10.1007/BF00335834>

288 Novais, Â., Freitas, A.R., Rodrigues, C., Peixe, L., 2019. Fourier transform infrared  
 289 spectroscopy: unlocking fundamentals and prospects for bacterial strain typing. *Eur. J. Clin.*  
 290 *Microbiol. Infect. Dis.* 38, 427–448. [https://doi.org/https://doi.org/10.1007/s10096-018-](https://doi.org/https://doi.org/10.1007/s10096-018-3431-3)  
 291 3431-3

292 Rumpel, C., Grootes, P.M., Kögel-Knabner, I., 2001. Characterisation of the microbial biomass  
 293 in lignite-containing mine soils by radiocarbon measurements. *Soil Biol. Biochem.* 33,  
 294 2019–2021. [https://doi.org/https://doi.org/10.1016/S0038-0717\(01\)00122-5](https://doi.org/https://doi.org/10.1016/S0038-0717(01)00122-5)

295 Uvarov, A. V., Tiunov, A. V., Scheu, S., 2011. Effects of seasonal and diurnal temperature

fluctuations on population dynamics of two epigeic earthworm species in forest soil. *Soil Biol. Biochem.* 43, 559–570. <https://doi.org/10.1016/j.soilbio.2010.11.023>

Van Groenigen, J.W., Van Groenigen, K.J., Koopmans, G.F., Stokkermans, L., Vos, H.M.J., Lubbers, I.M., 2019. How fertile are earthworm casts? A meta-analysis. *Geoderma* 338, 525–535. <https://doi.org/10.1016/j.geoderma.2018.11.001>

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* 131, 59–75. <https://doi.org/10.1016/j.geoderma.2005.03.007>

Vos, H.M.J., Koopmans, G.F., Beezemer, L., Goede, R.G.M. De, Hiemstra, T., Willem, J., Groenigen, V., 2019. Large variations in readily-available phosphorus in casts of eight earthworm species are linked to cast properties. *Soil Biol. Biochem.* 138, 107583. <https://doi.org/10.1016/j.soilbio.2019.107583>

Wenning, M., Scherer, S., 2013. Identification of microorganisms by FTIR spectroscopy: Perspectives and limitations of the method. *Appl. Microbiol. Biotechnol.* 97, 7111–7120. <https://doi.org/10.1007/s00253-013-5087-3>

319

## 320 **Figure captions**

321 **Fig. 1:** (a) Average absorption of pre-processed MIR spectra for earthworms sampled in the  
322 regions Provence-Alpes-Côte-d'Azur (A1, n=22 and A2, n=28) and Ile de France (V1, n= 21 and  
323 V2, n=29). The grey shadow represents the standard error; (b) scatter plot of the discriminant  
324 analysis of principal components (DAPC) performed on the MIR spectra using sites in the region  
325 Provence-Alpes-Côte-d'Azur at Avignon (A1 and A2) and in the region Ile de France at Versailles  
326 (V1 and V2) as priori groups. Ellipses correspond to 67% confidence intervals around the  
327 barycenter. The number of principal components retained was 30.

328 **Fig. 2:** Scatter plot of the discriminant analysis of principal components (DAPC) performed on  
329 the MIR spectra of *L. terrestris*, *A. chlorotica*, *A. caliginosa* and *A. icterica* using the three a  
330 priori groups represented by before incubation (V2 soil) and after incubation in the V2 and A2  
331 soils. Ellipses correspond to 67% confidence intervals around the barycenter. The number of  
332 principal components retained ranged between 5 to 8.

