MetaPlantCode Spring Survey 2025: Community Insights into Plant-related eDNA Metabarcoding

Ann-Sophie Mattern, Stephanie J. Swenson, Birgit Gemeinholzer

COPYRIGHT NOTICE

Introduction

DNA Metabarcoding is an emerging and innovative technique for taxonomic identification of complex environmental samples. Plant DNA metabarcoding is of interest for many different ecological and biomonitoring research questions, however it presents a number of unique challenges that may deter researchers from using the technique. The aim of the MetaPlantCode project [www.metaplantcode.eu], part of the 2022-2023 BiodivMon[1] joint call, funded by Biodiversa+, the European Biodiversity Partnership and co-funded by the European Commission (GA No. 101052342), is to provide tools that enable researchers to be confident in using this approach by providing best practice protocols for sample collection and laboratory protocols, as well as innovating data pipelines to ensure better accuracy in species level identification. To better understand the needs, challenges, and practices of researchers currently using, or interested in the methods, a short anonymous survey was conducted to evaluate stakeholder preconceptions in order to guide the deliverable products from the MetaPlantCode project. An integral part of this effort is ensuring that all output adhere to the FAIR Data principles[2], making data findable, accessible, interoperable and reusable. This supports transparency, reproducibility, long-term data usability and fosters collaboration across the scientific community and with data infrastructures such as GBIF (Global Biodiversity Information Facility), BOLD (Barcode of Life Data System), and NCBI (National Center for Biotechnology Information), PR2 (Protist Ribosomal Reference Database), SILVA (Comprehensive Ribosomal RNA Database), and UNITE (Unified system for the molecular identification of fungi), as well as broader initiatives like

BiCIKL (Biodiversity Community Integrated Knowledge Library), iBOL-Europe (International Barcode of Life – Europe), and BGE (Biodiversity Genomics Europe), which are relevant for data integration and sharing across platforms.

Scope

The online survey was conducted from March 25 to July 7, 2025 via Google Forms[3]. It targeted members of the scientific community, especially individuals already working with or planning to work with plant-related eDNA metabarcoding. The survey was available in English only and distributed through various channels: the internal MetaPlantCode mailing list, the project's Instagram account, and direct outreach to related research projects (Biodiversity Community Integrated Knowledge Library (BiCIKL), iBOL-Europe, Biodiversity Genomics Europe (BGE), and BOLD). Of these, only BiCIKL responded, offering to share the survey through their networks within the journals *Metabarcoding and Metagenomics* (MBMG) and the *PhytoKeys*. Within MetaPlantCode, one reminder was sent via the project mailing list, no additional follow-ups were issued. It consisted of 21 questions (listed in the Supplementary Information) divided into three sections: general information about the participants, a section for respondents with experience in metabarcoding, and a section for those with little or no prior exposure to the method. In total, 21 participants could be acquired. The results were intended to guide the development of standardized, accessible protocols and resources, such as those planned by the MetaPlantCode project.

Survey Findings

Participant Profile and Familiarity

Regarding academic background, most participants were at an early career stage. A majority held a PhD (38,1 % - Fig. 1) or were currently pursuing one (33,3 % - Fig. 1). Other respondents included master's graduates, professors and government employees (Fig. 1). This distribution reflects the strong representation of early-career researchers in the field and highlights the importance of targeted support for this group. More than a half of the respondents were affiliated with universities (52,4 %), and about 29 % worked at research institutes. Familiarity with metabarcoding was generally high: Over 50 % reported being familiar with the approach, 9,5 % planned to use it, and 33 % expressed interest.

