



HAL
open science

Finding the perfect pairs: A matchmaking of plant markers and primers for multi-marker eDNA metabarcoding

Armando Espinosa Prieto, Laurent Hardion, Nicolas Debortoli, Jean-Nicolas Beisel

► **To cite this version:**

Armando Espinosa Prieto, Laurent Hardion, Nicolas Debortoli, Jean-Nicolas Beisel. Finding the perfect pairs: A matchmaking of plant markers and primers for multi-marker eDNA metabarcoding. *Molecular Ecology Resources*, 2024, 24 (4), pp.e13937. 10.1111/1755-0998.13937 . hal-04568179

HAL Id: hal-04568179

<https://hal.science/hal-04568179>

Submitted on 7 May 2024

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.

RESOURCE ARTICLE

Finding the perfect pairs: A matchmaking of plant markers and primers for multi-marker eDNA metabarcoding

Armando Espinosa Prieto¹ | Laurent Hardion¹  | Nicolas Debortoli^{2,3} | Jean-Nicolas Beisel^{1,4}¹University of Strasbourg, CNRS, Laboratoire Image Ville Environnement, UMR 7362, Strasbourg, France²Namur Molecular Tech, CHU UCL Namur, Yvoir, Belgium³E-BIOM SA, Namur, Belgium⁴École Nationale du Génie de l'Eau et de l'Environnement de Strasbourg (ENGEEES), Strasbourg, France**Correspondence**

Armando Espinosa Prieto, University of Strasbourg, CNRS, Laboratoire Image Ville Environnement, UMR 7362, Strasbourg, France.

Email: armando.espinosa@live-cnrs.unistra.fr**Funding information**

French State as part of the Projet de territoire de Fessenheim, OHM Fessenheim - University of Strasbourg; Agence de l'Eau Rhin-Meuse

Handling Editor: Eric Coissac**Abstract**

As the scope of plant eDNA metabarcoding diversifies, so do the primers, markers and methods. A wealth of primers exists today, but their comparative evaluation is lacking behind. Similarly, multi-marker approaches are recommended but debates persist regarding barcode complementarity and optimal combinations. After a literature compilation of used primers, we compared *in silico* 102 primer pairs based on amplicon size, coverage and specificity, followed by an experimental evaluation of 15 primer pairs on a mock community sample covering 268 plant species and genera, and about 100 families. The analysis was done for the four most common plant metabarcoding markers, *rbcL*, *trnL*, ITS1 and ITS2 and their complementarity was assessed based on retrieved species. By focusing on existing primers, we identify common designs, promote alternatives and enhance prior-supported primers for immediate applications. The ITS2 was the best-performing marker for flowering vascular plants and was congruent to ITS1. However, the combined taxonomic breadth of ITS2 and *rbcL* surpassed any other combination, highlighting their high complementarity across Streptophyta. Overall, our study underscores the significance of comprehensive primer and barcode evaluations tailored to metabarcoding applications.

KEYWORDS*in silico*, internal transcribed spacer, mock community, primer pairs, *rbcL*, *trnL***1 | INTRODUCTION**

Plant eDNA metabarcoding, i.e. genetic material obtained directly from environmental samples (soil, sediment, water, etc.), is emerging as a robust and versatile approach, finding utility in diverse applications including herbivore diet analyses (da Silva et al., 2019; De Barba et al., 2014; Valentini, Miquel, et al., 2009), investigations in plant–pollinator interactions (Baensch et al., 2020; Bell et al., 2019; Richardson et al., 2019), assessment of complex food item composition (Bruno et al., 2019; Frigerio et al., 2020), characterization

of herbal products (Anthoons et al., 2021; Arulandhu et al., 2019; Chen et al., 2010; Xin et al., 2018) and vegetation surveys (Alsos et al., 2018; Espinosa Prieto et al., 2023). These eDNA metabarcoding studies consistently demonstrate that taxonomic resolution and the number of detected taxa are influenced by the choice of primers and markers. While finding a unique and universal plant barcoding marker akin to COI for animals remains challenging, the literature recognizes several commonly used plant markers (Taberlet et al., 2018), encompassing those associated with plastid DNA (*rbcL*, *matK*, *trnL* and *trnH-psbA*) (CBOL Plant Working Group

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Molecular Ecology Resources* published by John Wiley & Sons Ltd.

et al., 2009; Hollingsworth et al., 2011; Kress et al., 2015; Kress & Erickson, 2007) and nuclear ribosomal DNA (nrDNA) internal transcribed spacers (ITS) (Chen et al., 2010; Cheng et al., 2016; China Plant BOL Group et al., 2011). Adopting a multi-marker approach is recommended to increase the taxonomic resolution in plant eDNA research (da Silva et al., 2019; Fazekas et al., 2008; Kolter et al., 2021). However, the superior running cost associated with additional PCR amplifications, higher sequencing depths and extended laboratory and bioinformatics worktimes, makes this approach less cost-effective. Some experimental setups may offer promising solutions to reduce costs, such as multiplex PCRs (Bell et al., 2017; De Barba et al., 2014), high-level NGS multiplexing (da Silva et al., 2019; De Barba et al., 2014) and the use of degenerate primer sets (Elbrecht et al., 2019; Elbrecht & Leese, 2017b; Tournayre et al., 2020).

As the field of eDNA-based biomonitoring is relatively new and rapidly evolving, the methodological selection of primers and markers still requires validation and customization. This step is influential in eDNA studies as it directly impacts the taxonomic resolution (Kelly et al., 2019; Krehenwinkel et al., 2017; Li et al., 2020; Piñol et al., 2019). Original designs are being enhanced (Cheng et al., 2016; Moorhouse-Gann et al., 2018) and primers for specific floras are proposed (Coghlan et al., 2020; Drummond et al., 2021), but they lack direct comparative evaluations with a consistent metrics. To achieve optimal selection and combination of eDNA primers and markers, one should consider the following criteria: (i) the amplicon length should account for the expected level of DNA degradation in the studied environmental sample to increase the probability of detection and the insert size for NGS sequencing; (ii) the primers should be universal for a given taxonomic group, meaning they should ensure amplification for every species of this group and only them (Taberlet et al., 2018, Chapter 2); (iii) the markers should be able to distinguish between closely related taxa (specificity criteria); (iv) because one marker is not sufficient in plant eDNA metabarcoding, the choice a several markers should optimize their complementarity in order to enhance the accuracy and reliability of the identified taxa.

The trade-off between the length of eDNA fragments, its prevalence in the environment and the specificity of the barcode strongly influence amplification success and the detection of taxa (Freeland, 2017). Early on, the recommended average size of metabarcoding markers for eDNA was around 150bp (Valentini, Pompanon, & Taberlet, 2009), but this depends on the nature of the environmental sample (environmental matrices); for example, plant metabarcoding on soil samples could use whole barcodes of over 500bp (Fahner et al., 2016; but see Yoccoz et al., 2012). In water samples, the higher occurrence of small DNA fragments contributes to an elevated read abundance of short amplicons (Bylemans, Furlan, et al., 2018). In specific applications dealing with highly degraded eDNA, such as ancient DNA or diet metabarcoding, shorter barcodes are preferred to enhance PCR success rates (Bylemans, Gleeson, et al., 2018; Epp et al., 2012; Taberlet et al., 2007).

Universal primers have the potential to transcend applications due to their capacity to amplify a range of plant taxa. A good example

is primer pair *g-h* amplifying ~50bp of the *trnL*-UAA intron (Taberlet et al., 2007), which is used today in ancient DNA studies (Garcés-Pastor et al., 2022; Jorgensen et al., 2012; Sønstebø et al., 2010), for diet analysis (Baamrane et al., 2012; da Silva et al., 2019; De Barba et al., 2014; Valentini, Miquel, et al., 2009), for plant biodiversity assessment in aquatic ecosystems (Alsos et al., 2018; Cannon et al., 2016; Revéret et al., 2023; Tsukamoto et al., 2021) and in soil samples (Ariza et al., 2023; Fahner et al., 2016; Gaggini et al., 2019; Nichols et al., 2018), to cite a few. For a community approach, universal primers should target highly conserved binding sites common to all target taxa with minimal to no mismatches (Valentini, Pompanon, et al., 2009). A common feature of universal primers is the incorporation of degenerate nucleotides, as they accommodate for superior taxonomic coverage and help mitigate, to some extent, taxa amplification bias in complex environmental samples (Elbrecht & Leese, 2017b; Kolter & Gemeinholzer, 2021; Krehenwinkel et al., 2017; Moorhouse-Gann et al., 2018).

The capacity of a marker to resolve closely related species depends on its evolutionary history. Phylogenetic and barcoding studies have identified several universal plant barcodes (CBOL Plant Working Group et al., 2009), but eDNA applications use shorter sections of these barcodes at the expense of taxonomic resolution. Among the traditional plastid markers – *matK*, *rbcL* and *trnL* – metabarcoding generally achieves enhanced taxonomic resolution through the utilization of longer barcodes, given the limited polymorphism of these markers. In contrast, the higher mutation rate of the nrDNA ITS2 and its conserved small size (~220bp) for vascular plants makes it highly interesting for metabarcoding applications.

The combination of several markers is often used to increase taxonomic coverage, resolution and eventually quantitative estimates (Richardson et al., 2019). Different markers can exhibit varying degrees of specificity across different taxonomic groups, thus ensuring better coverage of the plant diversity within a study area (Cannon et al., 2016; De Barba et al., 2014). This is especially beneficial in ecosystems or samples characterized by a high level of species diversity (Vasconcelos et al., 2021), as well as when confronting plant groups marked by intricate taxonomic relationships (Kuzmina et al., 2018). However, determining the best marker combination requires careful consideration of marker specificity, PCR efficiency, computational demands and cost implications.

Designing and validating primers for metabarcoding necessitate expertise in both computational and wet laboratory techniques (Freeland, 2017). Primers are initially evaluated via *in silico* assessments, followed by *in vitro* analysis on mock communities or eDNA samples. Recognizing that primers were designed using different procedures, we emphasize the necessity of a comprehensive assessment of existing primers within a comparative study, using the aforementioned criteria: (i) to establish optimal primer pairs based on size, coverage and specificity, and (ii) to ascertain the most efficient combination of barcodes for multi-marker metabarcoding. Initially, we conducted a comprehensive search on the Web of Science (WOS) for published primer sets employed in eDNA-based plant metabarcoding, with a focus extending beyond

ecological research. Subsequently, we evaluated their performance using *in silico* PCR against all standard plant sequences from the EMBL database and whole plastid genomes from NCBI. We then conducted an *in vitro* metabarcoding assessment using a mock community sample encompassing 268 plant species from the four predominant plant groups: mosses, ferns, angiosperms and gymnosperms. This allowed for the evaluation of the retained primers' performance as identified in the *in silico* evaluation. The exhaustive and reproducible evaluation method combining *in silico* and *in vitro* approaches for four markers and with such a large mock community represents an unprecedented work. This pioneering methodology was used to offer primer selection recommendations and to suggest optimal barcode combinations for eDNA plant biomonitoring surveys.

2 | METHODS

2.1 | Identification of common metabarcoding primers

We first searched the online database WOS Core Collection for studies on plant DNA metabarcoding from all types of environmental samples (search date 26 June 2021). The search was carried out independently for each marker, *rbcl*, *trnL*, *matK* and ITS, with the terms 'plant' AND 'metabarcoding'. The query for the ITS contained the terms ('internal transcribed spacer' OR 'ITS1' OR 'ITS2') NOT 'fung*'. For the *rbcl*, we added 'NOT diatoms'. Following a thorough examination of titles and abstracts, we meticulously handpicked studies that exclusively met the following criteria: (i) analysis of environmental samples (including faeces, water, soil and sediment) or bulk samples like honey and plant-derived products, (ii) targeting several taxa and (iii) using high-throughput sequencing (HTS) rather than PCR-based detection. Studies on macro- and microalgae were excluded from the list because primers differ from those used for Embryophyta. We also discarded primers that were developed for specific taxonomic groups, such as Rosaceae (De Barba et al., 2014), marine vascular plants (Ortega et al., 2020) and freshwater vascular plants (Coghlan et al., 2020). We listed all identified primer pairs and incorporated combinations that had not been previously used as a pair in the literature.