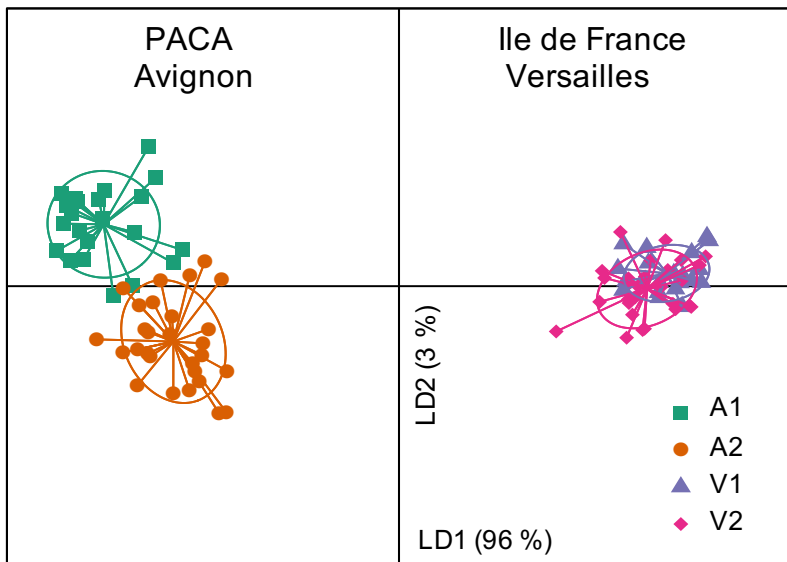
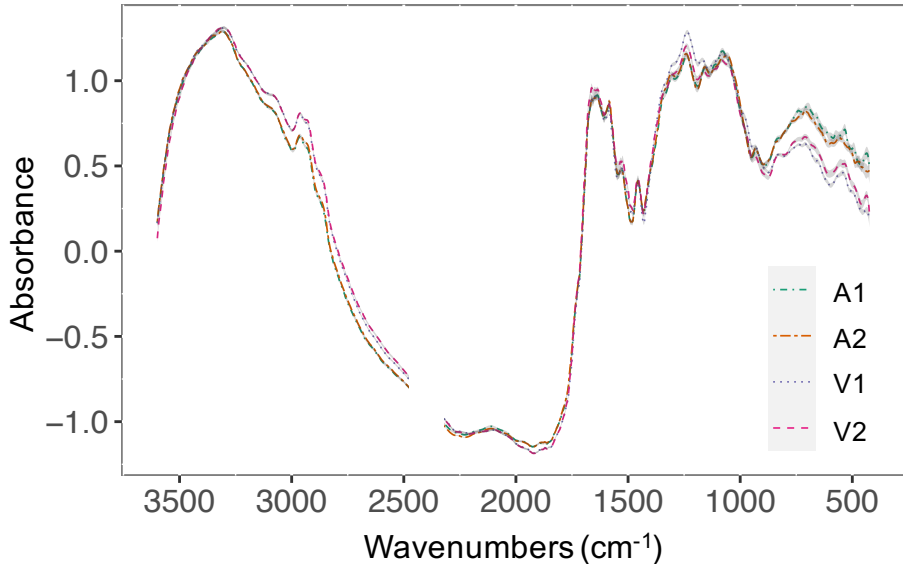
333 **Fig. 3:** (a) Average absorption of pre-processed MIR spectra of the training dataset for *L.*  
334 *terrestris* (n=26), *A. chlorotica* (n=31), *A. caliginosa* (n=29), and *A. icterica* (n=29). The grey  
335 shadow represents the standard error; (b) scatter plot of the discriminant analysis of principal  
336 components (DAPC) performed on the MIR spectra of the training dataset using the four  
337 earthworm species as a priori groups. Ellipses correspond to 67% confidence intervals around the  
338 barycenter. The number of principal components retained was 60; (c) Loadings of the  
339 wavenumbers contributing to the discriminant function 1 and 2. Loadings above the dashed line  
340 indicate 20% of the most important wavenumbers.

