




## Article

# Harnessing the Power of Metabarcoding in the Ecological Interpretation of Plant-Pollinator DNA Data: Strategies and Consequences of Filtering Approaches

Nicola Tommasi , Andrea Ferrari, Massimo Labra , Andrea Galimberti  and Paolo Biella \* 

ZooplantLab, Department of Biotechnology and Biosciences, University of Milano-Bicocca, 20126 Milano, Italy; nicola.tommasi@unimib.it (N.T.); a.ferrari79@campus.unimib.it (A.F.); massimo.labra@unimib.it (M.L.); andrea.galimberti@unimib.it (A.G.)

\* Correspondence: paolo.biella@unimib.it

**Abstract:** Although DNA metabarcoding of pollen mixtures has been increasingly used in the field of pollination biology, methodological and interpretation issues arise due to its high sensitivity. Filtering or maintaining false positives, contaminants, and rare taxa or molecular features could lead to different ecological results. Here, we reviewed how this choice has been addressed in 43 studies featuring pollen DNA metabarcoding, which highlighted a very high heterogeneity of filtering methods. We assessed how these strategies shaped pollen assemblage composition, species richness, and interaction networks. To do so, we compared four processing methods: unfiltering, filtering with a proportional 1% of sample reads, a fixed threshold of 100 reads, and the ROC approach (Receiver Operator Characteristic). The results indicated that filtering impacted species composition and reduced species richness, with ROC emerging as a conservative approach. Moreover, in contrast to unfiltered networks, filtering decreased network Connectance and Entropy, and it increased Modularity and Connectivity, indicating that using cut-off thresholds better describes interactions. Overall, unfiltering might compromise reliable ecological interpretations, unless a study targets rare species. We discuss the suitability of each filtering type, plead for justifying filtering strategies on biological or methodological bases and for developing shared approaches to make future studies more comparable.

**Keywords:** bioinformatics; cut-off thresholds; high throughput sequencing; rare taxa; false positives; molecular ecological network



**Citation:** Tommasi, N.; Ferrari, A.; Labra, M.; Galimberti, A.; Biella, P. Harnessing the Power of Metabarcoding in the Ecological Interpretation of Plant-Pollinator DNA Data: Strategies and Consequences of Filtering Approaches. *Diversity* **2021**, *13*, 437. <https://doi.org/10.3390/d13090437>

Academic Editors: Luc Legal and Oliver Gailing

Received: 15 June 2021

Accepted: 8 September 2021

Published: 10 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The study of Plant—Pollinator interactions is pivotal to address both theoretical and applicative issues, with important implications in evolutionary studies, conservation biology, and agrifood security, and it is relevant for providing reliable policies of land-use management and mitigation of anthropogenic stressors [1–4].

Traditionally, studies of Plant—Pollinator interactions have been carried out with direct field observations of animal foraging activity during flower visitation [5,6]. However, to unveil Plant—Pollinator interactions, it is a valuable approach to classify the pollen grains carried on the pollinator's body [7,8]. This pollen might be accidentally picked and carried by flower visitors when they touch plant reproductive structures. Alternatively, it can be actively collected and accumulated in specialized structures such as the scopa or the corbiculae in the case of some bee species. The identification of the transported pollen allows discovering of the foraging “history” of flower visitors prior to a sampling event. In this way, it is possible to retrieve complete behavioural and ecological information on flower resource exploitation and to address ecological research questions in a potentially fine detail. To achieve pollen identification, classical palynology based on morphology has traditionally been used. This approach could provide lot of information about sample

composition. However, it requires high expertise with pollen morphological assessment and, based on the operator's expertise, can be time-consuming [9,10]. In addition, a detailed taxonomic resolution through morphological criteria could be limited by the lack of diagnostic characters among taxa [11]. The morphological approach could be useful in cases of a low number of pollen samples to be identified or for gathering quantitative information on pollen amounts [12].

In the last decade, morphological difficulties have progressively been overcome by using DNA-based approaches that significantly reduce the time required for pollen identification [13,14]. Recent developments in DNA sequencing technologies, especially those based on High-Throughput Sequencing (HTS), made it possible to analyse the taxonomic composition of complex DNA matrices. For example, mixed pollen samples can be characterized by using standard DNA barcode regions in a so-called DNA metabarcoding approach [15,16]. For pollen-based studies, DNA metabarcoding is becoming a standard approach, being employed not only for the characterization of the pollen retrieved from animal bodies (see e.g., [17]), but also in the analysis of other kinds of samples such as cavity nests [18], honey [19], sediments [20,21], and forensics [10,22]. In the context of Plant—Pollinator interactions, the type of data retrieved from pollen DNA metabarcoding could potentially shed light on the foraging habits of flower visitors or evaluate the complexity and resilience of the interaction networks. This methodological revolution not only improved ecological knowledge, but it also offered new insights into the development of effective conservation and restoration actions [9].

Given the astounding number of sequences (hereafter “reads”) [23,24], the data from HTS techniques require a proper bioinformatic pipeline. This is a critical phase of the dry lab activities and usually consists in (i) the assembly of paired-end reads resulting from bidirectional sequencing of the DNA templates, (ii) the analysis of the variation among sequences and the clustering of molecular features (e.g., Operational Taxonomic Units OTUs sensu [25] or Exact Sequence Variants ESVs sensu [26]), and (iii) the removal of chimeras, artifacts, and spurious sequences [27]. However, this bioinformatics process does not completely solve all the potential biases. Additional artifacts, hereafter referred to as false positives, result from clusters of molecular features (i.e., OTUs and ESVs) generated as a consequence of inaccuracies during field sampling operations (e.g., cross-contamination among samples), laboratory processing (e.g., contamination of DNA extraction or amplification reagents), or from some steps of the bioinformatics analysis (e.g., taxonomic misidentification of molecular features) [24,28]. The presence of infrequently detected molecular features or taxa might add further background noise in the output of a DNA metabarcoding pipeline. Given the extreme sensitivity of DNA metabarcoding, it is crucial to filter out false positives and contaminants, which could significantly alter the reconstruction of samples composition. Moreover, rare features or taxa should be treated consciously during the postsequencing bioinformatics processing and possibly removed, depending on the study aims and the required sensitivity of the analysis [27–30]. However, the resulting species composition of a sample could be biased by the disapplication or misapplication of cut-off thresholds. For instance, DNA metabarcoding could detect the occurrence of particularly infrequent taxa, which may be of interest in some specific cases (e.g., tracking the origin of a sample based on rare pollen). On the other hand, these may produce a great impact on the ecological interpretation of results, especially when reads counts are converted into presence/absence data. Such a situation could lead to the overestimation of the generalist attitudes of the investigated pollinators and to misleading ecological interpretations.