2.2 | *In silico* testing

We first retrieved all available 'standard' plant sequences from the EMBL database (a total of 3,102,589,278 sequences, sourced from EMBL release 143 from March 2020), which we formatted in the ITS database following the exact procedure in (Bellemain et al. 2010) using the *ecoPCR* from the package *OBITools* (Boyer et al., 2016) with primers ITS-p5 and ITS-u4 (Cheng et al., 2016). The procedure extracted the sequences containing the entire ITS region along with

the primer binding sites for all evaluated primers and the flanking 50 nucleotides when available, with a taxonomic focus limited to Streptophyta (5670 species, 1229 genera, 227 families; Table S1, Appendix S1). In parallel, we acquired the complete plastid genomes from NCBI (>11,000 sequences, release March 2023 from <ftp.ncbi.nlm.nih.gov/refseq/release/plastid/>) and selected one genome per species within Streptophyta taxa. This curated collection of sequences was employed for the *in silico* evaluation of *rbcl* and *trnL* primers (10,339 species, 3023 genera, 424 families; Table S1, Appendix S1). Both strategies of database curation were used to standardize the available sequences, specifically to ensure equal representation of complete primer binding sites for all tested primer pairs in the reference database. This balanced approach eliminates any potential bias in amplification success due to variations in primer site availability in public reference databases.

We performed *in silico* PCRs (ePCR) using *ecoPCR* functions integrated within *OBITools* (Boyer et al., 2016). By assessing both published and primer combinations that were never used together before, our objective was to identify the most effective pairs. We first evaluated amplicon length by conducting ePCRs against the two custom databases described earlier, applying the following parameters: (i) a maximum allowance of three mismatches between each primer and its binding site and (ii) a minimum and maximum amplicon length unique for each marker (Appendix S1). The results from all the ePCRs were compared and we retained primer combinations yielding ~300-bp long amplicons. This mean amplicon length was selected as it would fit most applications based on the size of DNA in environmental samples and full-length barcodes could be sequenced using standard sequencing technologies (e.g. Illumina paired end 2×300). We evaluated barcode coverage (Bc) and specificity (Bs) as defined in Ficetola et al. (2010). Considering a specific taxonomic rank (e.g. genus or species), the Bc is the ratio between the number of amplified taxa and the total number of taxa in the reference sequence database, and the Bs is the ratio of unambiguously identified taxa over the total number of tested taxa. These measures were calculated using *ecotaxspecifity* and *ecotaxstat* commands, respectively, from *OBITools*. The Bc index was calculated for Streptophyta (NCBI TaxID 35,493), and Bs was calculated using the default parameters, wherein a species was unambiguously identified if an amplified sequence displayed a minimum nucleotide difference of one nucleotide from all other amplified sequences, regardless of the length. Subsequently, the melting temperature (T_m) and annealing temperature (T_a) for the selected primer pairs were calculated with the NEB Tm Calculator v1.15.0 (New England Biolabs, MA, USA) since the Q5 High-Fidelity DNA Polymerase (New England Biolabs, MA, USA) was used for the *in vitro* evaluation. The T_m of pairs was considered together with the level of degeneracy. Each new combination of primers was retained if the difference between their T_m was below the recommended threshold of 5°C, or otherwise if the T_a value was under 5°C from the T_m of either primer. For those with ambiguous nucleotides, the T_a was compared against the lowest T_m value of the variants. Combinations documented in the literature, even with a T_m difference surpassing the threshold, were still retained.

2.3 | In vitro evaluation: Mock community

2.3.1 | Sampling, DNA isolation and mock preparation

The primer pairs retained from the in silico evaluation were used to amplify a composite DNA community sample (mock community) of 268 plant species common to the upper Rhin River basin. Because closely related plant species could share similar barcodes, we deliberately chose only one species per genus, that is, 268 genera based on the French taxonomic repository TaxRef version 15 (Gargominy et al., 2021). In total, this mock community includes 98 families, comprising 241 angiosperms, six gymnosperms, 13 pteridophytes and 8 bryophytes. Briefly, DNA was extracted independently for every species following an in-lab CTAB protocol from 20 to 50 mg of dried leaves from herbarium voucher specimens prepared between 2020 and 2022 (STR Herbarium, University of Strasbourg). After quantification of DNA samples with the Qubit dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA), the mock community sample was obtained by mixing equal amounts of DNA per species (~5 ng). To eliminate PCR inhibitors from CTAB-based DNA extractions, we diluted this mock community (2 ng/μL) and cleaned 50 μL with the DNeasy Plant Mini Kit (Qiagen) for a final concentration of 0.79 ng/μL.

2.3.2 | DNA amplification

PCRs for the selected primers were carried out in 25 μL reaction volumes following recommendations for the Q5 High-Fidelity Polymerase by NEB with 5 μL 5X Q5 reaction buffer (1X), 1.25 μL of each primer (0.5 μM), 1 μL 5 mM dNTPs (200 μM), 0.25 μL Q5 High-Fidelity Polymerase (0.02 U/μL) and 2.5 μL template DNA. PCR products were visually checked on a 2% agarose gel under UV light. Primers had Illumina TruSeq adaptors used by Eurofins Genomics (Luxembourg) for MiSeq sequencing, on the 5' end of forward primers 5'-ACACTCTTCCCTACACGACGCTCTCCGATCT-[primer]-3' and on the reverse primers 5'-GACTGGAGTTCAGACGTGTGCTCTCCGATCT-[primer]-3'. The length of the adaptors and the primers altogether ranged from 48 to 59 bp. The PCR conditions were the same for each primer set except for the annealing temperature (T_a) which varied according to the primer pairs. An initial denaturation at 98°C for 2 min, followed by 30 cycles of 98°C for 5 s, T_a for 20 s and 72°C for 10 s and a final extension of 72°C for 2 min. The 15 PCR products, one for each primer pair tested, were pooled by three leading to a total of five libraries. Sequencing was done by Eurofins Genomics (Luxembourg) with MiSeq 2 × 300 bp at a theoretical coverage of 120k reads per library.

2.3.3 | Data analysis

To evaluate the coverage and specificity of these barcodes under a realistic setup, we established a regional database for the Alsatian Flora (Eastern France) encompassing the four markers:

rbcL, *trnL*, ITS1 and ITS2. It contained sequences from the 268 species featured in the mock community, when available, as well as additional sequences of ~1600 species from the regional flora (Table S1, Appendix S1). This approach was chosen because not all species from the mock community had public reference sequences. The additional closely related species and genera in the regional database would allow more representative estimates of coverage and specificity that is a marker's capacity to discriminate between closely related species. Nonetheless, including only one species per genera in the mock community might still lead to overestimating the taxonomic resolution. Expanding the taxonomic representation within genera could be a valuable avenue for future research, allowing a more nuanced assessment of specificity in vitro. Using the NSDPY package (Hebert & Megléc, 2022), we queried all accessions from the EMBL database for the four markers for the two lists of species described earlier. The Illumina reads were analysed with OBITools for the clearing and filtering steps (merge, trim, demultiplex and dereplicate), before taxonomic assignment with the *blastn* function of BLAST (Altschul et al., 1990). Briefly, raw forward and reverse sequences were aligned with the function *illuminapairedend* with a minimum score of 30. Successfully aligned sequences were then recovered with *obigrep* while also saving joined sequences in a separate file. We use the *ngsfitler* algorithm to our advantage by virtually adding random tags to both ends of each sequence as only one library was generated. We then manually demultiplexed sequences based on the unique primer pairs with the *obigrep* function. Once assigned to their respective libraries, we dereplicated and denoised using *obiuniq* and *obiclean* respectively. Of the remaining sequence dataset, we only selected sequences above 40 bp with *obigrep*. The resulting sequences were taxonomically assigned with *blastn* from BLAST against DB1600 with a similarity threshold of 95%. Finally, the fasta and blast files were passed to MEGAN (Huson et al., 2007, 2016) where the BLAST results were filtered through a lowest common ancestor (LCA) algorithm. We retained only the most significant hits based on two filters: a top-percent filter that retained only those hits whose bit scores were within 1% of the best score, and a min-score filter that used case-specific bit score thresholds to further narrow down the results.

3 | RESULTS

3.1 | Common metabarcoding markers and primers

From the literature on plant eDNA metabarcoding (Table S2, Appendix S1), we identified 16 *rbcL* primer pairs ($n=32$ studies), 16 ITS2 ($n=34$), 9 ITS1 ($n=10$) and 2 *trnL* ($n=47$). In the case of *matK*, we only retrieved six metabarcoding studies using universal plant primers and HTS, and four primer pairs. Only a handful of primers were used across different applications, for instance, the combination of the forward primer ITS2-S2F (Chen et al., 2010) with BEL-3 (Chiou et al., 2007) and ITS4 (White et al., 1990) has been used in

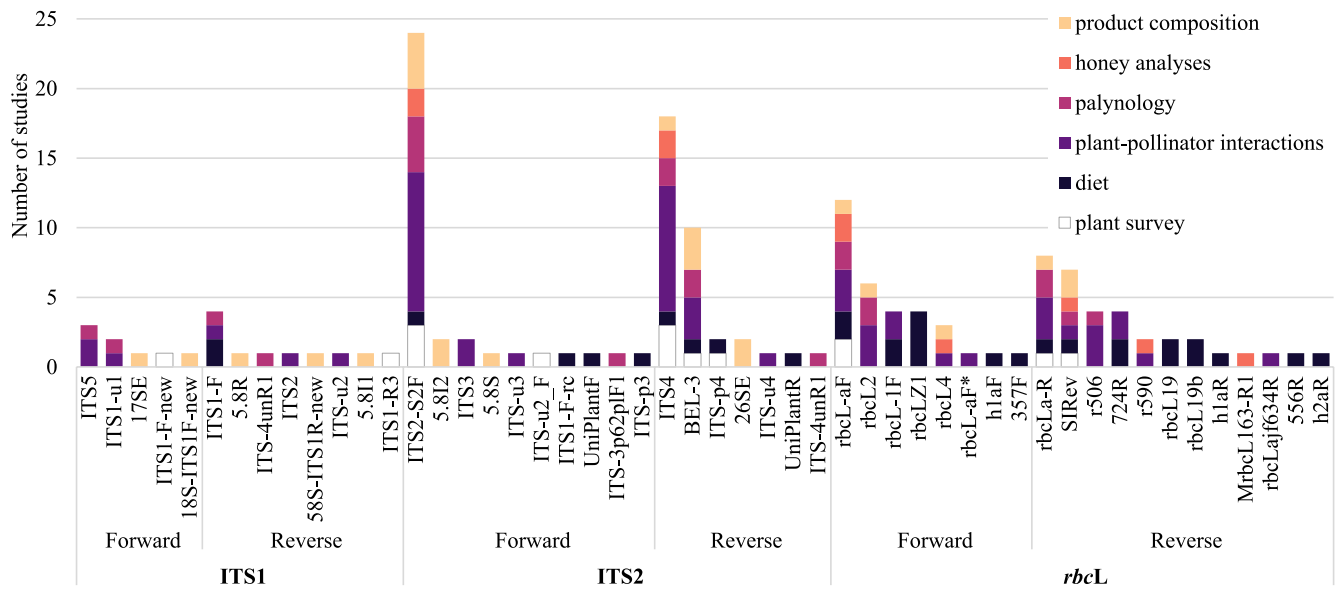


FIGURE 1 The use of plant primers across diverse applications for ITS1, ITS2 and *rbcL* markers. Novel metabarcoding primers were designed within specific research fields. *, primer variants.