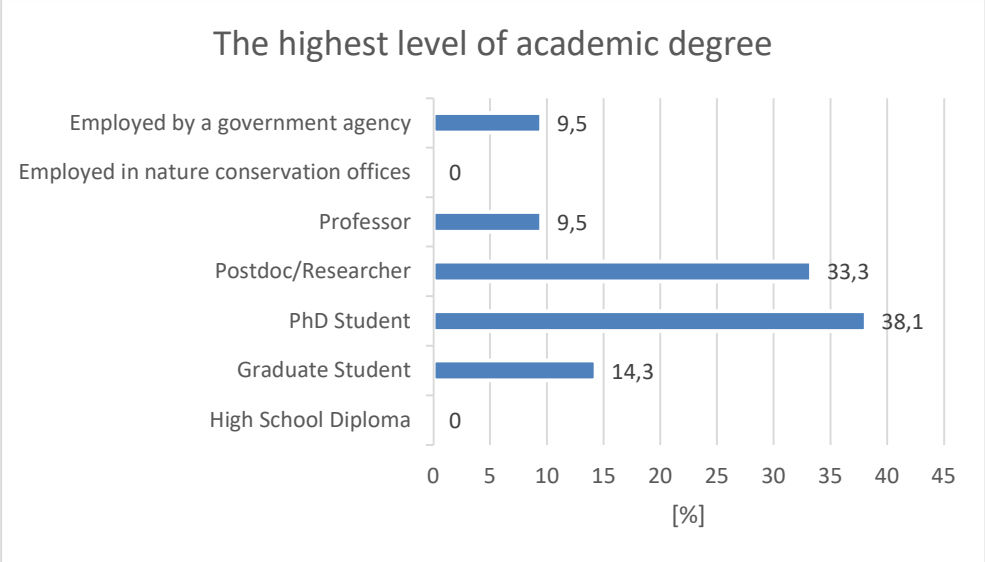

Fig. 1: Highest academic degree or current academic status of survey participants (n = 21). Percentages reflect the proportion of respondents in each category.

It should be noted that the survey design allowed all participants to proceed through the entire questionnaire regardless of their level of familiarity with plant metabarcoding. As a result, individuals without practical experience were able to respond to questions intended primarily for users of metabarcoding methods. While this has little impact on most results, particularly regarding interest in eDNA sample types or markers, where perspectives from outside users may still be informative, it does reduce the interpretive accuracy of responses concerning the use of software tools and standard protocols. In future surveys a clearer question structure should be implemented.

Sample Types and Barcoding Markers

Our respondents reported the highest interest in soil samples (61,1 % - Fig. 2), followed by insect body pollen (44,4 % - Fig. 2), airborne pollen and food or herbal authentication (each 33,3 % - Fig. 2). Other sample types included fecal samples and pollen from nest provisioning hymenoptera, indicating diverse research applications (Fig. 2).

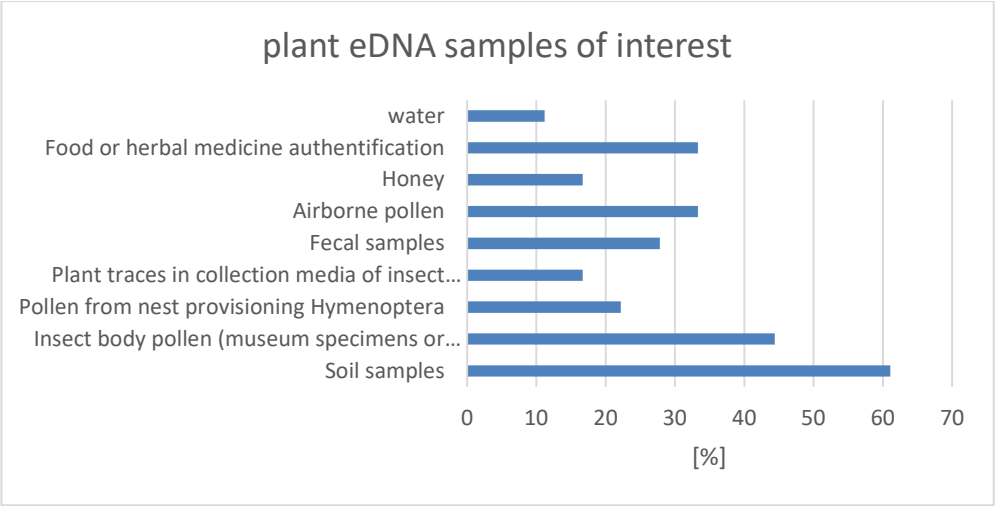

Fig. 2: Plant eDNA samples of interest among survey participants (n=21). Multiple answers were possible. The percentages indicate how many respondents selected each sample type.