The application of an appropriate cut-off threshold to filter the DNA metabarcoding data from the signal of possible false positives and rare taxa or features is therefore a critical step of the bioinformatics pipeline. Although some studies have not applied any cut-off threshold, different approaches for filtering false positives and rare taxa or features have been used so far in recent literature. In practice, some studies applied fixed cut-off thresholds, such as a defined number of reads used as reference level for accepting a

molecular feature or taxon in a sample (e.g., [31]). Other studies employed proportional cut-off thresholds, where molecular features or taxon are discarded if represented by less than a certain percentage of the total reads of a sample (e.g., [32]). Alternatively, statistical approaches have been used for estimating a variable threshold based on Receiver Operator Characteristic (ROC) curves, thus depending on the distribution of reads among molecular features or taxa within a sample [17]. This highlights the absence of agreement on whether and how to prune a DNA metabarcoding output. However, to date, no studies have investigated the effect of each of the abovementioned filtering strategies on molecular datasets concerning pollen samples (or honey) and Plant—Pollinator interactions.

In this study, we investigated and summarized the criteria and the strategies adopted for filtering out the false positives and rare features or taxa, focusing on published studies on pollen DNA metabarcoding. Moreover, we aimed at evaluating the direct ecological effects of the most commonly applied filtering methods on publicly available datasets of pollen/honey DNA metabarcoding. To do this, we measured how unfiltering or different cut-off thresholds impacted (i) plant species composition and species richness and (ii) the interactions among plants and pollinators described by network indexes. With these aims, we evaluated how the different filtering strategies could alter the identification of species and of interactions, and thus the ecological interpretation of the results.

## 2. Methods

### 2.1. Filtering Taxa from Pollen DNA Metabarcoding: Literature Overview

To revise the types of filtering and the methodology applied in the scientific literature used to remove (or not) false positives and rare taxa or features, bibliographical research was conducted in Scopus using the following keywords: “DNA” + “metabarcoding” + “pollen”. Within the results of the query, we selected only peer-reviewed original published articles that dealt with pollination, pollinator diet (pollen and honey), and plant—pollinator interactions by using a DNA metabarcoding approach (we excluded reviews, news, views, opinions, perspectives papers, and studies on airborne pollen or other pollen matrices when unrelated to pollinators). We selected studies spanning between 2012, when the term DNA metabarcoding was proposed for the first time [16], and 2021 (last update on 9 May 2021). The retrieved articles were used to create a summarizing table (Table 1) including: (i) the type of sample from which the DNA was extracted, (ii) the studied organism, (iii) the details of the filtering applied, and (iv) the DNA barcoding markers used to achieve the amplification reaction.

**Table 1.** List of published studies subjected to review, including details on referencing, used samples in the DNA metabarcoding analysis, the organisms from which the pollen samples were collected, and the cut-off threshold with a brief explanation of the filtering actually applied. Additional information is given on the DNA barcode marker(s) used and on whether the dataset was included in the present study.

| Source                      | Type of Sample             | Organism  | Type of Cut-Off Threshold   | Detail on the Application of the Cut-Off Threshold   | DNA Barcode Marker(s) | Nonfiltered Dataset Used in This Study |
|-----------------------------|----------------------------|---|-----------------------------|--|-----------------------|--|
| Baksay et al. (2020) [33]   | Mock pollen samples        | -   | Mixed                       | Sequences with a count of $\leq 10$ , with no variants and with a count $< 5\%$ of the total reads per sample  | ITS1, <i>trnL</i>     |  |
| Bänsch et al. (2020) [34]   | Pollen from legs           | <i>Apis mellifera</i> , <i>Bombus</i> spp. (Apidae)       | Not specified               | -  | ITS2                  |  |
| Bell et al. (2017b) [35]    | Mock pollen samples        | -   | Negative controls           | Removed identifications occurring at a frequency lower than those obtained in negative controls (isolation negative control = 34 reads, PCR negative control = 30 reads) | ITS2, <i>rbcL</i>     |  |
| Bell et al. (2017a) [12]    | Pollen from the whole body | Hymenoptera: Anthophila                                   | Negative controls           | Removed taxonomic classifications recorded from reads fewer than the maximum number of a negative control (21-936 <i>rbcL</i> and 42-1124 ITS2)                          | ITS2, <i>rbcL</i>     | X                                      |
| Bell et al. (2019) [24]     | Mock pollen samples        | -   | Negative controls           | Threshold based on the maximum sequence count from any negative control (11 and 34 ITS2, 8 and 30 <i>rbcL</i> )  | ITS2, <i>rbcL</i>     | X                                      |
| Beltramo et al. (2021) [36] | Honey                      | <i>Apis mellifera</i> (Apidae)                            | Proportional                | Removed OTUs with $< 0.2\%$ of the reads   | <i>trnL</i>           |  |
| Biella et al. (2019) [17]   | Pollen from legs           | <i>Bombus terrestris</i> (Apidae)                         | Variable: statistical-based | Receiver Operating Characteristics (ROC)   | ITS2                  | X                                      |
| Danner et al. (2017) [37]   | Pollen from legs           | <i>Apis mellifera</i> (Apidae)                            | Proportional                | Removed species $< 1\%$ of the relative reads count per sample   | ITS2                  |  |
| DeVere et al. (2017) [38]   | Honey                      | <i>Apis mellifera</i> (Apidae)                            | Not specified               | -  | <i>rbcL</i>           | X                                      |
| Elliott et al. (2021) [39]  | Pollen from legs or scopa  | Hymenoptera: Apidae, Halictidae, Megachilidae, Colletidae | Proportional                | Removed taxa $< 1\%$ of all reads per plant taxon for each bee species   | <i>rbcL</i>           |  |

Table 1. Cont.