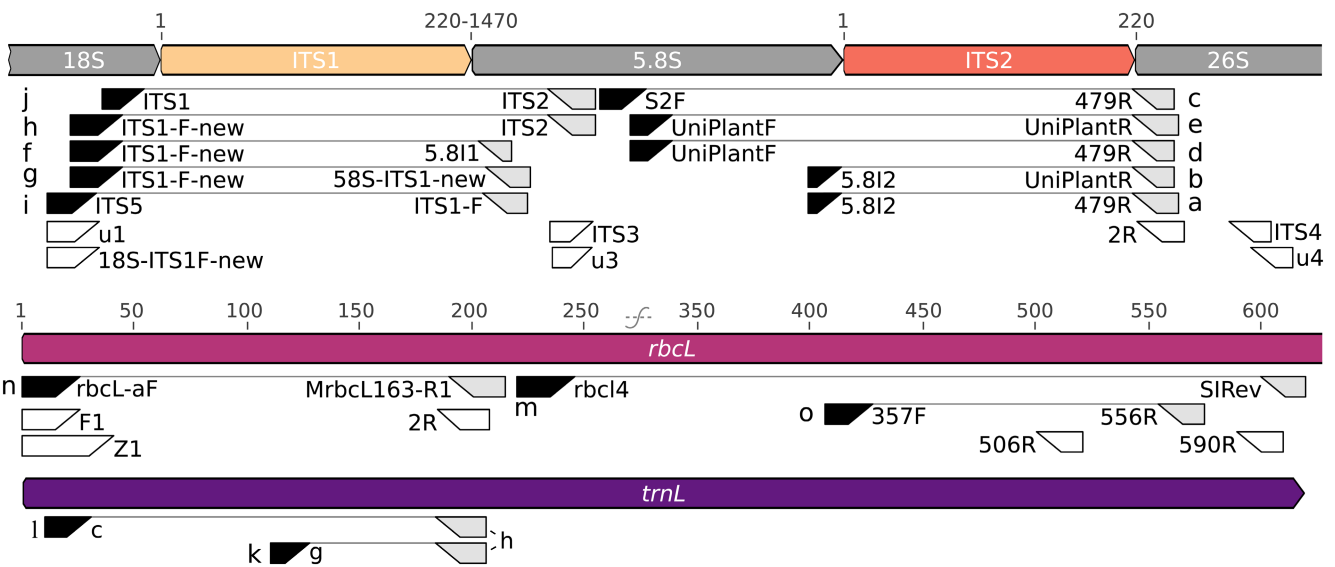


FIGURE 2 The map of the primers that passed the in silico tests. Primers pairs (linked) underwent subsequent in vitro testing (letters a–o).

plant eDNA surveys, honey and palynology analyses, studies on plant–pollination interactions, diet studies and product composition analyses (Figure 1). Similarly, *rbcL*-aF (Hasebe et al., 1994) together with *rbcL*-aR (Kress & Erickson, 2007) or with SI-Rev (Kress et al., 2009) have been used in the aforementioned applications. For the *trnL*, different combinations of primers c, d, g and h (Taberlet et al., 1991, 2007) have been used, but g–h was the most prevalent primer and was used for all applications. Metabarcoding primer pairs exclusive to plant eDNA surveys include the novel ITS1 primer pair ITS1-F/ITS1-R3 (Drummond et al., 2021) along with the ITS2 forward primer ITS-u2_F (Banchi et al., 2020) combined with the reverse ITS-p4 (Cheng et al., 2016). Additionally, the primer compilation includes 41 primer pairs designed and used in at least one study at the time

of our search. The inconsistent naming and referencing of primers throughout the literature was apparent from our compilation. The same primer can appear with different names; for instance, ITS1-R (Pornon et al., 2016) corresponds to primer ITS5 (White et al., 1990). An often-cited reverse primer for ITS2, referred to as ITS2-S3R, attributes its origin to Chen et al. (2010), though the primer was originally designed by Chiou et al. (2007) and named BEL-3. Although the original study was cited in Chen et al. (2010), subsequent studies using the primer pair only cited Chen et al. (2010). A similar situation arises with the primer *rbcL*-aF, variably referenced under different sources (CBOL Plant Working Group et al., 2009; de Vere et al., 2012; Kress & Erickson, 2007; Levin et al., 2003), but all are referring to the same primer *rbcL*-aF (Hasebe et al., 1994).

3.2 | In silico analyses

Among the 37 primer pairs for ITS1, 32 for ITS2, 29 for *rbcL* and 2 for *trnL* found in the literature (Table S3, Appendix S1), the summary statistics on ePCR amplicons allowed to retain 16, 10, 7 and

2 primer pairs respectively (Figure 2; Table S4, Appendix S1). The average amplicon lengths for the retained primer pairs ranged from 107 to 334 bp for *rbcL*, 49–147 bp for *trnL* (Figure 3), 281–330 bp for ITS1 and 225–309 bp for ITS2 (Figure 4). The highest standard deviation in amplicon length was observed for ITS1 (97–104 bp), followed

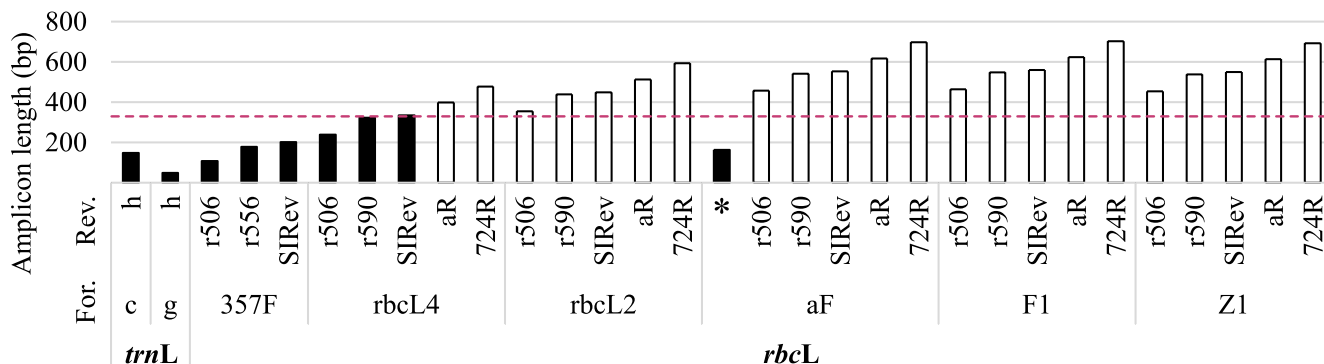


FIGURE 3 Mean amplicon lengths of plastid markers in silico. Primer pairs with an average amplicon length below the threshold of 330 bp (horizontal red line) were retained for subsequent screenings (black bars). *, MrbcL163-R1.

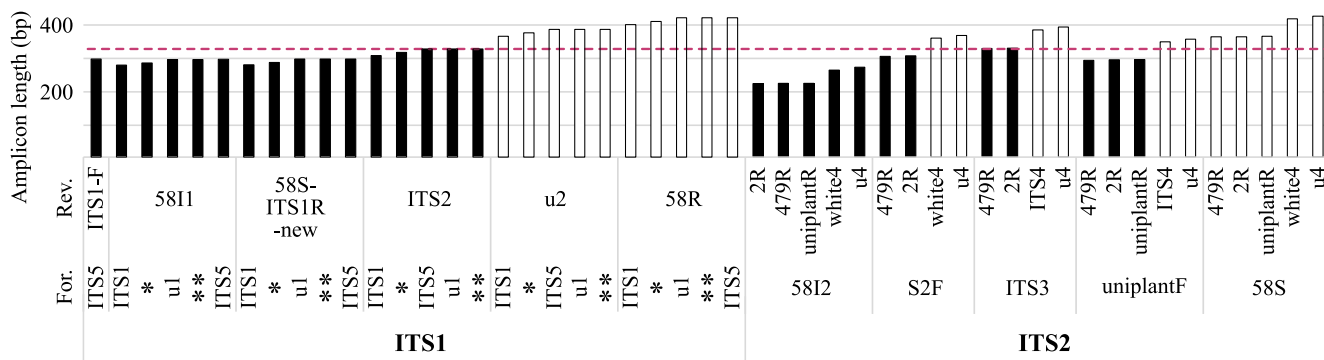


FIGURE 4 Mean amplicon lengths of ITS markers in silico. The primer pairs with an average amplicon length below the threshold of 315 bp (horizontal red line) were retained for subsequent screenings (black bars). *, ITS1-F-new; **, 18S-ITS1F-new.

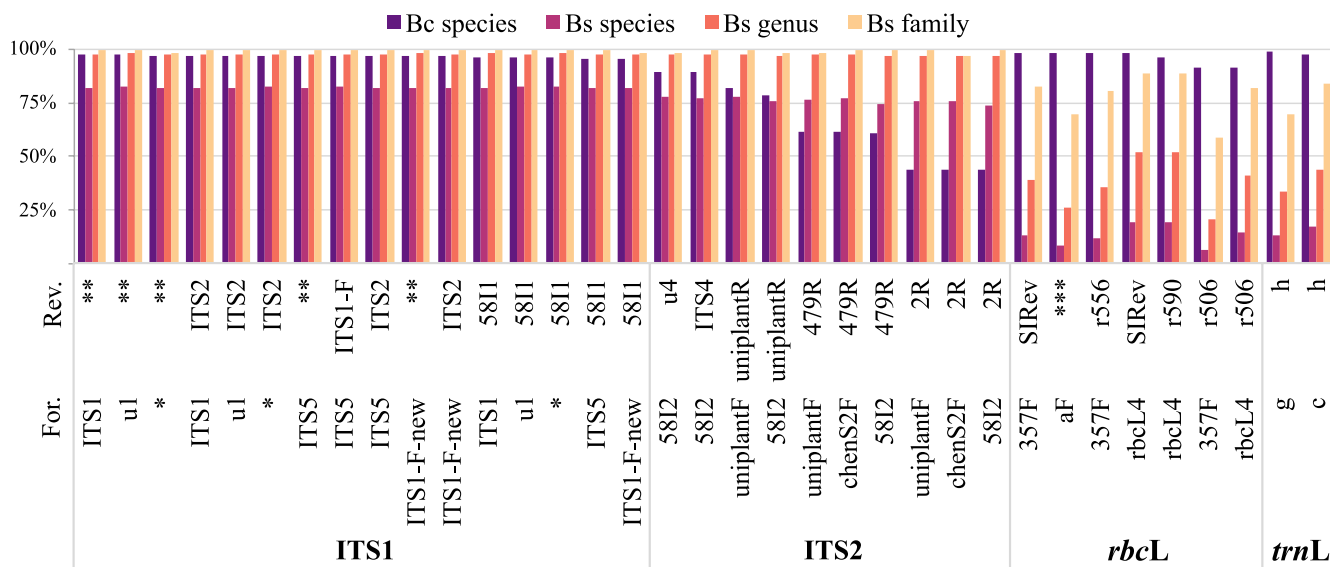
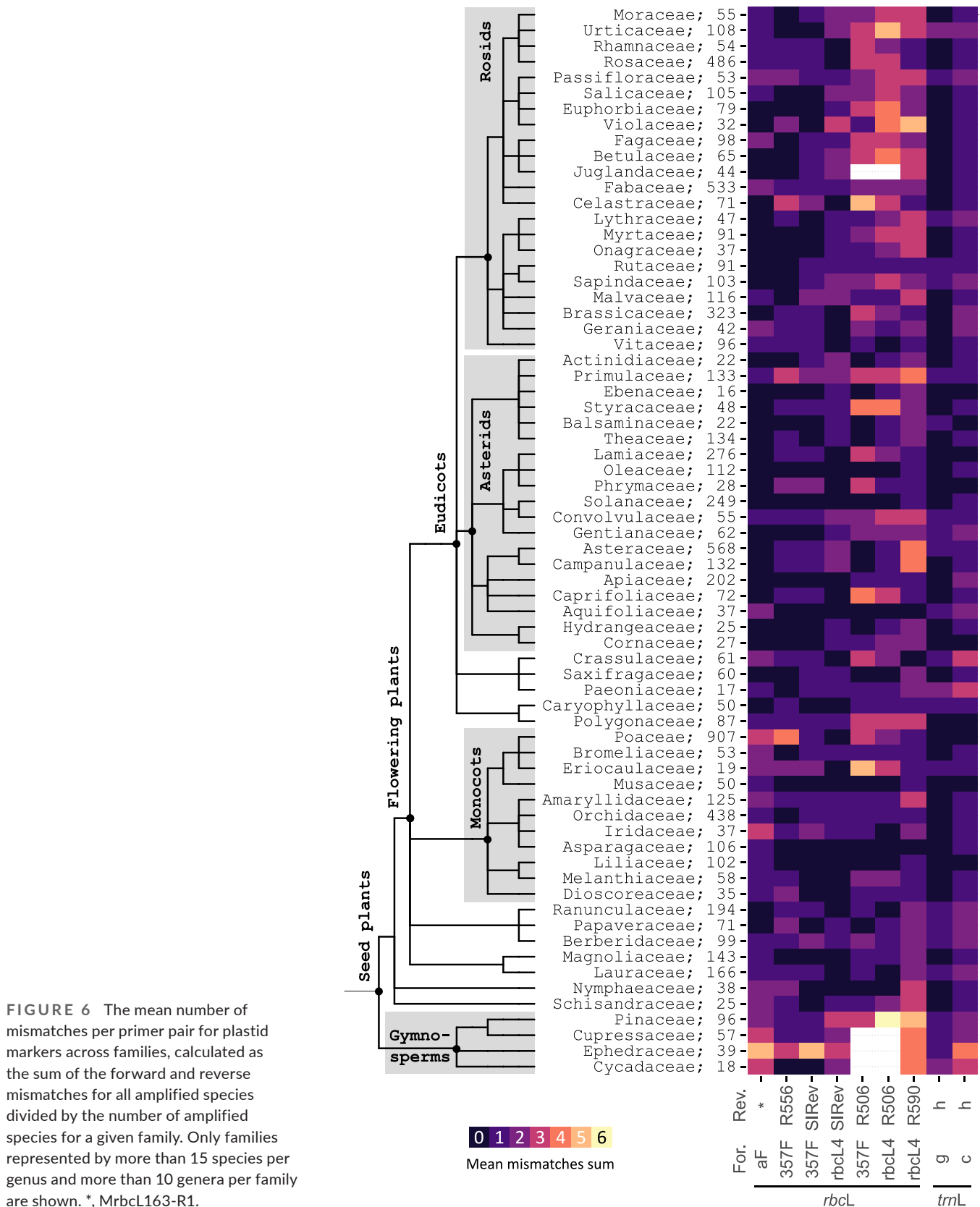


FIGURE 5 Coverage (Bc) and specificity (Bs) estimated in silico at different taxonomic ranks for the selected primer pairs. *, 18S-ITS1F-new; **, 58S-ITS1R-new; ***, MrbcL163-R1.

by ITS2 (32–47 bp), *trnL* (13 bp) and *rbcl* (0–3 bp). The *matK* primers exhibited coverage of up to 70% across the plastid database, with a mean amplicon length averaging 700 bp, too long for most eDNA studies. Consequently, *matK* was omitted from subsequent analyses.