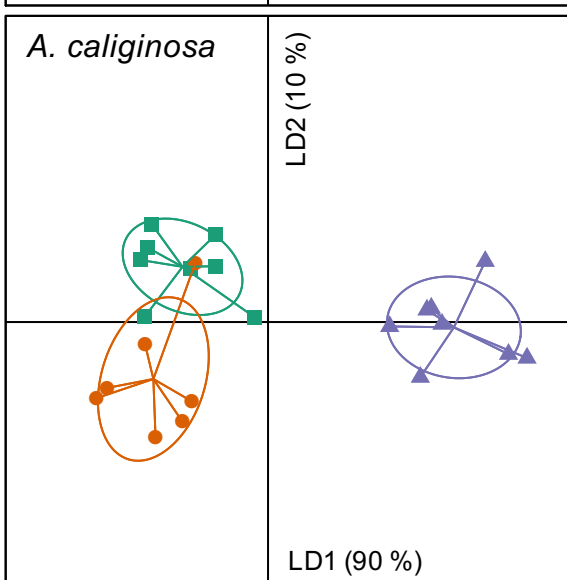
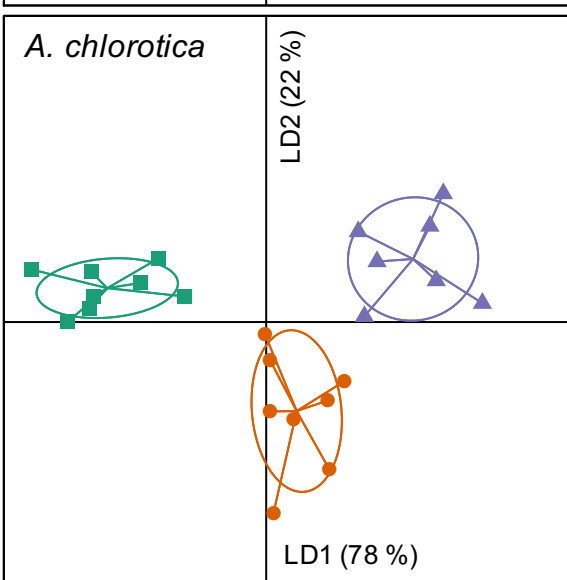
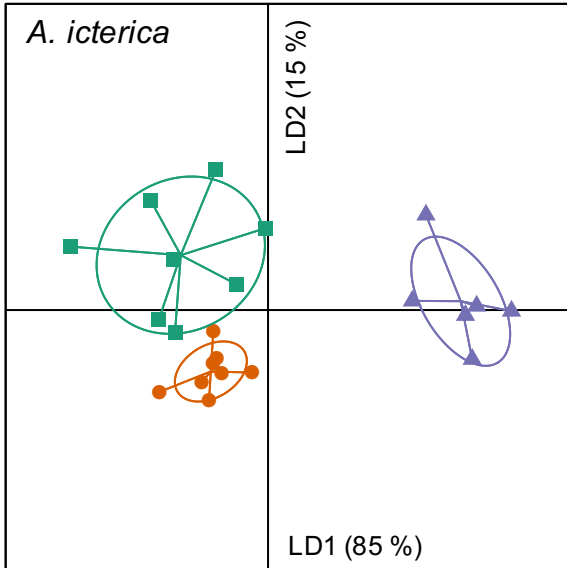
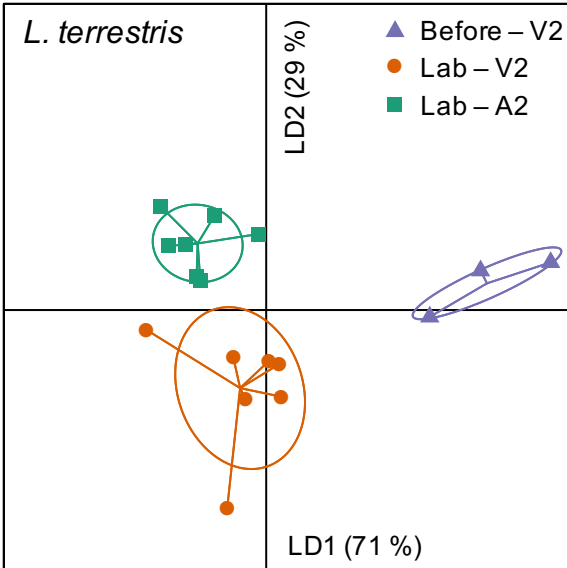
341