| Source                           | Type of Sample             | Organism  | Type of Cut-Off Threshold | Detail on the Application of the Cut-Off Threshold                             | DNA Barcode Marker(s)   | Nonfiltered Dataset Used in This Study |
|----------------------------------|----------------------------|---|---------------------------|--|-------------------------|--|
| Fahimee et al. (2021) [40]       | Pollen from the whole body | <i>Heterotrigona itama</i> (Apidae)<br>Diptera,                                 | Fixed—Not proportional    | Removed OTUs with <2 reads   | trnL                    |  |
| Galliot et al. (2017) [41]       | Pollen from the whole body | Hymenoptera,<br>Coleoptera,<br>Lepidoptera                                      | Negative controls         | Threshold of 3 reads based on negative controls                                | ITS2                    |  |
| Gous et al. (2019) [42]          | Pollen from the scopa      | <i>Megachile venusta</i> (Megachilidae)   | Proportional              | Removed taxa <0.1% of total reads number per sample                            | ITS1, ITS2, <i>rbcL</i> |  |
| Gous et al. (2021) [43]          | Pollen from the scopa      | <i>Megachile</i> spp. (Megachilidae)  | Proportional              | Removed taxa <0.1% of total reads number per sample                            | ITS2                    |  |
| Hawkins et al. (2015) [44]       | Honey                      | <i>Apis mellifera</i> (Apidae)  | Fixed—Not proportional    | Removed taxa <10 sequences   | <i>rbcL</i>             |  |
| Jones et al. (2021) [45]         | Honey                      | <i>Apis mellifera</i> (Apidae)  | Fixed—Not proportional    | Singletons discarded   | ITS2, <i>rbcL</i>       | X                                      |
| Khansaritoreh et al. (2020) [46] | Honey                      | <i>Apis mellifera</i> (Apidae)  | Not specified             | -  | ITS2, <i>rbcL</i>       |  |
| Leidenfrost et al. (2020) [47]   | Pollen from legs           | <i>Bombus terrestris</i> (Apidae)   | Not specified             | -  | ITS2                    |  |
| Lucas et al. (2018a) [48]        | Pollen from the whole body | Syrphidae   | Not specified             | -  | <i>rbcL</i>             |  |
| Lucas et al. (2018b) [49]        | Pollen from the whole body | Syrphidae   | Not specified             | -  | <i>rbcL</i>             |  |
| Lucek et al. (2019) [50]         | Honey                      | <i>Apis mellifera</i> (Apidae)  | Fixed—Not proportional    | 5 reads per sequences cluster  | ITS2                    | X                                      |
| Macgregor et al. (2019) [51]     | Pollen from proboscid      | Lepidoptera (moths)   | Negative controls         | Threshold of 50 reads based on positive and negative controls                  | <i>rbcL</i>             |  |
| Nürnbergger et al. (2019) [52]   | Pollen from legs           | <i>Apis mellifera</i> (Apidae)  | Not specified             | -  | ITS2                    |  |
| Peel et al. (2019) [53]          | Pollen from legs           | <i>Apis mellifera</i> , <i>Bombus</i> spp. (Apidae)                             | Proportional              | Removed taxa <1% of the total assigned long reads per sample                   | Genomic DNA             |  |
| Piko et al. (2021) [54]          | Pollen from the whole body | <i>Bombus terrestris</i> ,<br><i>B.pascuorum</i> ,<br><i>B.lucorum</i> (Apidae) | Mixed                     | Removed taxa <100 reads each sample and <1% of the total read count per sample | ITS2                    |  |

Table 1. Cont.

| Source                         | Type of Sample                                 | Organism  | Type of Cut-Off Threshold | Detail on the Application of the Cut-Off Threshold  | DNA Barcode Marker(s)                         | Nonfiltered Dataset Used in This Study |
|--------------------------------|--|---|---------------------------|---|---|--|
| Pornon et al. (2016) [55]      | Mock pollen samples,<br>Pollen from whole body | <i>Hippeastrum</i> sp.,<br><i>Chrysanthemum</i> sp.,<br><i>Lilium</i> sp.; Diptera,<br>Hymenoptera,<br>Coleoptera,<br>Lepidoptera<br>Diptera, | Mixed                     | Removed taxa <0.1% of the most common sequences and <10 reads each sample   | ITS1, <i>trnL</i>                             |  |
| Pornon et al. (2017) [56]      | Pollen from the whole body                     | Hymenoptera,<br>Coleoptera,<br>Lepidoptera  | Fixed—Not proportional    | Sequences less than <1000   | ITS1, <i>trnL</i>                             |  |
| Pornon et al. (2019) [31]      | Pollen from the whole body                     | Syrphidae,<br>Empididae, Apidae   | Fixed—Not proportional    | Sequences less than <1000   | ITS1, <i>trnL</i>                             |  |
| Potter et al. (2019) [57]      | Pollen from the whole body                     | Hymenoptera:<br>Anthophila  | Not specified             | -   | <i>rbcL</i>                                   |  |
| Richardson et al. (2015a) [19] | Pollen from legs                               | <i>Apis mellifera</i><br>(Apidae)   | Not specified             | -   | ITS2  |  |
| Richardson et al. (2015b) [58] | Pollen from legs                               | <i>Apis mellifera</i><br>(Apidae)   | Not specified             | -   | ITS2, <i>rbcL</i> , <i>matK</i>               |  |
| Richardson et al. (2019) [59]  | Pollen from legs                               | <i>Apis mellifera</i><br>(Apidae)   | Proportional              | Removed genera identified with only one marker and taxa with proportion of sequences <0.01%   | ITS2, <i>rbcL</i> , <i>trnL</i> , <i>trnH</i> |  |
| Richardson et al. (2021) [60]  | Pollen from legs                               | <i>Apis mellifera</i><br>(Apidae)   | Proportional              | Removed genera identified with only one marker and with <0.001 proportional abundance of sequences  | ITS2, <i>rbcL</i> , <i>trnL</i>               |  |
| Sickel et al. (2015) [61]      | Pollen from nest                               | <i>Osmia bicornis</i> ,<br><i>O.truncorum</i><br>(Megachilidae)   | Proportional              | Removed taxa <0.1% of reads per sample  | ITS2  |  |
| Simanonok et al. (2021) [62]   | Pollen from legs                               | <i>Bombus affinis</i><br>(Apidae)   | Mixed                     | Removed OTU <10 reads and taxa with <2% reads per sample  | ITS2  |  |
| Smart et al. (2017) [63]       | Pollen from legs                               | <i>Apis mellifera</i><br>(Apidae)   | Fixed—Not proportional    | Removed taxa <50 reads  | ITS1, ITS2                                    |  |
| Suchan et al. (2019) [64]      | Pollen from the whole body                     | <i>Vanessa cardui</i><br>(Lepidoptera)  | Fixed—Not proportional    | Removed taxa <100 reads per sample  | ITS2  |  |
| Swenson et al. (2021) [65]     | Mock pollen samples                            | -   | Mixed                     | Removed taxa <0.1% of the sample reads of ITS1 and ITS2; removed taxa occurring at a lower frequency than those obtained from negative controls | ITS1, ITS2, <i>rbcL</i>                       |  |

Table 1. Cont.

| Source                      | Type of Sample             | Organism                                    | Type of Cut-Off Threshold | Detail on the Application of the Cut-Off Threshold   | DNA Barcode Marker(s) | Nonfiltered Dataset Used in This Study |
|-----------------------------|----------------------------|---|---------------------------|--|-----------------------|--|
| Tanaka et al. (2020) [66]   | Pollen from honeycomb      | <i>Apis mellifera</i> (Apidae)              | Not specified             | -  | <i>rbcL</i>           |  |
| Tommasi et al. (2021) [67]  | Pollen from the whole body | Hymenoptera: Anthophila, Diptera: Syrphidae | Not specified             | -  | ITS2                  | X                                      |
| Tremblay et al. (2019) [68] | Pollen from legs           | <i>Apis mellifera</i> (Apidae)              | Fixed—Not proportional    | Removed taxa <100 reads  | ITS2                  |  |
| Vaudo et al. (2020) [69]    | Pollen from nest           | <i>Osmia cornifrons</i> (Megachilidae)      | Proportional              | Removed taxa <1% sample read abundance and genera <0.3% of the total read counts per site across all sites | ITS2                  | X                                      |
| Wilson et al. (2021) [32]   | Pollen from nest           | <i>Tetragonula carboniaria</i> (Apidae)     | Proportional              | Removed taxa identified in blank controls with abundance <1% of the relative read abundance in real sample | ITS2, <i>rbcL</i>     |  |