The Bc values extrapolated from the ePCR output exhibited comparable outcomes among the top four *rbcl*, all amplifying above 97% of the species within the plastid database (Figure 5). Among these, the longer 334 bp barcode generated with *rbcl*L4+SIRev (Figure 3)



yielded superior Bs values across all taxonomic ranks (Figure 5). Conversely, the primer pair *rbcl*-af+MrbcL163-R1 achieved a remarkable 99% Bc but had the third lowest Bs value (8%). The performance of the two *trnL* primer pairs was on par with the best *rbcl* primer pairs in terms of Bc, while exhibiting moderate outcomes in terms of Bs values (13% and 17% for g+h and c+h respectively). The marginal higher Bc values for the *trnL* primer pair g+h (99%) over c+h (97%) is attributable to the non-amplification of ferns according to ePCR conditions (Appendix S2). The *trnL* barcodes obtained with c+h were twice as long as those generated with g+h (Figure 3) and had superior Bs values for all taxonomic ranks, peaking at 17% and 13% at the rank species respectively (Figure 5).

Despite the broad taxonomic coverage of *rbcl* and *trnL* primer pairs, the mean total number of mismatches varied greatly across families and was specific to each primer pair family (Figure 6). Non-flowering vascular plants (conifers and ferns) and mosses (Bryophyta) cumulated overall more mismatches than flowering plants. The forward *rbcl* primers 357F, *rbcl*L4 and *rbcl*-aF covered the majority of seed plants (Spermatophyta) with up to one mismatch, which could occur at different positions, and other seedless vascular plants and mosses with up to three mismatches (Appendix S2). In comparison, reverse primers SIRev exhibited a match to most Streptophyta with less than one mismatch, while primers R556, R506, R590 and MrbcL163-R1 required up to three mismatches. To reduce the number of mismatches and enhance PCR efficiency across seed plants and other lower Streptophyta, we propose the following modifications (in bold letters) to the original designs (primers forward (F) and reverse (R) oriented 5' to 3'): (F) 357F-bis ATTGTRGTAAYGDTTTGG, (F) *rbcl*L4-bis CTTACBAGYCTTGATCGTTAYAAAGG, (F) *rbcl*aF-bis ATGTCACAMAACAGARACTAAAGC, (R) 556R-bis ACAYTCATAAMMHGCTCTACC, and (R) SIRev-bis 5'-GTRAAATCAAGBCCACCRG-3'. The *trnL* primer pair g+h had fewer mismatches than c+h against seed plants. The forward primer c displayed a single mismatch with most Streptophyta, typically occurring at the third or seventh nucleotides from the 5' end. For the reverse primer h, while it perfectly matched most seed plants with less than one mismatch, we observed variations in some taxa at the third and fifth nucleotides from the 5' end. Their sequences could be revised as follows to incorporate our observations across more than 10,000 seed plants (Appendix S2): c-bis CGRAATYGGTAGACGCTACG and h-bis CCDTYGAGTCTCTGCACCTATC. The suggested alterations involving degenerate nucleotides for both *rbcl* and *trnL* primers should not affect their performance since the Ta remains within 5°C and no additional secondary structures or dimers were inferred.

The ITS1 primer pairs all amplified 97% of the species from the ITS database, encompassing over 5100 species, while the ITS2 primer pairs exhibited relatively lower Bc values, ranging from 44% to 90%. This observed trend is aligned with the taxonomic coverage of ITS1 primer pairs, which encompassed most Streptophyta, while ITS2 primer pairs mostly covered seed plants (Spermatophyta), hence the lower Bc values (Figure 7). Nonetheless, some flowering plant families exhibited more mismatches than others for both ITS1 and ITS2, such as Melianthaceae, Orchidaceae, Solanaceae, Brassicaceae and Malvaceae. Additionally, specific ITS2 mismatches

were observed within families such as Nymphaeaceae, Piperaceae and Crassulaceae. When assessing species-level Bs, the ITS1 barcodes unambiguously identified 82% of the species, while the ITS2 values ranged between 74% and 78% (Figure 5). At the genus rank, ITS1 and ITS2 achieved an identification rate of 97%–98%.

The primer pairs for the ITS2 marker could be categorized into three groups. The first group amplified over 75% of the species present within the ITS database, including UniPlantF + UniPlantR and three primer pairs with 5.8I2 (u4, ITS4 and UniPlantR). The highest Bc values were observed for the primer pairs 5.8I2+u4 (90%) and 5.8I2+ITS4 (89%), corresponding to the successful amplification of over 4700 species. However, these two primer pairs cumulated up to three mismatches in total, due to the forward primer 5.8I2 whose binding site lies closer to the ITS2 region (Figure 2). These mismatches located at the 3' end negatively affected the in vitro results for this primer. The reverse primers ITS4 and u4 shared half of their sequence (Figure 2) and perfectly matched most seed plants with at most one and two mismatches respectively (Appendix S2). In contrast, primer UniPlantF + UniPlantR had the third highest Bc and amplified most flowering plants without mismatches, conifers with two or more mismatches. Mosses were poorly amplified with our ePCR conditions (Figure 7; Appendix S2). Our findings are congruent with Kolter et al. (2021) observation that UniPlantF cumulates seemingly unnecessary degenerate nucleotides towards the 3' end. Although this does not seem to interfere with our in vitro evaluation, we suggest the following simplified version of the primer: UniPlantF-bis 5'-TGTGAATTGCAGRATCCCC-3'. Only u4 exhibited a perfect match with most mosses, consequently accounting for the reduced number of mismatches for the primer pair 5.8I1+u4 in comparison to 5.8I1+ITS4.

The second group exhibited a moderate species Bc of 60% and consisted of primer pairs featuring the reverse 479R. This primer is identical to a variant of the degenerate primer UniPlantR except for two fewer nucleotides on the 5' end (Figure 2; Appendix S2). Contrary to UniPlantR, the primer 479R does not have degenerate nucleotides and cumulates a minimum of two mismatches against most seed plants, while the degeneracy of UniPlantR allowed for a perfect match across most taxa (Figure 7). Kolter et al. (2021) reported that this version of the primer matched the most taxa against their database (c. 72%), a result that we confirmed with the reverse 479R when allowing up to four mismatches in the ePCR (76%). The forward primers ITS2-S2F and UniPlantF had seven nucleotides in common (Figure 2), but ITS2-S2F had one more mismatch for some families (Figure 7). Kolter et al. (2021) suggested an alternative to this primer by introducing degenerate nucleotides, and our results also support this suggestion.

The third group exhibited a species-level Bc of 44% and consisted of primer pairs with the reverse ITS-2R. The reverse primer ITS2-2R targeted the same binding site as 479R and UniPlantR (Figure 2). However, ITS2-2R possesses three additional nucleotides on the 5' end, which resulted in an increased number of mismatches surpassing the maximum threshold set for our ePCR conditions, subsequently leading to the omission of certain families. The 3' end of

these three primers played a pivotal role in either facilitating or inhibiting the amplification of other Streptophyta (Appendix S2). The G on the 3' end of ITS2-2R and 479R was present in most seed plants but was substituted by a degenerate nucleotide in UniPlantR.

Regarding the forward primers for the ITS1 marker, 18S-ITS1F-new (Omelchenko et al., 2019), ITS5 (White et al., 1990) and u1 (Cheng et al., 2016), all shared the identical 5'-GGAAGGAGAAGTCGTAACAAGG-3' primer binding site (Figure 2). The primers ITS5 and 18S-ITS1F-new differed by two nucleotides at positions six and eight, while u1 incorporated these differences as ambiguous bases (K and R; Appendix S2). We observed that the primer pairs featuring the forward u1 and 18S-ITS1F-new had a relatively smaller total number of mismatches across all Streptophyta, in comparison to primer pairs utilizing ITS5, which could accumulate up to five mismatches for some families like for Araceae, Solanaceae and Brassicaceae (Figure 7). Similarly, the forward primers ITS1 (White et al., 1990) and ITS1-F-new (Drummond et al., 2021) shared an overlap of nine nucleotides (Figure 2). They both perfectly matched most seed plants and mosses, though with up to two mismatches for certain families. The reverse primers ITS1-F (Baamrane et al., 2012) and 58S-ITS1R-new (Omelchenko et al., 2019) targeted the same binding site GATATCCGTTGCCGAGAGT, albeit with one-nucleotide shift (Figure 2). They also shared 12 nucleotides with 5.811 (Sun et al., 1994), which exhibited two mismatches on either end against seed plants (Appendix S2). We observed that the primer pairs best matching most Streptophyta were those with the reverse ITS1-F and 58S-ITS1R-new. However, we observed that the incorporation of a degenerate nucleotide (highlighted in bold) would reduce the number of mismatches without altering the primers' functionality: 58S-ITS1R-bis 5'-AGATATCCRTTGYCGAGAGT-3'. The 13th nucleotide could take a degenerate nucleotide Y, though most seed plants from our database had a C and this may not be necessary.

3.3 | In vitro analyses

A total of 15 primer pairs were evaluated in vitro (Figure 2). Among the 268 species present in the mock community, 233 were identified at the species level, 16 at the genus level (due to inter-specific barcodes in the reference database), while 19 were never retrieved with any marker (Table S5, Appendix S1). The species not referenced in the corresponding marker database were subtracted from the initial number of species in the mock community to determine the maximum expected number of species (N_{max}) that could be retrieved by a given primer pair. The highest proportion of retrieved species was observed for the ITS2 marker, with 184 (77%) species identified, followed by ITS1 (71%), *rbcl* (58%) and *trnL* (53%) with 170, 143 and 109 species respectively (Figure 8; Table S6, Appendix S1). However, the primer pairs of both ITS markers showed lower species retrieval rates in the in vitro evaluation compared to the in silico evaluation. All primer pairs for the ITS1 marker exhibited similar list of species from the mock community (Figure 8), indicating homogeneity in the in silico evaluation and indicative of their redundancy in terms of coverage. Compared

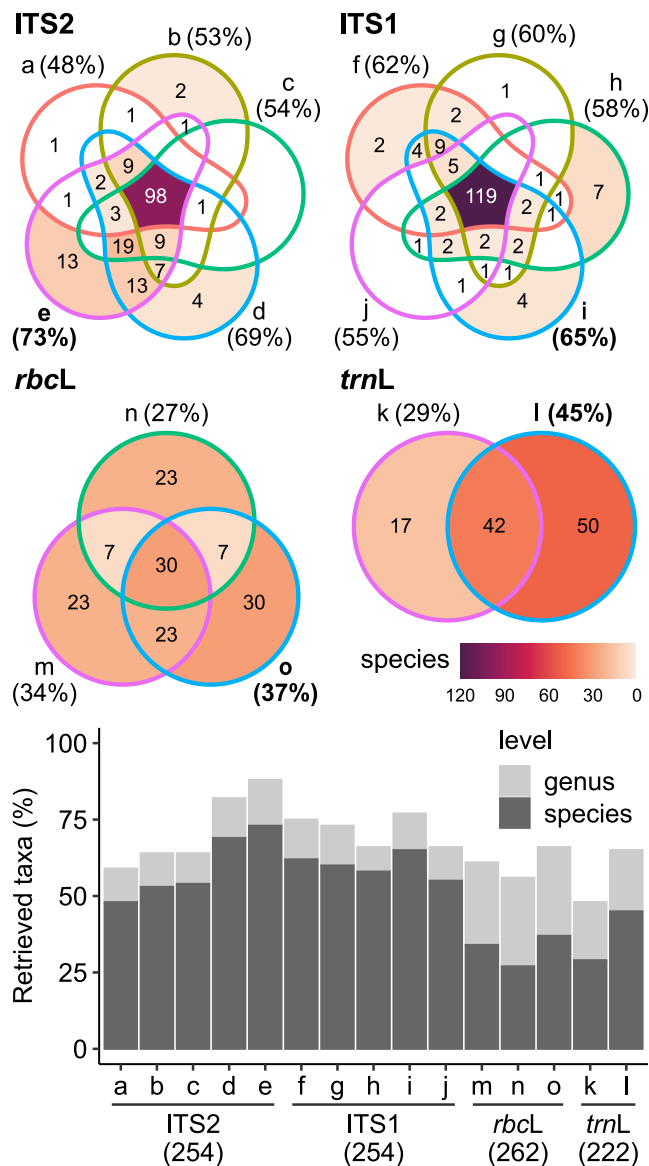


FIGURE 8 Combined analysis of Venn diagrams and bar graphs illustrating the species retrieval outcomes in vitro using different primer pairs and markers. The Venn diagrams display the number of species commonly and uniquely retrieved by each primer pair. The accompanying bar graphs provide the percentage of retrieved taxa at species and genus levels per primer pair relative to the maximum expected number of species (N_{max}) per marker. N_{max} represents the total number of species in the mock community, accounting for species absent from a marker's reference database. (a) 5812 + 479R; (b) 5812 + uniplantR; (c) chenS2F + 479R; (d) uniplantF + 479R; (e) uniplantF + uniplantR; (f) ITS1-F-new + 5811; (g) ITS1-F-new + 58S-ITS1R-new; (h) ITS1-F-new + ITS2; (i) ITS5 + ITS1-F; (j) ITS1 + ITS2; (k) g + h; (l) c + h; (m) *rbcl*4 + SIRev; (n) aF + MrbcL163-R1; (o) 357F + r556.