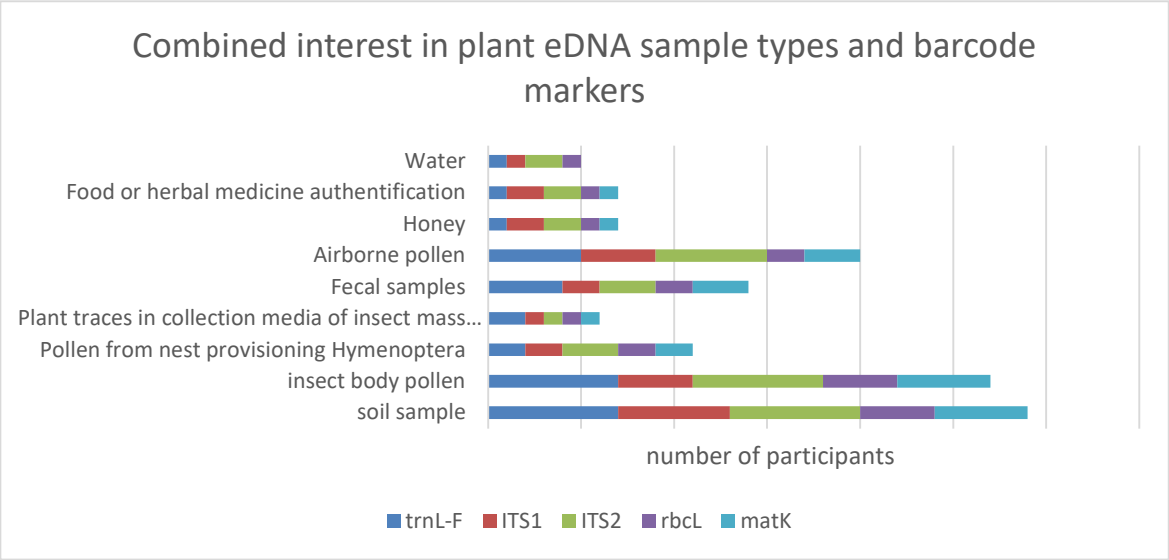

Fig. 3: Combined interest in plant eDNA sample types and barcode markers. This chart visualizes the total number of responses (n = 12) in which participants indicated interest in both specific sample types and corresponding barcode markers (trnL-F, ITS1, ITS2, rbcL, matK).

Preferred barcoding markers were ITS2 (52,9 %), trnL-F (47,1 %) and ITS1 & ITS2 in combination (47,1 %), with many using multiple markers. Figure 3 shows the combined results of preferred sample types and markers to explore possible correlations. A total of 12 responses could be matched, meaning participants who answered both relevant questions. This matching was necessary to identify any relationship. The chart indicates

that soil samples, insect body pollen and airborne pollen are the most preferred sample types, as we already could observe in Figure 2. For these, the markers trnL-F and ITS2 are most frequently requested. Interest exists for all sample types across all markers, though to varying degrees.

Protocols, Software, and Data Practices

Most participants were aware of and applied FAIR data principles[2], although about 28 % were not. Half of the respondents reported using standardized protocols for sampling, but noted that these are mostly 'internal' procedures, as no official standards exist , though some steps appear to be consistent across studies or sample types. The remaining respondents cited unclear standards or lack of involvement in sampling as reasons for not using standardized protocols. For data processing, 41 % used standardized protocols, while 35 % did not but showed interest in partial adoption. Commonly used open-source software tools included Cutadapt[4], VSEARCH[5], and OBITools4[6]. However, nearly 30 % of respondents (7 out of 17) reported not using any software tools. Among these, two participants indicated familiarity with plant metabarcoding but likely rely on in-house pipelines. The remaining five participants reported no involvement in metabarcoding, which, as already described, may have influenced the overall distribution of responses. Quality control was performed by 75 % of respondents, and widely used reference databases included BOLD[7], NCBI[8], PR2[9], Silva[10], and UNITE[11]. Data submission practices varied: about one third had submitted both raw and processed data, another third planned to do so, while a minority did not plan to submit. Nearly all participants supported the development of standardized best practices, and 87 % stated they would adopt such standards.

Challenges and Needs

The main challenges identified were lack of method standardization, issues with reference databases, amplification primer choice, budget constraints, subsampling bias, and insufficient open-source software. Participants not currently working with plant eDNA cited limited access to resources and knowledge as key barriers, along with the complexity and costs involved.

Requested resources to facilitate broader adoption included simple protocols, training

workshops, user-friendly tools, support networks, and cost reduction strategies.

These findings highlight the importance of clear, accessible guidance, resource availability, and community-agreed standards to foster growth in plant metabarcoding research.

Alignment with MetaPlantCode and Outlook

The results align well with the goals of the MetaPlantCode project, which aims to provide marker-independent, reproducible protocols, workflows, and documentation to enhance accessibility and standardization. While the project addresses many community needs, further emphasis on communication, training, and low-threshold resources could improve uptake, especially among less experienced users.

In conclusion, continued community engagement and feedback will be essential to develop tools and standards that meet diverse requirements and promote reproducibility and collaboration across the plant eDNA metabarcoding field.

References

[1] *2022 – 2023 BioDivMon*. (2022, 7. Oktober). Biodiversa +. https://www.biodiversa.eu/2022/10/07/2022-2023-joint-call/

[2] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., . . . Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3(1). https://doi.org/10.1038/sdata.2016.18

[3] Google. (o.D.). Google Forms [Software]. Available at: docs.google.com/forms/ [17.07.2025].