## 2.2. Evaluating the Consequences of Filtering (or Not) Taxa

To evaluate how the application (or not) of different cut-off thresholds could lead to changes in the results and their interpretation, we retrieved publicly available DNA metabarcoding datasets based on the ITS2 DNA barcode marker (that is the most used in pollen DNA metabarcoding studies) from the previously mentioned literature search. Only those datasets which were not preliminarily filtered were kept for our analysis (see Table 1). In detail, we retrieved published nonfiltered datasets (hereafter named as “no cut”, equivalent to a 0-reads threshold), and we derived several subsequent filtered versions by separately applying three different approaches for filtering false positives and rare taxa or features. The filtering approaches were chosen based on utilization frequency in the literature or on their biological reliability (i.e., the ROC approach). Specifically, the first method is based on a fixed threshold, and it removes from a sample the molecular features or taxa represented by less than 100 reads (hereafter “fixed 100 reads”), thus mimicking studies where exclusion thresholds are based on reads found in sequencing blanks (e.g., [12]). The second method is proportional and discards what is represented in a sample by a number of reads lower than 1% of the total sample count of reads (hereafter “proportional 1%”) as used for example in [37]. The third one estimates a cut-off threshold accounting for the distribution of reads among molecular features, thus providing a customized proportion for each sample through the statistical ROC curve approach, as indicated in [17] (hereafter “statistical ROC”). This strategy is commonly applied in several disciplines, and it was specifically proposed for the detection of false positives [70]. We applied the ROC approach in the same way as it was done in the pollen-based literature, thus following the procedure of [17]; see Supplementary Material Text S1 for a script, although different implementations of ROC are possible and they might affect the final estimations. We associated a variable coded as “negative” or “positive” to each taxon of a sample. Specifically, “negative” was assigned if its reads were 0; otherwise, “positive”. We fitted a Generalized Linear Regression with an overdispersed Poisson distribution (quasi-Poisson) for each sample to model the distribution of the amount of reads per taxa (quantitative response) between “positives”/“negatives” (categorical predictor). Fitting a regression is a necessary step for later estimating the false positives of a sample that otherwise are usually not known in DNA metabarcoding data. The predicted reads distribution was processed with the *pROC* package [71] (in the R environment) that uses the *roc* function to build ROC curves between the reads per sample estimated by the GLM and true “positives”/“negatives” (those used to fit the regression). The optimal threshold of reads below which taxa should be excluded was obtained with the function *coord* in the same package based on the Youden’s J statistic [72] (see [71] for further details).

For each dataset, changes in plant species composition and species richness (that was standardized for the maximum number of species observed in a sample) for each pollen sample was evaluated in response to the type of filtering used (i.e., no cut, fixed 100 reads, proportional 1%, and statistical ROC). For the comparison of pollen species composition, we used a Permutational Manova based on distance matrices (with Jaccard distance index), which is an analysis of variance that uses a permutations test with pseudo-F ratio [73]. This analysis was performed through the *adonis* function with R-package *vegan* [74], where each dataset was analysed independently. The effect of the different cut-off thresholds on species richness was evaluated through a Generalized Linear Mixed Model (GLMM) with species richness as response variable and the type of filtering used (i.e., no cut, fixed 100 reads, proportional 1%, and statistical ROC) as covariate. The identity of the pollinator animal nested within the dataset was set as a random effect.

Network indices describing the interactions between plants and pollinators were calculated. Specifically, the analysed indices were Connectance (i.e., proportion of possible links actually recorded), Modularity (i.e., a measure of how interactions are distributed into modules, where species within modules mostly interact with each other), and Shannon Entropy (i.e., a measure of the overall diversity and complexity in the interactions of a network). Furthermore, at the level of a single individual pollinator, the Connectivity



index was calculated. This index of centrality quantifies the putative central role of an individual or of a species in connecting different parts of the whole network [75]. It could provide information in ranking individuals or species according to their contribution to the stability of the interactions and the cohesion among network participants. Network indices were calculated through the R-package *bipartite* (specifically for Connectance, Modularity, and Entropy) and *rnetcarto* (for Connectivity) [76,77]. For this purpose, only those datasets originated from direct characterization of pollinator foraging were used, excluding a study on mock samples [24] and a study with an incomparable experimental design [38] that instead were used in the other analyses. Changes in interaction indices at the network level (Connectance, Modularity, Entropy) were evaluated through either a Linear or a Generalized Mixed Model depending on the distribution and range of the response variable. The type of filtering used was included in the models as covariate, and the dataset identity, as random effect. The individual level (i.e., samples) Connectivity was analysed as response variable, the type of filtering used, as covariate in interaction with the normalized degree of the pollinator individuals. This normalized degree was calculated as the number of plant species found in each sample divided by the overall number of plants. The inclusion of the normalized degree in this analysis allowed us to describe the variation of Connectivity across the entire specialism–generalism spectrum of an individual and in relation to the applied filtering approach. In this case, the sample identity nested within the dataset was included in the model as a random effect.

For all the mentioned regression analyses, a comparison among the adopted strategies of filtering was performed through a post hoc test (Tukey's HSD test). All the statistical analyses explained above were carried out with R (Version 3.6.1 of R).

### 3. Results

#### 3.1. Filtering Taxa from Pollen DNA Metabarcoding: Literature Overview

Overall, 43 research articles on pollen DNA metabarcoding were found and reviewed concerning the strategy of filtering of false positive and rare taxa or features (Table 1). About one quarter of studies did not apply any filtering approach, while the remaining ones applied at least a filtering type. Specifically, the proportional cut-off threshold was the most applied method (11 studies, 28%). In these studies, the cut-off threshold calculated as 1% of the number of reads produced by each sample was the most recurrent. Other filtering types were less common. Only one study used a statistical approach (i.e., the ROC curve; [16]) to set a proportional cut-off threshold. Nine studies (21%) used a fixed number of reads chosen arbitrarily as cut-off threshold (e.g., 100 or 1000 reads), while five other studies (12%) used the number of reads produced by negative controls to set the threshold to remove false positives and rare taxa or features. Finally, five studies (12%) used more than one filtering approach simultaneously. Details and a brief explanation of the strategies applied to set the cut-off threshold for the reviewed studies are reported in Table 1.

Twenty-eight published studies (65%) recovered the pollen samples from the whole insect's body or from specific body parts such as scopa and corbiculae. Four studies (9%) focused on pollen stored in cavity nests or in hives, while five (12%) investigated mixed pollen mock samples to address methodological issues (e.g., the optimization of DNA extraction or quantitative use of DNA metabarcoding reads). Finally, six studies (14%) analysed the taxonomic composition of honey by looking at the pollen grains contained in it.