to the N_{max} for the ITS markers ($N_{max} = 254$ species), the ITS2 primer pair UniPlantF + UniPlantR (e) exhibited the highest species retrieval rate from the mock community (73%), as expected from the in silico evaluation (Figure 8; Table S7, Appendix S1). This primer retrieved 15 additional species to UniPlantF + 479R (d) which retrieved 69% of

the mock community at the same 62°C Ta as for the previous primer. Unexpectedly, primer pairs UniPlantF +479R (d) and ITS2-S2F+479R (c) retrieved some species within the Pinaceae family (Figure 9), contrary to the expectations based on the in silico results for UniPlantF + UniPlantR. The primer pairs a and b with the forward 5.812 retrieved the least number of species (~50%), a stark contrast from the in silico results. The poor performance of these primers in vitro could be attributed to their poor binding properties due to the observed number of mismatches for some families. Within the rosids, families such as Brassicaceae, Fabaceae and Euphorbiaceae were missed (Figure 9) as they cumulated more mismatches than others (4–5 mismatches for primer pair b). For the ITS1, the highest retrieval rate was observed for primer pair i (65%) closely followed by f (62%).

In contrast, plastid primer pairs exhibited a lower retrieval rate at the species level compared to the ITS markers as expected by the in silico analysis. Nevertheless, the percentage of retrieved species was higher than the in silico Bs values. The highest percentages were observed for *rbcL* when employing the primer pair 357F+R556 (o) with 37% ($N_{max}=262$), and for *trnL* using primer pair c+h (l) with 45% ($N_{max}=222$). Each of the three *rbcL* primer pairs exhibited unique taxonomic detection capabilities, capturing a distinct subset of species, with 9%–12% of species exclusively retrieved by a single primer pair.

The optimal combination of markers for achieving the highest percentage of retrieved species consistently included ITS2 primer

pair e (Figure 8). When comparing groups of two primer pairs, the highest species retrieval rate observed was 85% for e+m, though all combinations with e differed by at most 3%. Compared to the species retrieval rate of e alone (75%), the addition of another marker resulted in a notable 7%–10% increase in the percentage of retrieved species (Figure 10). Moreover, when comparing the highest performing combinations for groups two and three, we observed an additional 6% of newly retrieved species. The benefit of using all four markers becomes less quantifiable and more qualitative as the probability of detection reaches a plateau (Figure 10). The highest percentage of retrieved species was 97% for the combination of the four markers and five primer pairs e+i+l+m+o, including two *rbcL* primer pairs.

4 | DISCUSSION

4.1 | Which is the best marker for plant metabarcoding?

The research on plant eDNA metabarcoding is expanding, with new primers and different marker combinations being regularly proposed. Yet, this also hinders the selection of the most appropriate primers, especially in the absence of comprehensive comparative

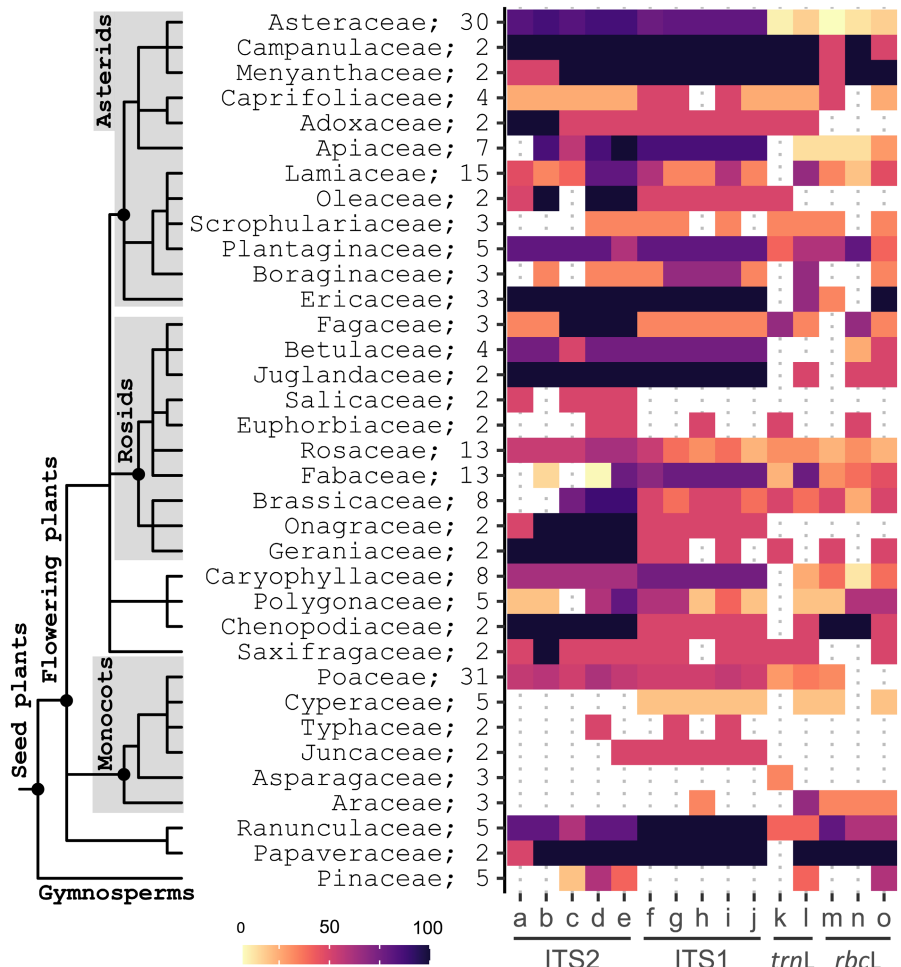


FIGURE 9 Species retrieval rate of primer pairs in vitro, calculated as the ratio of the number of retrieved species to the total number of species per family in the mock community. Only families with two or more species are shown.

from our analysis since metabarcoding studies reported poor amplification success (poor recovery and taxonomic coverage), long amplicons and inconsistent taxonomic resolution (Fahner et al., 2016; Yang et al., 2016).

The *trnL* primer pair g+h exhibits high coverage and generates short ~50bp amplicons for vascular plants, making it a preferred choice for environmental samples with highly degraded DNA, such as in diet and ancient DNA studies (Taberlet et al., 2007; Valentini, Miquel, et al., 2009; Willerslev et al., 2014). However, our ePCR and mock community results showed that primer pair c+h amplified an equal number of species while exhibiting higher taxonomic resolution, probably due to the additional ~100bp. This echoes earlier observations in eDNA studies using this primer pair for contemporary plant studies (Varotto et al., 2021; Yang et al., 2016), supporting its continued use in future eDNA research (Table 1).

We observed for all *rbcl* primer pairs broad coverage across the taxonomically rich plastid database with less than three mismatches, and highly conserved amplicon lengths. These observations collectively contribute to improving the success of both PCR and sequencing processes, aligning with the constrained evolution of the *rbcl*. We identified 16 primer pairs used for metabarcoding, among the most popular are *rbcl*-af+SIRev, *rbcl*-af+*rbcl*a-R, *rbcl*-aF+*rbcl*LR506 and *rbcl*2+*rbcl*a-R (Kress & Erickson, 2007; Palmieri et al., 2009) generating barcodes exceeding 500bp. These long barcodes may be ill-suited for some eDNA metabarcoding studies, though they yielded higher Bs values, for example 28% for 600bp barcodes against 19% for 330bp for the same region and with the same reverse primers. However, we focused on primer pairs that generated whole small barcodes (<300bp) in line with the prevalence of small fragments of eDNA and to take advantage of standard sequencing services (mainly paired

TABLE 1 Recommended primer pairs with the suggested modifications from this study (bold letters) and the original design sources.

Marker	Pair	Primer name	Direction	Primer sequence (5'→3')	Design study
<i>rbcl</i>	m	<i>rbcl</i> L4	Forward	CTTACCAGYCTTGATCGTTACAAAGG	Erickson et al. (2017)
		SIRev	Reverse	GTAATAAACAAGTCCACCRCG	Kress et al. (2009)
		<i>rbcl</i> L4-bis	Forward	CTTACBAGYCTTGATCGTTAYAAAGG	This work
		SIRev-bis	Reverse	GTRAAATCAAGBCCACCRCG	This work
	n	<i>rbcl</i> -aF	Forward	ATGTCACCACAAACAGAGACTAAAGC	Hasebe et al. (1994)
		Mr <i>rbcl</i> L163-R1	Reverse	CGGTCCAYACAGYBGTTCCAKGTACC	Ivanova et al. (2016)
		<i>rbcl</i> -aF-bis	Forward	ATGTCACCAMAAACAGARACTAAAGC	This work
		<i>rbcl</i> L163-bis	Reverse	GTCCAHCADGKTGTCAYGTACC	This work
	o	357F	FORWARD	CATTGTRGGTAATGTATTGG	Aziz et al. (2017)
		556R	Reverse	ACATTCATAAACHGCYCTACC	Aziz et al. (2017)
		357F-bis	Forward	ATTGTRGGTAAYGTDTTGG	This work
		556R-bis	Reverse	ACAYTCATAAMMHGCYCTACC	This work
<i>trnL</i>	k	g	Forward	GGGCAATCCTGAGCCAA	Taberlet et al. (2007)
	l	c	Forward	CGAAATCGGTAGACGCTACG	Taberlet et al. (1991)
	k, l	h	Reverse	CCATTGAGTCTCTGCACCTATC	Taberlet et al. (1991)
		g-bis	Forward	GGGYAATCCTGAGCCAA	This work
		c-bis	Forward	CGRAATYGGTAGACGCTACG	This work
		h-bis	Reverse	CCDTYGAGTCTCTGCACCTATC	This work
ITS1	g	ITS1-F-new	Forward	GTCGTAACAAGGTTTCCGTAGGT	Drummond et al. (2021)
		58S-ITS1R-new	Reverse	AGATATCCGTTGCCGAGAGT	Omelchenko et al. (2019)
		58S-ITS1R-bis	Reverse	AGATATCCRTTGYCGAGAGT	This work
		u1	Forward	GGAAGKARAAGTCGTAACAAGG	Cheng et al. (2016)
ITS2	e	UniplantF	Forward	TGTGAATTGCARRATYCMG	Moorhouse-Gann et al. (2018)
		UniplantR	Reverse	CCCGHYTGAYTGRGGTCDC	Moorhouse-Gann et al. (2018)
		UniplantF-bis	Forward	TGYGAATTGCAKRATYCCG	This work
		UniplantR-bis	Reverse	CYYGYCTGAYCTGRGGTCDC	This work
		ITS-3p62pIF1	Forward	ACBTRGTGTGAATTGCAGRATC	Kolter et al. (2021)
		ITS-4unR1	Reverse	TCCTCCGCTTATTKATATGC	Kolter et al. (2021)

150bp NGS sequencing). For instance, the primer pair *rbcl4*+SIRev amplified 334bp and showed the highest coverage and specificity from our ePCR and in vitro results and was successfully used in plant-pollinator interactions (Khansaritoreh et al., 2020; Tanaka et al., 2020) and product composition studies (Handy et al., 2021) supporting its continued use (Table 1). When comparing the two primer pairs amplifying the smallest *rbcl* barcodes tested here, we demonstrated that primer pair 357F+R556 was superior to *rbcl*-af+Mr*rbcl*163-R1 both in silico and in vitro, but can recommend both as they target complementary regions of the *rbcl* (Table 1). The latter generated the barcodes with the lowest species-level resolution in both evaluations, demonstrating that the first 163 nucleotides on the 5' end of the *rbcl* gene, part of the *rbcl*a region of the gene, are well conserved within families compared to the middle sections of the gene (*rbcl* b region). Our results are consistent with previous observations suggesting that the *rbcl* b's region was better suited for barcoding, partly because it is more variable (Dong et al., 2014). We advocate for the evaluation of barcodes from the *rbcl* b region as alternative candidates for metabarcoding. We need not start from scratch as reference databases are the same and the reverse primers evaluated here could be repurposed as forward primers paired with the reverse *rbcl*bR designed by Dong et al. (2014).