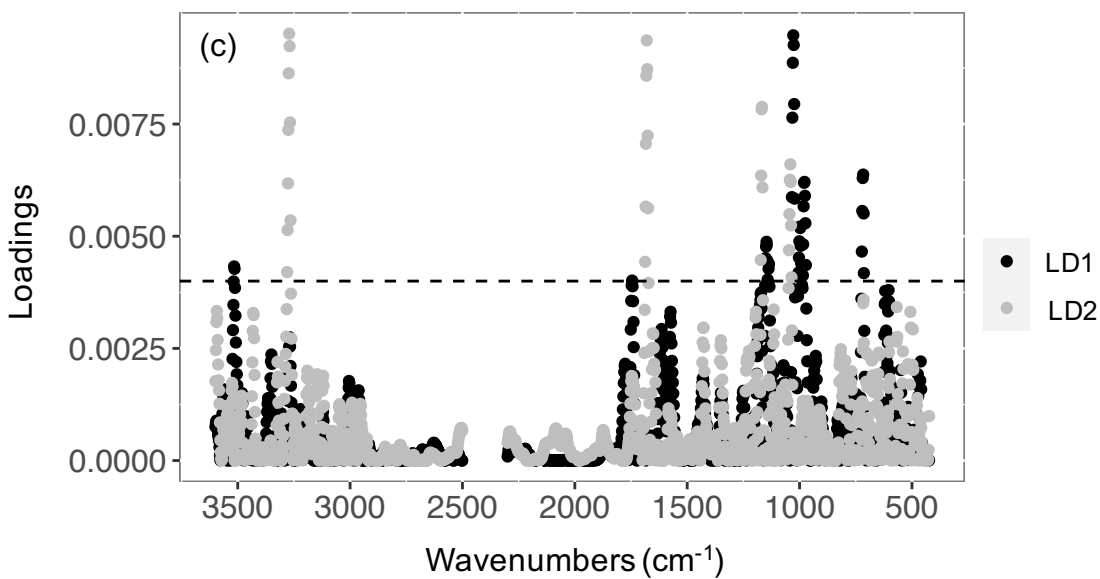
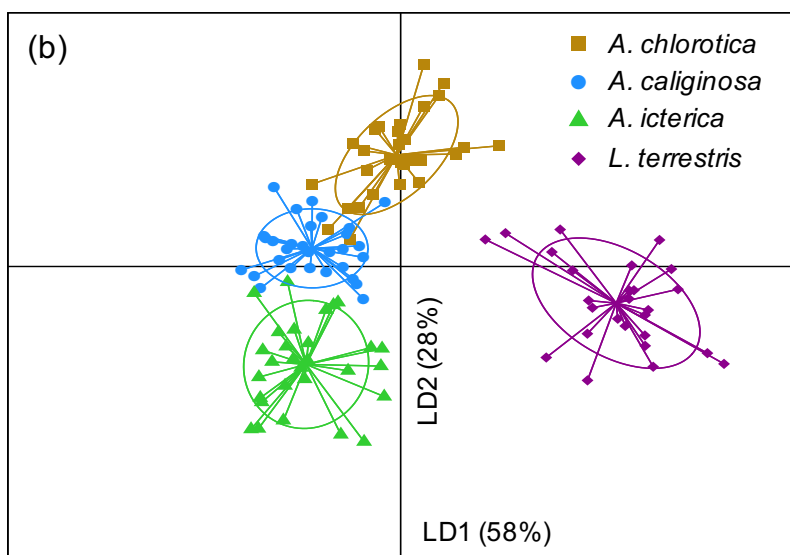
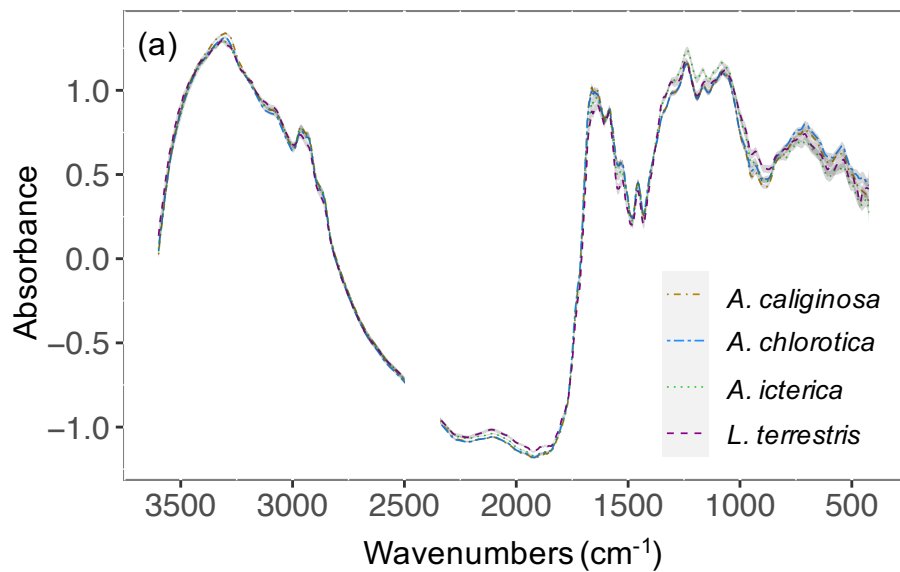


**Table 1:** Soil properties (0-15 cm) for the region Provence-Alpes-Côte-d’Azur at Avignon (A1 and A2) and the region Ile de France at Versailles (V1 and V2) where *L. terrestris*, *A. chlorotica*, *A. caliginosa* and *A. icterica* were collected. OM indicates organic matter.

Fields	OM	Clay	Silt	Sand	pH (H <sub>2</sub> O)	Land use
	%					
V1	3.3	23	48	29	7.5	Fallow
V2	1.9	17	61	22	6.7	Fallow
A1	2.1	16	35	47	7.7	Apple orchard
A2	5.1	31	49	21	8.3	Apple orchard



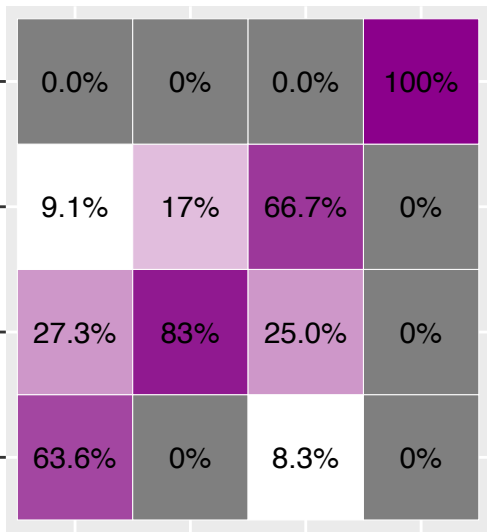




Overall accuracy 78%

Prediction

*L. terrestris*  
*A. icterica*  
*A. chlorotica*  
*A. caliginosa*



*A. caliginosa*  
*A. chlorotica*  
*A. icterica*  
*L. terrestris*

Reference