Most of these studies (65%) used the ITS2 marker as a DNA barcode region for species identification, although in some these cases (28%), this marker was combined with other barcode loci (e.g., *rbcL*).

#### 3.2. Evaluating the Consequences of Filtering (or Not) Taxa

From the 43 reviewed studies, eight nonfiltered and publicly available ITS2 DNA metabarcoding datasets were retrieved. Among these, four were obtained by processing the pollen found in nests or carried on insects' bodies [17,35,67,69]. Three datasets contained

data from honey samples [38,45,50], and one was obtained from the analysis of pollen mock samples [24].

Significant changes in the composition of pollen samples depending on the filtering approach are summarized in Table 2. Specifically, the main differences occurred between the no-cut and all the other filtering approaches, in all datasets (Table 2). Minor changes in community composition among fixed 100 reads, proportional 1%, and statistical ROC were only occasionally found (Table 2).

**Table 2.** Comparison of cut-off thresholds applied on pollen species composition of samples from several datasets, based on Permutational Manova. Dataset names (entitled with main author and year; see Table 1 for further details) are reported in the first column “Dataset”. The column “F-value” reports the pseudo-F ratio value and the associated significance  $p$  ( $\alpha = 0.05$ ). Significant cases are reported in bold.

| Dataset                    | F-Value | Significance $p$ of Full Model | Significance $p$ of Pairwise Comparisons |                            |                                     |                            |                        |                                     |
|----------------------------|---------|--------------------------------|--|----------------------------|-------------------------------------|----------------------------|------------------------|-------------------------------------|
|                            |         |                                | Proportional 1% vs. No Cut               | Fixed 100 Reads vs. No Cut | Proportional 1% vs. Fixed 100 Reads | Statistical ROC vs. No Cut | Statistical ROC vs. 1% | Statistical ROC vs. Fixed 100 Reads |
| Tommasi et al. (2021) [67] | 0.819   | 0.806                          | 1  | 1                          | 1                                   | <b>0.031</b>               | 0.314                  | <b>0.045</b>                        |
| Bell et al. (2017a) [12]   | 87.264  | <b>0.001</b>                   | <b>0.001</b>                             | <b>0.001</b>               | 0.14                                | <b>0.001</b>               | <b>0.001</b>           | <b>0.001</b>                        |
| Bell et al. (2019) [24]    | 39.817  | <b>0.001</b>                   | <b>0.001</b>                             | <b>0.001</b>               | <b>0.01</b>                         | <b>0.001</b>               | 0.658                  | <b>0.001</b>                        |
| Biella et al. (2019) [17]  | 29.671  | <b>0.001</b>                   | <b>0.001</b>                             | <b>0.001</b>               | 0.035                               | <b>0.001</b>               | 0.725                  | <b>0.008</b>                        |
| Jones et al. (2021) [45]   | 6.538   | <b>0.001</b>                   | <b>0.001</b>                             | <b>0.001</b>               | 0.944                               | <b>0.001</b>               | 0.233                  | 0.855                               |
| Lucek et al. (2019) [50]   | 5.465   | <b>0.001</b>                   | <b>0.001</b>                             | <b>0.001</b>               | <b>0.038</b>                        | <b>0.001</b>               | 0.975                  | <b>0.058</b>                        |
| DeVere et al. (2017) [38]  | 2.415   | <b>0.024</b>                   | <b>0.004</b>                             | 0.212                      | 0.538                               | <b>0.003</b>               | 0.704                  | 0.152                               |
| Vaudo et al. (2020) [69]   | 11.553  | <b>0.001</b>                   | <b>0.001</b>                             | <b>0.001</b>               | 0.556                               | <b>0.001</b>               | 1                      | 0.578                               |

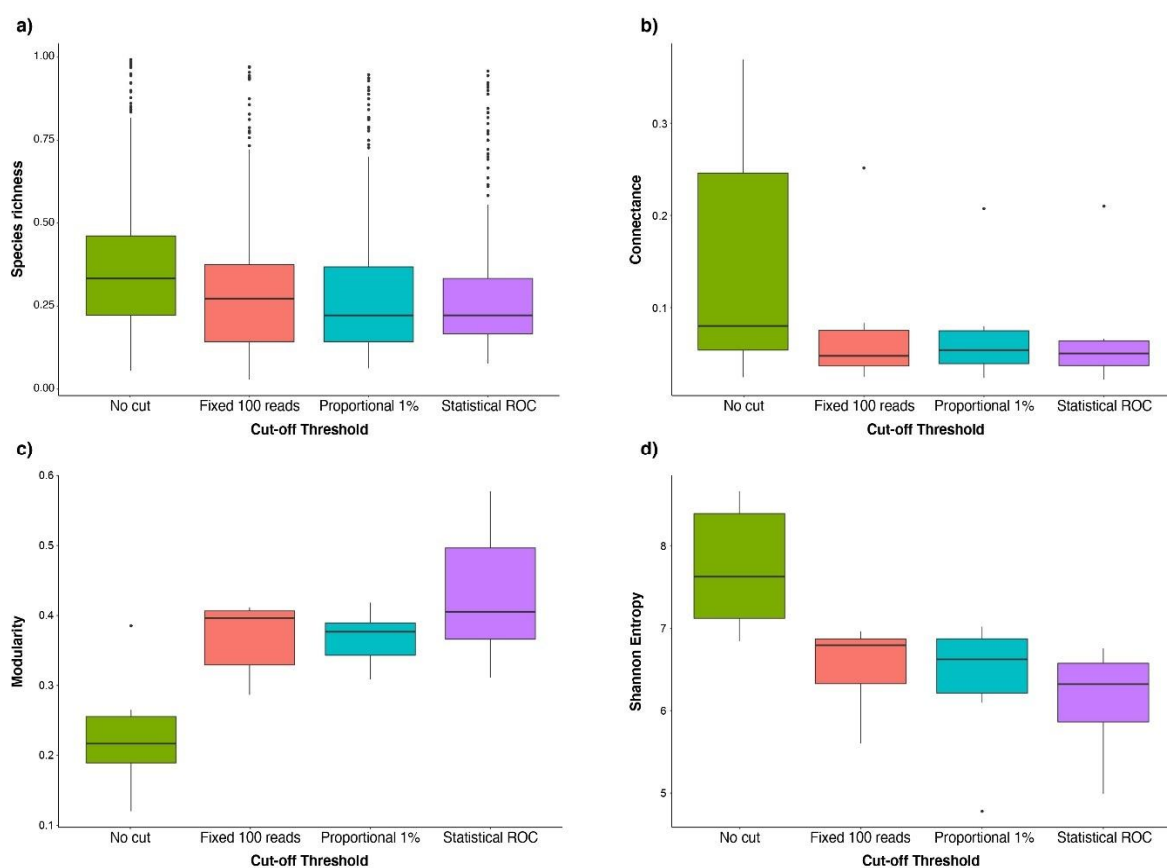
Plant species richness inferred from pollen samples was significantly influenced by the filtering approach ( $X_3^2 = 468.22$ ,  $p < 0.001$ ). Specifically, higher species richness per sample was found in the unfiltered type (i.e., no cut) compared to all the other filtering approaches. A significant difference between the proportional 1% and the statistical ROC approaches was also found, with the latter reducing species richness even more (Figure 1a, Table 3).