4.4 | Considerations for nrDNA ITS markers

The high coverage of ITS1 observed in silico comes at the cost of long barcodes (>700bp) for non-flowering land plants (mosses, ferns and gymnosperms). This significant length may not be suitable for environmental samples containing highly fragmented DNA, and it may also pose challenges for standard sequencing technologies, which typically cover a range of 200–500bp. Based on the ePCR results, the forward ITS1 primer u1 (Cheng et al., 2016) was the best alternative and recommended (Table 1) over ITS5 (White et al., 1990) and 18S-ITS1F-new (Omelchenko et al., 2019), because it accounts for the differences between both using degenerate nucleotides. However, by design, u1 can still perfectly match fungi sequences in silico and should be paired with plant-specific reverse primers to preclude undesired co-amplifications (Cheng et al., 2016; Kolter & Gemeinholzer, 2021). From our results, we also recommend primer ITS1-F-new (Drummond et al., 2021) as an alternative to u1 (Table 1). This primer does not require degenerate nucleotides to match most Streptophyta and laps over u1 and ITS1 (White et al., 1990), avoiding altogether the stable hairpin of ITS1 (Kolter et al., 2021). The hairpin structure of ITS1 was not hindering in our study because we used a 68°C Ta which is above its melting temperature. Together with the previously recommended forward primers, we suggest the use of the following reverse primer 58S-ITS1R-bis 5'-AGATATCCRTTGCGAGAGT-3' (Table 1) which is a modification of ITS1-F (Baamrane et al., 2012) and 58S-ITS1R-new (Omelchenko et al., 2019).

The ITS2 primer pair UniPlantF + UniPlantR (Moorhouse-Gann et al., 2018) showed promising outcomes in both our in silico and in vitro evaluations, and in accordance with prior research findings

supporting its continued use (Table 1) (Castle et al., 2020; Kolter & Gemeinholzer, 2021; Meyer et al., 2020; Moorhouse-Gann et al., 2022; and except Timpano et al., 2020). During our in silico evaluation, we did not retrieve Kolter et al. (2021) work on ITS2 primers, and the primer pair ITS-3p62pIF1 + ITS-4unR1 developed therein could not be evaluated in our study. Primer ITS-3p62pIF1 overlaps UniPlantF and is an alternative to the commonly used ITS2-S2F (Chen et al., 2010), which perfectly matched most taxa from our database during in silico analysis. The incorporation of degenerate nucleotides proposed by Kolter et al. (2021) aligns with our observations regarding ITS2-S2F, potentially enhancing the primer's universality. Their primer pair ITS-3p62pIF1 + ITS-4unR1 should be a valuable contribution to eDNA metabarcoding but remain to be used in eDNA studies to ascertain its effectiveness (Table 1).

4.5 | The more the better? The advantages of a multi-marker approach

Assessing and comparing primer pairs are only half of the equation. Our findings underscore the requirement of a multi-marker approach to enhance species-level taxonomic resolution. While the ITS2 emerges from our results as an indispensable marker for seed plants, it falls short for other Streptophyta taxa such as horsetails, mosses and ferns. In this regard, the higher complementarity between ITS2 and *rbcl* yielded superior species retrieval rates than the more congruent ITS2 and ITS1. For bio-monitoring and contemporary ecological studies targeting most land plants, we suggest the *rbcl* as it offers small barcodes and superior complementarity. The *rbcl* amplified and identified species of mosses and other seedless vascular plants in both evaluations as it is known to have superior inter-specific genetic variability for these groups (Dong et al., 2014; Li et al., 2011; Liu et al., 2010). In contrast, the ITS2 has superior species-level taxonomic resolution for seed plants (Fahner et al., 2016; Prosser & Hebert, 2017). A three-marker approach, while costlier, increased species retrieval by 6% from our mock community with ITS2, *rbcl*, and ITS1 compared to ITS2 + *rbcl*. In addition to the superior probability of retrieving a species, the increased congruence over a breadth of taxa instils greater confidence in detection. Additionally, in scenarios with incomplete reference databases, the use of three or four markers would be beneficial as it increases the probability that the species is represented in at least one database. However, a multi-marker approach does not compensate for the necessity for a curated regional reference database for taxonomic assignment, particularly if species-level assignment is required. Bioinformatic pipelines could enhance interpretability by integrating confidence indices. For instance, incorporating prior-known marker bias to a group of taxa or weighing species similarities between markers can assist in multi-marker metabarcoding studies. This approach would benefit eDNA biodiversity assessments and increase adoption by practitioners seeking standard and comprehensive results.

4.6 | Perspectives

By comparing new combinations of primers, re-evaluating established ones, suggesting changes and identifying functional combinations of markers we contribute to laying the foundations for the standardization of plant metabarcoding. We strongly advocate for the use of the same markers and primers across studies to facilitate comparability, transferability and reproducibility. It should help us gain a better understanding of the underlying processes of eDNA in different environments by reducing the stochasticity due to the molecular tools. The standardization of markers and primers has major implications for biodiversity monitoring as we cannot afford differences between studies to be defined by the improper selection of markers and primers.

AUTHOR CONTRIBUTIONS

AEP, LH and JNB conceived the study. AEP scanned the literature and conducted the laboratory work. AEP and ND contributed to bioinformatics. AEP wrote the initial draft, and all co-authors improved the final version of the manuscript.

ACKNOWLEDGEMENTS

This work was supported by the French State as part of the Projet de territoire de Fessenheim, OHM Fessenheim - University of Strasbourg. The authors acknowledge the High Performance Computing Center of the University of Strasbourg for supporting this work by providing scientific support and access to computing resources. Part of the computing resources were funded by the Equipex Equip@Meso project (Programme Investissements d'Avenir) and the CPER Alsacalcul/Big Data.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Raw sequence reads and related metadata are deposited in the SRA (BioProject PRJNA1065652).

ORCID

Laurent Hardion  <https://orcid.org/0000-0002-4574-6354>

REFERENCES

- Alsos, I. G., Lammers, Y., Yoccoz, N. G., Jørgensen, T., Sjögren, P., Gielly, L., & Edwards, M. E. (2018). Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. *PLoS One*, 13(4), e0195403. <https://doi.org/10.1371/journal.pone.0195403>
- Anthoens, B., Karamichali, I., Schröder-Nielsen, A., Drouzas, A. D., de Boer, H., & Madesis, P. (2021). Metabarcoding reveals low fidelity and presence of toxic species in short chain-of-commercialization of herbal products. *Journal of Food Composition and Analysis*, 97, 103767. <https://doi.org/10.1016/j.jfca.2020.103767>
- Ariza, M., Fouks, B., Mauvisseau, Q., Halvorsen, R., Alsos, I. G., & de Boer, H. J. (2023). Plant biodiversity assessment through soil eDNA reflects temporal and local diversity. *Methods in Ecology and Evolution*, 14, 415–430. <https://doi.org/10.1111/2041-210X.13865>
- Arulandhu, A. J., Staats, M., Hagelaar, R., Peelen, T., & Kok, E. J. (2019). The application of multi-locus DNA metabarcoding in traditional medicines. *Journal of Food Composition and Analysis*, 79, 87–94. <https://doi.org/10.1016/j.jfca.2019.03.007>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/s0022-2836\(05\)80360-2](https://doi.org/10.1016/s0022-2836(05)80360-2)
- Aziz, S., Clements, G., Peng, L., Campos-Arceiz, A., McConkey, K., Forget, P., & Gan, H. (2017). Elucidating the diet of the Island flying fox (*Pteropus hypomelanus*) in peninsular Malaysia through Illumina next-generation sequencing. *PeerJ*, 5, e3176. <https://doi.org/10.7717/peerj.3176>
- Baamrane, M. A. A., Shehzad, W., Ouhammou, A., Abbad, A., Naimi, M., Coissac, E., Taberlet, P., & Znari, M. (2012). Assessment of the food habits of the Moroccan Dorcas gazelle in M'Sabih Talaa, west Central Morocco, using the trnL approach. *PLoS One*, 7(4), e35643. <https://doi.org/10.1371/journal.pone.0035643>
- Baensch, S., Tschardtke, T., Wuenschiers, R., Netter, L., Brenig, B., Gabriel, D., & Westphal, C. (2020). Using ITS2 metabarcoding and microscopy to analyse shifts in pollen diets of honey bees and bumble bees along a mass-flowering crop gradient. *Molecular Ecology*, 29(24), 5003–5018. <https://doi.org/10.1111/mec.15675>
- Banchi, E., Ametrano, C. G., Tordoni, E., Stankovi, D., Ongaro, S., Tretiac, M., Pallavicini, A., Muggia, L., Verardo, P., Tassan, F., Trobiani, N., Moretti, O., Borney, M. F., & Lazzarin, S. (2020). Environmental DNA assessment of airborne plant and fungal seasonal diversity. *Science of the Total Environment*, 738, 140249. <https://doi.org/10.1016/j.scitotenv.2020.140249>
- Bell, K. L., Burgess, K. S., Botsch, J. C., Dobbs, E. K., Read, T. D., & Brosi, B. J. (2019). Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Molecular Ecology*, 28(2), 431–455. <https://doi.org/10.1111/mec.14840>
- Bell, K. L., Fowler, J., Burgess, K. S., Dobbs, E. K., Gruenewald, D., Lawley, B., Morozumi, C., & Brosi, B. J. (2017). Applying pollen DNA metabarcoding to the study of plant–pollinator interactions. *Applications in Plant Sciences*, 5(6), 1600124. <https://doi.org/10.3732/apps.1600124>
- Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., & Kausrud, H. (2010). ITS as an environmental DNA barcode for fungi: An in silico approach reveals potential PCR biases. *BMC Microbiology*, 10(1), 189. <https://doi.org/10.1186/1471-2180-10-189>
- Boyer, F., Mercier, C., Bonin, A., Bras, Y. L., Taberlet, P., & Coissac, E. (2016). Obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16(1), 176–182. <https://doi.org/10.1111/1755-0998.12428>
- Brandon-Mong, G.-J., Gan, H.-M., Sing, K.-W., Lee, P.-S., Lim, P.-E., & Wilson, J.-J. (2015). DNA metabarcoding of insects and allies: An evaluation of primers and pipelines. *Bulletin of Entomological Research*, 105(6), 717–727. <https://doi.org/10.1017/S0007485315000681>
- Bruno, A., Sandionigi, A., Agostinetto, G., Bernabovi, L., Frigerio, J., Casiraghi, M., & Labra, M. (2019). Food tracking perspective: DNA metabarcoding to identify plant composition in complex and processed food products. *Genes*, 10(3), 248. <https://doi.org/10.3390/genes10030248>
- Bylemans, J., Furlan, E. M., Gleeson, D. M., Hardy, C. M., & Duncan, R. P. (2018). Does size matter? An experimental evaluation of the relative abundance and decay rates of aquatic environmental DNA. *Environmental Science & Technology*, 52(11), 6408–6416. <https://doi.org/10.1021/acs.est.8b01071>
- Bylemans, J., Gleeson, D. M., Hardy, C. M., & Furlan, E. (2018). Toward an ecoregion scale evaluation of eDNA metabarcoding primers: A case study for the freshwater fish biodiversity of the Murray-Darling