[4] Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

[5] Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

[6] OBITools4 Documentation. (o. D.). Retrieved July 21, 2025, from https://obitools4.metabarcoding.org/docs/about/#what-are-obitools4

[7] *BOLD – The Barcode of Life Data Systems*. (o. D.). Retrieved July 21, 2025, from https://boldsystems.org/

[8] *National Center for Biotechnology Information*. (o. D.). Retrieved July 21, 2025, from

https://www.ncbi.nlm.nih.gov/

[9] Vaulot, D. (o. D.). *The PR2 databases*. Retrieved July 21, 2025, from https://pr2-database.org/

[10] SILVA. (o. D.). *Silva.* Retrieved July 21, 2025, from https://www.arb-silva.de/

[11] University of Tartu, Natural History Museum (2017). *UNITE.* Retrieved July 21, 2025, from https://unite.ut.ee/

Supplementary

Catalog of questions

General Information

1. What is your highest level of academic degree?
    - High School Diploma
    - Graduate Student
    - PhD Student
    - Postdoc/Researcher
    - Professor
    - Employed in nature conservation offices
    - Employed by a government agency
    - Other…
2. Where are you currently employed or continuing education?
    - University
    - Research Institute
    - Government Agency
    - Non-governmental Organization (NGO)
    - Private Industry
    - Other…
3. Are you already familiar with plant metabarcoding, or are you interested in conducting a plant metabarcoding study in the future?
    - Yes, I am already familiar.
    - No, but I plan to conduct such a study in the future.
    - No, but I am interested.
    - Other…
4. Which types of plant eDNA samples are of interest to you?
    - Soil Samples
    - Insect body pollen (museum specimens or new collections)
    - Pollen from nest provisioning Hymenoptera
    - Plant traces in collection media of insect mass collection traps
    - Fecal samples
    - Airborne Pollen

- Honey
- Food or Herbal medicine authentication
- Other…

5. Are you familiar with the FAIR principles (Findable, Accessible, Interoperable, Reusable)?
   - Yes, I am familiar with the FAIR principles and apply them in my work.
   - Yes, I am familiar with the FAIR principles, but I do not currently apply them in my work.
   - No, I am not familiar with the FAIR principles.
   - Other…

Plant eDNA Metabarcoding Practices and Challenges

6. If you are familiar with the FAIR principles but do not apply them or you are opposed to FAIR data repositories, please briefly explain why.

   (Free Text)

7. Do you use standardized protocols for sampling eDNA in the field? (Please specify your answer in the "Other" field.)
   - Yes (If "Yes", please specify which protocols you use.)
   - No (If "No", would you be interested in using standardized protocols?)
   - Lack of clear protocols and standards make me reluctant to implement plant metabarcoding in my research program
   - Other…

8. Do you use standardized protocols for processing eDNA samples in the laboratory? (Please specify your answer in the "Other" field.)
   - Yes (If "Yes", please specify which protocols you use.)
   - No (If "No", would you be interested in using standardized protocols?)
   - Other…

9. Which plant barcoding region are you using or are you most interested in?
   - trnL-F
   - ITS1
   - ITS2
   - ITS1 & ITS2
   - rbcL
   - matK
   - Other…

10. Do you use any open-source bioinformatics software for your analysis? (Please specify your answer in the "Other" field.)
    - Yes (If "Yes", please provide a link, name etc.)
    - No (If "No", do you use any in-house custom made software?)
    - Other…
11. Are you encountering challenges with bioinformatics software?
    - Yes/No
12. Which reference databases or taxonomic assignment systems are you currently using or would you like to use? (Please specify.)

    (Free Text)

13. Do you manually perform quality control of your metabarcoding results?
    - Yes/No
14. Do you submit your sampled and analyzed metabarcoding data to a FAIR repository?
    - Yes, I submit both sampled and analyzed data.
    - Yes, but I only submit sampled data (e.g., sequencing results, metadata about samples).
    - Yes, but I only submit analyzed data (e.g., taxonomic assignments, statistical results).
    - No, I do not submit any data, and I am not planning to. (If applicable, please answer the next question.)
    - No, but I plan to in the future.
    - Other…
15. Are you in favor of developing standardized best practices for plant metabarcoding? (Please specify.)

    (Free Text)

16. Would you use established standards for plant metabarcoding if they were available? (Please specify.)

    (Free Text)

17. What are the biggest obstacles you face in plant eDNA metabarcoding? (Please specify.)

    (Free Text)

Exploring Plant eDNA Metabarcoding

18. What interests you most about plant eDNA metabarcoding?
    - Biodiversity monitoring
    - Environmental conservation

- Innovation in genetic research
- Developing new research methods
- Other…

19. What are the main reasons you have not yet used plant metabarcoding?

- Lack of technical knowledge or training
- Limited access to equipment or resources
- Uncertainty about protocols and best practices
- No clear use case for my research
- Other…

20. What type of resources would help you get started with plant eDNA metabarcoding?

- Easy-to-follow protocols and guidelines
- Hands-on training workshops
- Access to open-source analysis tools
- Support networks or expert consultations
- Other…

21. What would make you more likely to start using plant metabarcoding in your research?

    (Free Text)