**Table 3.** Statistical comparison of the selected cut-off thresholds (i.e., no cut, fixed 100 reads, proportional 1%, statistical ROC) on plant species richness, Connectance, Modularity, and Entropy of interaction networks after Tukey’s pairwise comparison test ( $\alpha = 0.05$ ). Significant differences are highlighted in bold.

| Variable         | Comparison                      | Estimated Difference | Significance $p$ |
|------------------|---------------------------------|----------------------|------------------|
| Species richness | No cut—Fixed 100 reads          | 0.383                | <b>&lt;0.001</b> |
|                  | Proportional 1%—Fixed 100 reads | 0.043                | 0.178            |
|                  | Statistical ROC—Fixed 100 reads | −0.013               | 0.934            |
|                  | Proportional 1%—No cut          | −0.340               | <b>&lt;0.001</b> |
|                  | Statistical ROC—No cut          | −0.396               | <b>&lt;0.001</b> |
|                  | Statistical ROC—Proportional 1% | −0.055               | <b>0.040</b>     |
| Connectance      | No cut—Fixed 100 reads          | 0.660                | <b>0.008</b>     |
|                  | Proportional 1%—Fixed 100 reads | −0.068               | 0.991            |
|                  | Statistical ROC—Fixed 100 reads | −0.131               | 0.941            |
|                  | Proportional 1%—No cut          | −0.729               | <b>0.004</b>     |
|                  | Statistical ROC—No cut          | −0.792               | <b>0.001</b>     |
|                  | Statistical ROC—Proportional 1% | −0.063               | 0.993            |

Table 3. Cont.

| Variable   | Comparison                      | Estimated Difference | Significance <i>p</i> |
|------------|---------------------------------|----------------------|-----------------------|
| Modularity | No cut—Fixed 100 reads          | −0.678               | <0.001                |
|            | Proportional 1%—Fixed 100 reads | 0.000                | 1                     |
|            | Statistical ROC—Fixed 100 reads | 0.259                | 0.176                 |
|            | Proportional 1%—No cut          | 0.679                | <0.001                |
|            | Statistical ROC—No cut          | 0.937                | <0.001                |
| Entropy    | Statistical ROC—Proportional 1% | 0.259                | 0.177                 |
|            | No cut—Fixed 100 reads          | 1.189                | <0.001                |
|            | Proportional 1%—Fixed 100 reads | −0.191               | 0.819                 |
|            | Statistical ROC—Fixed 100 reads | −0.411               | 0.237                 |
|            | Proportional 1%—No cut          | −1.380               | <0.001                |
|            | Statistical ROC—No cut          | −1.600               | <0.001                |
|            | Statistical ROC—Proportional 1% | −0.220               | 0.746                 |

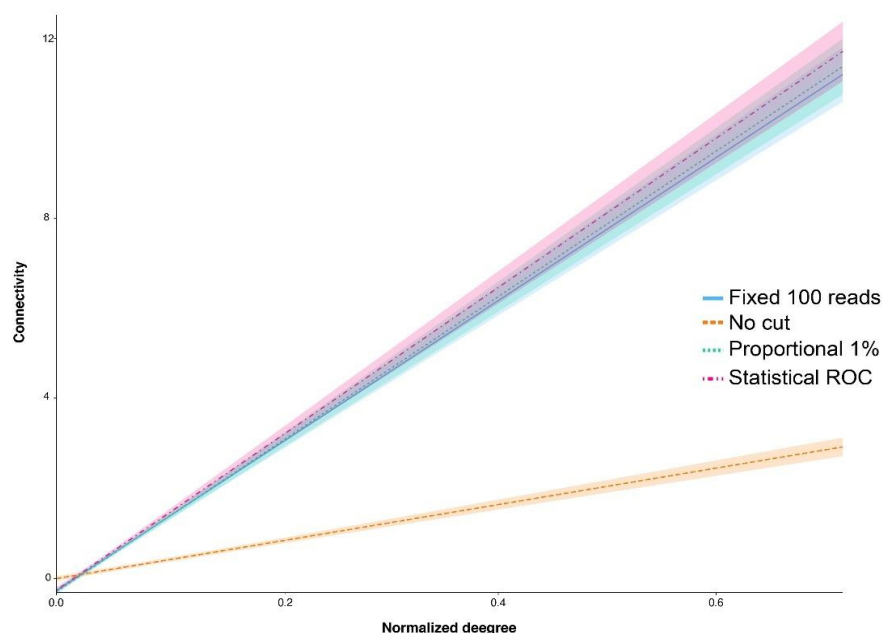


**Figure 1.** Changes in plant species richness (a), Connectance (b), Modularity (c), and Entropy (d) of plant—pollinator networks after using different cut-off thresholds (i.e., no cut, fixed 100 reads, proportional 1%, statistical ROC; see methods for details). Significances are reported in Table 3.

Significant effects of filtering on the network level indices were found on Connectance ( $X_3^2 = 11.642$ ,  $p = 0.008$ ), Modularity ( $X_3^2 = 25.273$ ,  $p < 0.001$ ), and Entropy ( $X_3^2 = 29.907$ ,  $p < 0.001$ ). Connectance (Figure 1b, Table 3) and Entropy indices (Figure 1d, Table 3) were significantly higher, while Modularity significantly lower (Figure 1c, Table 3) in the unfiltered (i.e., no cut) compared to all the other filtering approaches. In addition, in most cases, the ROC filtering changed the network indices even more than the “proportional 1%” and “fixed 100 reads” approaches (Figure 1c,d).

The individual level index of Connectivity showed a significant effect of the interaction between the filtering approach and the normalized degree index ( $X_3^2 = 609.2$ ,  $p < 0.001$ ).

Specifically, the Connectivity was lower in the unfiltered (no cut) compared to all the other filtering types for any value of the normalized degree (i.e., for both generalist and specialist individual pollinators; Figure 2, Table 4).



**Figure 2.** Response of the individual-level index of Connectivity to normalized degree in statistical interaction with the type of filtering (i.e., no cut, fixed 100 reads, proportional 1%, statistical-based ROC).