- basin (Australia). *Ecology and Evolution*, 8(17), 8697–8712. <https://doi.org/10.1002/ece3.4387>
- Cannon, M. V., Hester, J., Shalkhauser, A., Chan, E. R., Logue, K., Small, S. T., & Serre, D. (2016). In silico assessment of primers for eDNA studies using PrimerTree and application to characterize the biodiversity surrounding the Cuyahoga River. *Scientific Reports*, 6, 22908. <https://doi.org/10.1038/srep22908>
- Capo, E., Giguet-Covex, C., Rouillard, A., Nota, K., Heintzman, P. D., Vuillemin, A., Ariztegui, D., Arnaud, F., Belle, S., Bertilsson, S., Bigler, C., Bindler, R., Brown, A. G., Clarke, C. L., Crump, S. E., Debroas, D., Englund, G., Ficetola, G. F., Garner, R. E., ... Parducci, L. (2021). Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: Overview and recommendations. *Quaternary*, 4(1), 6. <https://doi.org/10.3390/quat4010006>
- Castle, S. T., Allan, N., Clifford, D., Aylward, C. M., Ramsey, J., Fascetti, A. J., Pesapane, R., Roy, A., Statham, M., Sacks, B., & Foley, J. (2020). Diet composition analysis provides new management insights for a highly specialized endangered small mammal. *PLoS One*, 15(10), e0240136. <https://doi.org/10.1371/journal.pone.0240136>
- CBOL Plant Working Group, Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M. W., Cowan, R. S., Erickson, D. L., Fazekas, A. J., Graham, S. W., James, K. E., Kim, K.-J., Kress, W. J., Schneider, H., van AlphenStahl, J., Barrett, S. C. H., van den Berg, C., ... Little, D. P. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*, 106(31), 12794–12797. <https://doi.org/10.1073/pnas.0905845106>
- Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y., & Leon, C. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One*, 5(1), e8613. <https://doi.org/10.1371/journal.pone.0008613>
- Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., & Zhou, S. (2016). Barcoding the kingdom plantae: New PCR primers for ITS regions of plants with improved universality and specificity. *Molecular Ecology Resources*, 16(1), 138–149. <https://doi.org/10.1111/1755-0998.12438>
- China Plant BOL Group, Li, D.-Z., Gao, L.-M., Li, H.-T., Wang, H., Ge, X.-J., Liu, J.-Q., Chen, Z.-D., Zhou, S.-L., Chen, S.-L., Yang, J.-B., Fu, C.-X., Zeng, C.-X., Yan, H.-F., Zhu, Y.-J., Sun, Y.-S., Chen, S.-Y., Zhao, L., Wang, K., ... Duan, G.-W. (2011). Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National Academy of Sciences*, 108(49), 19641–19646. <https://doi.org/10.1073/pnas.1104551108>
- Chiou, S.-J., Yen, J.-H., Fang, C.-L., Chen, H.-L., & Lin, T.-Y. (2007). Authentication of medicinal herbs using PCR-amplified ITS2 with specific primers. *Planta Medica*, 73(13), 1421–1426. <https://doi.org/10.1055/s-2007-990227>
- Coghlan, S. A., Shafer, A. B. A., & Freeland, J. R. (2020). Development of an environmental DNA metabarcoding assay for aquatic vascular plant communities. *Environmental DNA*, 3(2), 372–387. <https://doi.org/10.1002/edn3.120>
- da Silva, L. P., Mata, V. A., Lopes, P. B., Pereira, P., Jarman, S. N., Lopes, R. J., & Beja, P. (2019). Advancing the integration of multi-marker metabarcoding data in dietary analysis of trophic generalists. *Molecular Ecology Resources*, 19(6), 1420–1432. <https://doi.org/10.1111/1755-0998.13060>
- De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources*, 14(2), 306–323. <https://doi.org/10.1111/1755-0998.12188>
- de Vere, N., Rich, T. C. G., Ford, C. R., Trinder, S. A., Long, C., Moore, C. W., Satterthwaite, D., Davies, H., Allainguillaume, J., Ronca, S., Tatarinova, T., Garbett, H., Walker, K., & Wilkinson, M. J. (2012). DNA barcoding the native flowering plants and conifers of Wales. *PLoS One*, 7(6), e37945. <https://doi.org/10.1371/journal.pone.0037945>
- Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters*, 10(9), 20140562. <https://doi.org/10.1098/rsbl.2014.0562>
- Dong, W., Cheng, T., Li, C., Xu, C., Long, P., Chen, C., & Zhou, S. (2014). Discriminating plants using the DNA barcode rbcLb: An appraisal based on a large data set. *Molecular Ecology Resources*, 14(2), 336–343. <https://doi.org/10.1111/1755-0998.12185>
- Drummond, J. A., Larson, E. R., Li, Y., Lodge, D. M., Gantz, C. A., Pfrender, M. E., Renshaw, M. A., Correa, A. M. S., & Egan, S. P. (2021). Diversity metrics are robust to differences in sampling location and depth for environmental DNA of plants in Small Temperate Lakes. *Frontiers in Environmental Science*, 9, 617924. <https://doi.org/10.3389/fenvs.2021.617924>
- Elbrecht, V., Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Hajibabaei, M., Wright, M., Zakharov, E. V., Hebert, P. D. N., & Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, 7, e7745. <https://doi.org/10.7717/peerj.7745>
- Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass – Sequence relationships with an innovative metabarcoding protocol. *PLoS One*, 10(7), e0130324. <https://doi.org/10.1371/journal.pone.0130324>
- Elbrecht, V., & Leese, F. (2017a). PrimerMiner: An R package for development and in silico validation of DNA metabarcoding primers. *Methods in Ecology and Evolution*, 8(5), 622–626. <https://doi.org/10.1111/2041-210X.12687>
- Elbrecht, V., & Leese, F. (2017b). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science*, 5, 11. <https://doi.org/10.3389/fenvs.2017.00011>
- Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., Erséus, C., Gusarov, V. I., Edwards, M. E., Johnsen, A., Stenøien, H. K., Hassel, K., Kausarud, H., Yoccoz, N. G., Bråthen, K. A., Willerslev, E., Taberlet, P., Coissac, E., & Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Molecular Ecology*, 21(8), 1821–1833. <https://doi.org/10.1111/j.1365-294X.2012.05537.x>
- Erickson, D. L., Reed, E., Ramachandran, P., Bourg, N. A., McShea, W. J., & Ottesen, A. (2017). Reconstructing a herbivore's diet using a novel rbcL DNA mini-barcode for plants. *AoB Plants*, 9(3), plx015. <https://doi.org/10.1093/aobpla/plx015>
- Espinosa Prieto, A., Beisel, J. N., Verschuren, P., & Hardion, L. (2023). Toward freshwater plant diversity surveys with eDNA barcoding and metabarcoding. *Environmental DNA*, 5(4), 648–670. <https://doi.org/10.1002/edn3.407>
- Fahner, N. A., Shokralla, S., Baird, D. J., & Hajibabaei, M. (2016). Large-scale monitoring of plants through environmental DNA metabarcoding of soil: Recovery, resolution, and annotation of four DNA markers. *PLoS One*, 11(6), e0157505. <https://doi.org/10.1371/journal.pone.0157505>
- Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., Percy, D. M., Hajibabaei, M., & Barrett, S. C. H. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One*, 3(7), e2802. <https://doi.org/10.1371/journal.pone.0002802>
- Ficetola, G. F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P., & Pompanon, F. (2010). An in silico approach for the evaluation of DNA barcodes. *GMC Genomics*, 11(1), 434. <https://doi.org/10.1186/1471-2164-11-434>
- Freeland, J. R. (2017). The importance of molecular markers and primer design when characterizing biodiversity from environmental DNA. *Genome*, 60(4), 358–374. <https://doi.org/10.1139/gen-2016-0100>

- Frigerio, J., Agostinetto, G., Sandionigi, A., Mezzasalma, V., Berterame, N. M., Casiraghi, M., Labra, M., & Galimberti, A. (2020). The hidden 'plant side' of insect novel foods: A DNA-based assessment. *Food Research International*, 128, 108751. <https://doi.org/10.1016/j.foodres.2019.108751>
- Fuertes Aguilar, J., Rossello, J. A., & Nieto Feliner, G. (1999). Nuclear ribosomal DNA (nrDNA) concerted evolution in natural and artificial hybrids of *Armeria* (Plumbaginaceae). *Molecular Ecology*, 8(8), 1341–1346. <https://doi.org/10.1046/j.1365-294X.1999.00690.x>
- Gaggini, L., Rusterholz, H.-P., & Baur, B. (2019). The invasion of an annual exotic plant species affects the above- and belowground plant diversity in deciduous forests to a different extent. *Perspectives in Plant Ecology, Evolution and Systematics*, 38, 74–83. <https://doi.org/10.1016/j.ppees.2019.04.004>
- Garcés-Pastor, S., Coissac, E., Lavergne, S., Schwörer, C., Theurillat, J.-P., Heintzman, P. D., Wangenstein, O. S., Tinner, W., Rey, F., Heer, M., Rutzer, A., Walsh, K., Lammers, Y., Brown, A. G., Goslar, T., Rijal, D. P., Karger, D. N., Pellissier, L., Heiri, O., ... Alsos, I. G. (2022). High resolution ancient sedimentary DNA shows that alpine plant diversity is associated with human land use and climate change. *Nature Communications*, 13(1), 6559. <https://doi.org/10.1038/s41467-022-34010-4>
- Gargominy, O., Terceire, S., Régnier, C., Ramage, T., Dupont, P., Daszkiewicz, P., & Poncet, L. (2021). TAXREF v15, référentiel taxonomique pour la France: Méthodologie, mise en œuvre et diffusion. In *Muséum national d'Histoire naturelle* (pp. 1–63). Rapport UMS PatriNat (OFB-CNRS-MNHN).
- Handy, S. M., Pawar, R. S., Ottesen, A. R., Ramachandran, P., Sagi, S., Zhang, N., Hsu, E., & Erickson, D. L. (2021). HPLC-UV, metabarcoding and genome skims of botanical dietary supplements: A case study in *Echinacea*. *Planta Medica*, 87(4), 314–324. <https://doi.org/10.1055/a-1336-1685>
- Hasebe, M., Omori, T., Nakazawa, M., Sano, T., Kato, M., & Iwatsuki, K. (1994). RbcL gene sequences provide evidence for the evolutionary lineages of leptosporangiate ferns. *Proceedings of the National Academy of Sciences*, 91(12), 5730–5734. <https://doi.org/10.1073/pnas.91.12.5730>
- Hebert, R., & Megléc, E. (2022). NSDPY: A python package to download DNA sequences from NCBI. *SoftwareX*, 18, 101038. <https://doi.org/10.1016/j.softx.2022.101038>
- Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). Choosing and using a plant DNA barcode. *PLoS One*, 6(5), e19254. <https://doi.org/10.1371/journal.pone.0019254>
- Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Research*, 17(3), 377–386. <https://doi.org/10.1101/gr.5969107>
- Huson, D. H., Beier, S., Flade, I., Górski, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., & Tappu, R. (2016). MEGAN community edition—Interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Computational Biology*, 12(6), e1004957. <https://doi.org/10.1371/journal.pcbi.1004957>
- Ivanova, N. V., Kuzmina, M. L., Braukmann, T. W. A., Borisenko, A. V., & Zakharov, E. V. (2016). Authentication of herbal supplements using next-generation sequencing. *PLoS One*, 11(12), e0156426. <https://doi.org/10.1371/journal.pone.0156426>
- Jorgensen, T., Kjaer, K. H., Haile, J., Rasmussen, M., Boessenkool, S., Andersen, K., Coissac, E., Taberlet, P., Brochmann, C., Orlando, L., Gilbert, M. T. P., & Willerslev, E. (2012). Islands in the ice: Detecting past vegetation on Greenlandic nunataks using historical records and sedimentary ancient DNA meta-barcoding. *Molecular Ecology*, 21(8), 1980–1988. <https://doi.org/10.1111/j.1365-294X.2011.05278.x>
- Kelly, R. P., Shelton, A. O., & Gallego, R. (2019). Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Scientific Reports*, 9(1), 12133. <https://doi.org/10.1038/s41598-019-48546-x>
- Khansaritoreh, E., Salmaki, Y., Ramezani, E., Azirani, T. A., Keller, A., Neumann, K., Alizadeh, K., Zarre, S., Beckh, G., & Behling, H. (2020). Employing DNA metabarcoding to determine the geographical origin of honey. *Heliyon*, 6(11), e05596. <https://doi.org/10.1016/j.heliyon.2020.e05596>
- Kolter, A., & Gemeinholzer, B. (2021). Internal transcribed spacer primer evaluation for vascular plant metabarcoding. *Metabarcoding and Metagenomics*, 5, e68155. <https://doi.org/10.3897/mbmg.5.68155>
- Kolter, A., Gemeinholzer, B., & Boatwright, J. S. (2021). Plant DNA barcoding necessitates marker-specific efforts to establish more comprehensive reference databases. *Genome*, 64(3), 265–298. <https://doi.org/10.1139/gen-2019-0198>
- Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, 7(1), 17668. <https://doi.org/10.1038/s41598-017-17333-x>
- Kress, W. J., & Erickson, D. L. (2007). A two-locus global DNA barcode for land plants: The coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS One*, 2(6), e508. <https://doi.org/10.1371/journal.pone.0000508>
- Kress, W. J., Erickson, D. L., Jones, F. A., Swenson, N. G., Perez, R., Sanjurjo, O., & Bermingham, E. (2009). Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceedings of the National Academy of Sciences*, 106(44), 18621–18626. <https://doi.org/10.1073/pnas.0909820106>
- Kress, W. J., Garcia-Robledo, C., Uriarte, M., & Erickson, D. L. (2015). DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology & Evolution*, 30(1), 25–35. <https://doi.org/10.1016/j.tree.2014.10.008>
- Kuzmina, M. L., Braukmann, T. W. A., & Zakharov, E. V. (2018). Finding the pond through the weeds: EDNA reveals underestimated diversity of pondweeds. *Applications in Plant Sciences*, 6(5), e01155. <https://doi.org/10.1002/aps3.1155>
- Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, 10(1), 34. <https://doi.org/10.1186/1742-9994-10-34>
- Levin, R. A., Wagner, W. L., Hoch, P. C., Nepokroeff, M., Pires, J. C., Zimmer, E. A., & Sytsma, K. J. (2003). Family-level relationships of Onagraceae based on chloroplast rbcL and ndh F data. *American Journal of Botany*, 90(1), 107–115. <https://doi.org/10.3732/ajb.90.1.107>
- Li, F.-W., Kuo, L.-Y., Rothfels, C. J., Ebihara, A., Chiou, W.-L., Windham, M. D., & Pryer, K. M. (2011). rbcL and matK earn two thumbs up as the Core DNA barcode for ferns. *PLoS One*, 6(10), e26597. <https://doi.org/10.1371/journal.pone.0026597>
- Li, S., Deng, Y., Wang, Z., Zhang, Z., Kong, X., Zhou, W., Yi, Y., & Qu, Y. (2020). Exploring the accuracy of amplicon-based internal transcribed spacer markers for a fungal community. *Molecular Ecology Resources*, 20(1), 170–184. <https://doi.org/10.1111/1755-0998.13097>
- Liu, Y., Yan, H.-F., Cao, T., & Ge, X.-J. (2010). Evaluation of 10 plant barcodes in Bryophyta (Mosses). *Journal of Systematics and Evolution*, 48(1), 36–46. <https://doi.org/10.1111/j.1759-6831.2009.00063.x>
- Meyer, J. M., Leempoel, K., Losapio, G., & Hadly, E. A. (2020). Molecular ecological network analyses: An effective conservation tool for the assessment of biodiversity, trophic interactions, and community structure. *Frontiers in Ecology and Evolution*, 8, 588430. <https://doi.org/10.3389/fevo.2020.588430>
- Moorhouse-Gann, R. J., Dunn, J. C., de Vere, N., Goder, M., Cole, N., Hipperson, H., & Symondson, W. O. C. (2018). New universal ITS2 primers for high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. *Scientific Reports*, 8, 8542. <https://doi.org/10.1038/s41598-018-26648-2>