**Table 4.** Comparison of Connectivity index after different filtering types in statistical interaction with normalized degree of individual pollinator (Tukey’s multiple comparison test,  $\alpha = 0.05$ ). Significant cases are highlighted in bold.

| Filtering Type x Normalized Degree | Estimated Difference | Significance <i>p</i> |
|------------------------------------|----------------------|-----------------------|
| No cut—Fixed 100 reads             | −11.896              | <b>&lt;0.001</b>      |
| Proportional 1%—Fixed 100 reads    | 0.279                | 0.756                 |
| Statistical ROC—Fixed 100 reads    | 0.694                | 0.456                 |
| Proportional 1%—No cut             | 12.175               | <b>&lt;0.001</b>      |
| Statistical ROC—No cut             | 12.590               | <b>&lt;0.001</b>      |
| Statistical ROC—Proportional 1%    | −0.414               | 0.662                 |

#### 4. Discussion

Since its “formalization” in 2012, the DNA metabarcoding approach has revolutionized the field of biodiversity investigation, and it has even provided insights for studying biological interactions. Its application has rapidly spread and has contributed to research contexts such as microbiome [78,79], food [80–82], trophic ecology [83,84], and environmental DNA-based analyses [85]. In spite of its usefulness, methodological choices during the whole DNA metabarcoding pipeline and specifically the bioinformatics processing could deeply influence the obtained results and their interpretation [86–88]. Therefore, in this study, we attempted to evaluate the effects of the approach used to filter DNA metabarcoding outputs. Specifically, we focused on the analysis of DNA metabarcoding of pollen in the framework of plant—pollinator interactions, being aware that the outputs of our investigation could be extended to the other typologies of DNA metabarcoding-based studies. Although the issue of removing false positives and rare taxa or features is quite neglected in the scientific literature in relation to the bioinformatic pipeline (but see [28,30]), the choices made when analysing a HTS output could generate relevant effects on the

community composition, species richness, and species interactions. These aspects would deeply impact the ecological outcomes of the investigated system.

The high potential of using DNA metabarcoding outputs for pollen analysis would also merit a robust bioinformatic pipeline that should be coherent and comparable among different studies. Instead, our literature overview highlighted a high heterogeneity of filtering approaches adopted to remove false positives and infrequent taxa. This is particularly appreciable even among studies that focused on similar analytical matrices (e.g., pollen from animal bodies). Surprisingly, a similarly high heterogeneity in filtering approaches emerges from studies using morphological palynology ([7] used a minimum number of 10 pollen grains, while [89,90] used a threshold of 5, and [91,92] retain species with a frequency of pollen grains above 10%, and [93], above 1%). In the case of DNA metabarcoding of pollen, our literature review showed that the proportional approach is the most recurrent, that is, to remove those molecular features or species present in the sample with reads under a certain proportion of the total reads per sample. This is quite expected, as it is an approach also well represented in other studies using DNA metabarcoding (e.g., [83,94,95]). Probably due to the ease of calculating proportions, they bear advantages when comparing different samples or when samples would be too depauperated after a fixed raw number of reads. However, we found no concordance between different authors about the exact amount of proportion to be used as threshold, and surprisingly about the reasons justifying the choice of a particular percentage or another one (e.g., [61] used 0.1%, while [59] used 0.01%; see Table 1). It should be noticed that Peel and colleagues [53], while analysing samples of pollen prepared ad hoc with a known composition, highlighted that false positives occurred at a rate lower than 1% per sample, thus supporting this filtering strategy. On the other hand, caution should be recommended prior to generalizing the 1% threshold as a universally effective filtering practice; for samples represented by extremely high total reads count, it might be better to use a lower value. Conversely, 1% threshold can also be ineffective with almost empty samples, such as in the case of a fly that has never visited a flower but that was contaminated by airborne pollen. In such cases, it might be worth using even higher threshold values to better safeguard from misleading information.

The second most recurrent cut-off approach found in literature is based on a fixed number of read counts, used as general threshold across all samples (e.g., 50 as in [63], 100 as in [68]; see Table 1), the most frequent amount being 100 reads per sample. As reported above, with this approach, the specific value of the cut-off threshold is poorly supported by clear biological reasons. The subjectivity of authors is an important factor, and it could be a source of biases, as, for instance, studies using high threshold values would likely remove a high proportion of truly occurring taxa. For example, [56,57] observed how a threshold of 1000 reads per plant species ensures the removal of the vast majority of grass pollen species, which, however, were taxa occurring at the study areas and shall be considered true positives from airborne pollen. Therefore, low or high cutting values could have been chosen depending on the need to remove false positives but also on potential environmental contamination or infrequent species. Conversely, in other studies, the fixed cut-off value is clearly derived from sequenced negative controls. In those cases, the maximum number of reads found in blank samples is usually set as threshold (see Table 1). The assumption behind this approach is that it would allow removal of false positives exclusively originating from laboratory activities (i.e., during DNA isolation, PCR, and sequencing) [24]. However, the impact of using blanks to yield thresholds is not so clear when it comes to rare or infrequent species that might have fewer sequencing reads than controls, and those cases would be systematically removed by this approach. Regarding this, the development of practices to retrieve negative controls for field contaminations (as hypothesized in [24]) could probably further improve the potential of this filtering method, allowing for better discrimination between species originated from field contamination and rare but truly occurring taxa.

Unfiltering seems controversial. The literature survey (Table 1) showed that nearly a quarter of studies did not report a filtering approach (based on reads count) and possibly

did not filter the datasets with quantitative thresholds. However, in some of these cases, a manual filtering was used to remove the species that were not plausible in the study area [48,49,57,67], while in others, the concordance between multiple DNA barcoding markers is employed [58], thus at least partly following the recommendations to remove false positives and possibly rare taxa or features [96]. The analyses of our study clearly suggest that using a cut-off threshold for filtering the HTS output leads to significant differences compared to the unfiltered output matrix, especially in species composition, species richness and plant–pollinator interactions, impacting the ecological interpretation of the data. Our results indicated that, firstly, any of the cut-off thresholds yielded a community composition different from those obtained through unfiltered data. Moreover, filtering decreased species richness in comparison to nonfiltered data. These differences between unfiltered and filtered data could even be amplified under particular research aims. For example, in studies focusing on pollinator foraging behaviours, the unfiltering could overestimate the number of plants foraged by an animal, and it could obviously lead to an overestimation of generalism, foraging niche, and delivered ecosystem service of pollination. Another example derives from studies on honey composition, where a no-cut strategy could mislead on the purity of products, with consequences that could involve commercial issues. In our investigation, the filtering of false positives or rare taxa impacted not only the species composition and richness but also the ecological networks of species interactions. Specifically, we detected significant differences when comparing networks calculated from filtered to nonfiltered data. The implications of this network variation could potentially be very high, as, for instance, network Entropy, Connectance, Modularity, and Connectivity refer to network stability and resilience, to the ability to buffer perturbations, and to the stabilizing role of central hub species [40,97–99]. Thus, the higher the difference between filtering or unfiltering strategies, the higher the potential for misleading ecological results obtained from the networks associated to each filtering type. For instance, our results showed that filtering decreased network Connectance and Entropy. This aligns well with the lower species richness per sample found in filtered datasets, and it can be explained by a decrease in network number of realized links (i.e., fewer plant species found on pollinator bodies or samples). In other words, by decreasing the numerosity of links, filtering likely yields networks with slightly higher element-specific linkage compared to nonfiltered networks. From an ecological point of view, this translates in a lower chance of overestimating generalism after filtering. Moreover, filtering increases Modularity and Connectivity of networks. This result further clarifies that removing ambiguous taxa decreases the ubiquity of links among elements, thus allowing for better emergence of ordered patterns of well-defined compartments of interactions (Modularity) and important hub species connecting them (characterized by a higher Connectivity). As a consequence, filters seem to increase the ecological reliability when describing how flower resources are used by foragers (i.e., Modularity) and the importance of certain species in contributing to interactions stability (i.e., Connectivity). In other words, unfiltering returns networks richer in links, which tend to be ubiquitously distributed among elements, with the high potential of overestimating foraging strategies and network resilience. Based on these considerations, researchers could prefer filtering their data. One exception to this would be when the role of rare plant species is targeted in the study [100]. In this case, manually checking an unfiltered dataset for unplausible taxa could limit the amount of false negatives. Moreover, integrating DNA metabarcoding data with traditional (quantitative) morphology of pollen could improve results reliability and interpretation [101]. This could possibly control for the presence of spurious information from DNA metabarcoding.