- Moorhouse-Gann, R. J., Vaughan, I. P., Cole, N. C., Goder, M., Tatayah, V., Jones, C. G., Mike, D., Young, R. P., Bruford, M. W., Rivers, M. C., Hipperson, H., Russo, I.-R. M., Stanton, D. W. G., & Symondson, W. O. C. (2022). Impacts of herbivory by ecological replacements on an Island ecosystem. *Journal of Applied Ecology*, 59(9), 2245–2261. <https://doi.org/10.1111/1365-2664.14096>
- Nichols, R. V., Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., Green, R. E., & Shapiro, B. (2018). Minimizing polymerase biases in metabarcoding. *Molecular Ecology Resources*, 18(5), 927–939. <https://doi.org/10.1111/1755-0998.12895>
- Omelchenko, D. O., Speranskaya, A. S., Ayginin, A. A., Khafizov, K., Krinitsina, A. A., Fedotova, A., Pozdyshev, D., Shtratnikova, V. Y., Kupriyanova, E., Shipulin, G. A., & Logacheva, M. D. (2019). Improved protocols of ITS1-based metabarcoding and their application in the analysis of plant-containing products. *Genes*, 10(2), 122. <https://doi.org/10.3390/genes10020122>
- Ortega, A., Geraldí, N. R., Diaz-Rua, R., Orberg, S. B., Wesselmann, M., Krause-Jensen, D., & Duarte, C. M. (2020). A DNA mini-barcode for marine macrophytes. *Molecular Ecology Resources*, 20(4), 920–935. <https://doi.org/10.1111/1755-0998.13164>
- Palmieri, L., Bozza, E., & Giongo, L. (2009). Soft fruit traceability in food matrices using real-time PCR. *Nutrients*, 1(2), 316–328. <https://doi.org/10.3390/nu1020316>
- Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, 15(4), 819–830. <https://doi.org/10.1111/1755-0998.12355>
- Piñol, J., Senar, M. A., & Symondson, W. O. C. (2019). The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. *Molecular Ecology*, 28(2), 407–419. <https://doi.org/10.1111/mec.14776>
- Pornon, A., Escaravage, N., Burrus, M., Holota, H., Khimoun, A., Mariette, J., Pellizzari, C., Iribar, A., Etienne, R., Taberlet, P., Vidal, M., Winterton, P., Zinger, L., & Andalo, C. (2016). Using metabarcoding to reveal and quantify plant-pollinator interactions. *Scientific Reports*, 6, 27282. <https://doi.org/10.1038/srep27282>
- Prosser, S. W. J., & Hebert, P. D. N. (2017). Rapid identification of the botanical and entomological sources of honey using DNA metabarcoding. *Food Chemistry*, 214, 183–191. <https://doi.org/10.1016/j.foodchem.2016.07.077>
- Revéret, A., Rijal, D. P., Heintzman, P. D., Brown, A. G., Stoof-Leichsenring, K. R., & Alsos, I. G. (2023). Environmental DNA of aquatic macrophytes: The potential for reconstructing past and present vegetation and environments. *Freshwater Biology*, 68, 1929–1950. <https://doi.org/10.1111/fwb.14158>
- Richardson, R. T., Curtis, H. R., Matcham, E. G., Lin, C.-H., Suresh, S., Sponsler, D. B., Hearon, L. E., & Johnson, R. M. (2019). Quantitative multi-locus metabarcoding and waggle dance interpretation reveal honey bee spring foraging patterns in Midwest agroecosystems. *Molecular Ecology*, 28(3), 686–697. <https://doi.org/10.1111/mec.14975>
- Sipos, R., Székely, A. J., Palatinszky, M., Révész, S., Márialiget, K., & Nikolausz, M. (2007). Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targeting bacterial community analysis. *FEMS Microbiology Ecology*, 60(2), 341–350. <https://doi.org/10.1111/j.1574-6941.2007.00283.x>
- Sønstebo, J. H., Gielly, L., Brysting, A. K., Elven, R., Edwards, M., Haile, J., Willerslev, E., Coissac, E., Rioux, D., Sannier, J., Taberlet, P., & Brochmann, C. (2010). Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. *Molecular Ecology Resources*, 10(6), 1009–1018. <https://doi.org/10.1111/j.1755-0998.2010.02855.x>
- Sun, Y., Skinner, D. Z., Liang, G. H., & Hulbert, S. H. (1994). Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theoretical and Applied Genetics*, 89(1), 26–32. <https://doi.org/10.1007/bf00226978>
- Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA: For biodiversity research and monitoring* (First ed.). Oxford University Press.
- Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermet, T., Corthier, G., Brochmann, C., & Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, 35(3), e14. <https://doi.org/10.1093/nar/gkl938>
- Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, 17(5), 1105–1109. <https://doi.org/10.1007/BF00037152>
- Tanaka, K., Nozaki, A., Nakadai, H., Shiwa, Y., & Shimizu-Kadota, M. (2020). Using pollen DNA metabarcoding to profile nectar sources of urban beekeeping in Kt-ku, Tokyo. *BMC Research Notes*, 13(1), 515. <https://doi.org/10.1186/s13104-020-05361-2>
- Timpano, E., Scheible, M., & Meiklejohn, K. (2020). Optimization of the second internal transcribed spacer (ITS2) for characterizing land plants from soil. *PLoS One*, 15(4), e0231436. <https://doi.org/10.1371/journal.pone.0231436>
- Tournayre, O., Leuchtmann, M., Filippi-Codaccioni, O., Trillat, M., Piry, S., Pontier, D., Charbonnel, N., & Galan, M. (2020). In silico and empirical evaluation of twelve metabarcoding primer sets for insectivorous diet analyses. *Ecology and Evolution*, 10(13), 6310–6332. <https://doi.org/10.1002/ece3.6362>
- Tsukamoto, Y., Yonezawa, S., Katayama, N., & Isagi, Y. (2021). Detection of endangered aquatic plants in rapid streams using environmental DNA. *Frontiers in Ecology and Evolution*, 8, 622291. <https://doi.org/10.3389/fevo.2020.622291>
- Valentini, A., Miquel, C., Nawaz, M. A., Bellemain, E., Coissac, E., Pompanon, F., Gielly, L., Cruaud, C., Nascetti, G., Wincker, P., Swenson, J. E., & Taberlet, P. (2009). New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: The trnL approach. *Molecular Ecology Resources*, 9(1), 51–60. <https://doi.org/10.1111/j.1755-0998.2008.02352.x>
- Valentini, A., Pompanon, F., & Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in Ecology & Evolution*, 24(2), 110–117. <https://doi.org/10.1016/j.tree.2008.09.011>
- Varotto, C., Pindo, M., Bertoni, E., Casarotto, C., Camin, F., Girardi, M., Maggi, V., & Cristofori, A. (2021). A pilot study of eDNA metabarcoding to estimate plant biodiversity by an alpine glacier core (Adamello glacier, North Italy). *Scientific Reports*, 11, 1208. <https://doi.org/10.1038/s41598-020-79738-5>
- Vasconcelos, S., Nunes, G., Dias, M., Lorena, J., Oliveira, R., Lima, T., Pires, E., Valadares, R., Alves, R., Watanabe, M., Zappi, D., Hiura, A., Pastore, M., Vasconcelos, L., Mota, N., Viana, P., Gil, A., Simões, A., Imperatriz-Fonseca, V., ... Oliveira, G. (2021). *Unravelling the plant diversity of the Amazonian canga through DNA barcoding*. <https://doi.org/10.22541/au.161882857.78207100/v1>
- Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., Geiger, M. F., Grabowski, M., Rimet, F., Rulik, B., Strand, M., Szucsich, N., Weigand, A. M., Willassen, E., Wyler, S. A., Bouchez, A., Borja, A., Čiamporová-Zaťovičová, Z., Ferreira, S., ... Ekrem, T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Science of the Total Environment*, 678, 499–524. <https://doi.org/10.1016/j.scitotenv.2019.04.247>
- White, T., Bruns, T., Lee, S., Taylor, J., Innis, M., Gelfand, D., & Sninsky, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *Pcr Protocols: A Guide to Methods and Applications*, 31, 315–322.
- Wilkinson, M. J., Szabo, C., Ford, C. S., Yarom, Y., Croxford, A. E., Camp, A., & Gooding, P. (2017). Replacing Sanger with next generation

- sequencing to improve coverage and quality of reference DNA barcodes for plants. *Scientific Reports*, 7(1), 46040. <https://doi.org/10.1038/srep46040>
- Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergard, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L. S., Pearman, P. B., Cheddadi, R., Murray, D., Brathen, K. A., Yoccoz, N., ... Taberlet, P. (2014). Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature*, 506(7486), 47–51. <https://doi.org/10.1038/nature12921>
- Xin, T., Xu, Z., Jia, J., Leon, C., Hu, S., Lin, Y., Ragupathy, S., Song, J., & Newmaster, S. G. (2018). Biomonitoring for traditional herbal medicinal products using DNA metabarcoding and single molecule, real-time sequencing. *Acta Pharmaceutica Sinica B*, 8(3), 488–497. <https://doi.org/10.1016/j.apsb.2017.10.001>
- Yang, Y., Zhan, A., Cao, L., Meng, F., & Xu, W. (2016). Selection of a marker gene to construct a reference library for wetland plants, and the application of metabarcoding to analyze the diet of wintering herbivorous waterbirds. *PeerJ*, 4, e2345. <https://doi.org/10.7717/peerj.2345>
- Yoccoz, N. G., Bråthen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., Von Stedingk, H., Brysting, A. K., Coissac, E., Pompanon, F., Sønstebo, J. H., Miquel, C., Valentini, A., De Bello, F., Chave, J., Thuiller, W., Wincker, P., Cruaud, C., Gavory, F., ... Taberlet, P. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology*, 21(15), 3647–3655. <https://doi.org/10.1111/j.1365-294X.2012.05545.x>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Espinosa Prieto, A., Hardion, L., Debortoli, N., & Beisel, J.-N. (2024). Finding the perfect pairs: A matchmaking of plant markers and primers for multi-marker eDNA metabarcoding. *Molecular Ecology Resources*, 24, e13937. <https://doi.org/10.1111/1755-0998.13937>