Among the filtering strategies analysed here, the statistical ROC approach appears to be the most conservative one, since it tends to yield the lowest species richness, the highest Modularity, and the lowest Connectance and Entropy. Thus, ROC-based filtering might remove not only the false positives from samples but also the infrequent species. It should be noted that ecological patterns emerging or confirmed even in a conservative framework are more likely to be trustworthy. Even if this approach has rarely been applied,

to date, in the pollen DNA metabarcoding literature (Table 1), it was specifically developed to distinguish “true signals” from “noise” [102] in molecular biology and could constitute a promising avenue for processing data of pollen DNA metabarcoding. For instance, it has been used in other DNA-based research fields [103,104], such as for eDNA where it is proved to increase the reliability of data [105]. However, it would be promising to investigate how different parametrizations of the ROC approach would impact the estimations of cutting thresholds. Because ROC is a conservative approach, it may be favoured in studies willing to highlight ecologically meaningful species composition, richness, and interactions, while sacrificing the pursuit of high species richness based on keeping elements of rarity, potential contaminants, and false positives.

## 5. Conclusions

Our survey shed light on the possible consequences of using (un)filtering strategies of pollen DNA metabarcoding data in ecological and biological research. To date, this powerful molecular tool still requires the development of shared approaches on the bioinformatic filtering of molecular features. This would improve the provision of reliable, repeatable, and comparable data. In particular, we recommend that researchers (i) always make both raw unfiltered and filtered data easily accessible, thus improving the possibility of exploring large amounts of data and, consequently, the growing rate of human knowledge in strategic research fields such as pollination ecology. The authors should (ii) apply filtering from false positives and possibly also from infrequent species, depending on research aims. Moreover, (iii) the specific type of filtering must be clearly justified under a biological perspective, evaluating the efficiency and universality of the loci selected for species identification and the consequent taxonomic resolution of molecular feature assignments. Furthermore, (iv) the specific strategy has to be decided based on whether the research aim would benefit from a conservative filtering. Without an appropriate filtering, DNA metabarcoding reads converted to presence/absence data certainly yield spurious results [30,106,107]. To avoid this, conservative approaches like the ROC filtering must be preferably adopted. Otherwise, the application of a filter either based on a percentage with a clear biological support or based on a fixed value from negative controls is possible, although greater awareness should be placed on the risk of excluding only false positives while keeping environmental contamination and infrequent species.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/d13090437/s1>, Text S1: Script for estimating filtering thresholds with an approach based on ROC curves.

**Author Contributions:** Conceptualization, P.B., A.G. and N.T.; methodology, P.B. and N.T.; formal analysis, P.B. and N.T.; investigation, A.F.; data curation, N.T. and A.F.; writing—original draft preparation, N.T., P.B., A.G., M.L. and A.F.; supervision, A.G., P.B. and M.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** All the relevant data are within the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mitchell, R.J.; Irwin, R.E.; Flanagan, R.J.; Karron, J.D. Ecology and evolution of plant–pollinator interactions. *Ann. Bot.* **2009**, *103*, 1355–1363. [[CrossRef](#)]
2. Schweiger, O.; Biesmeijer, J.C.; Bommarco, R.; Hickler, T.; Hulme, P.E.; Klotz, S.; Kühn, I.; Moora, M.; Nielsen, A.; Ohlemüller, R.; et al. Multiple stressors on biotic interactions: How climate change and alien species interact to affect pollination. *Biol. Rev.* **2010**, *85*, 777–795. [[CrossRef](#)]
3. Burkle, L.A.; Alarcón, R. The future of plant–pollinator diversity: Understanding interaction networks across time, space, and global change. *Am. J. Bot.* **2011**, *98*, 528–538. [[CrossRef](#)]

4. Burkle, L.A.; Delphia, C.M.; O'Neill, K.M. A dual role for farmlands: Food security and pollinator conservation. *J. Ecol.* **2017**, *105*, 890–899. [[CrossRef](#)]
5. CaraDonna, P.J.; Waser, N.M. Temporal flexibility in the structure of plant–pollinator interaction networks. *Oikos* **2020**, *129*, 1369–1380. [[CrossRef](#)]
6. de Manincor, N.; Hautekèete, N.; Mazoyer, C.; Moreau, P.; Piquot, Y.; Schatz, B.; Schmitt, E.; Zélazny, M.; Massol, F. How biased is our perception of plant–pollinator networks? A comparison of visit- and pollen-based representations of the same networks. *Acta Oecologica* **2020**, *105*, 103551. [[CrossRef](#)]
7. Bosch, J.; Martín González, A.M.; Rodrigo, A.; Navarro, D. Plant–pollinator networks: Adding the pollinator’s perspective. *Ecol. Lett.* **2009**, *12*, 409–419. [[CrossRef](#)]
8. Cullen, N.; Xia, J.; Wei, N.; Kaczorowski, R.; Arceo-Gómez, G.; O'Neill, E.; Hayes, R.; Ashman, T.-L. Diversity and composition of pollen loads carried by pollinators are primarily driven by insect traits, not floral community characteristics. *Oecologia* **2021**, *196*, 131–143. [[CrossRef](#)]
9. Bell, K.L.; de Vere, N.; Keller, A.; Richardson, R.; Gous, A.; Burgess, K.S.; Brosi, B.J. Pollen DNA barcoding: Current applications and future prospects. *Genome* **2016**, *59*, 629–640. [[CrossRef](#)]
10. Bell, K.L.; Burgess, K.S.; Okamoto, K.C.; Aranda, R.; Brosi, B.J. Review and future prospects for DNA barcoding methods in forensic palynology. *Forensic Sci. Int. Genet.* **2016**, *21*, 110–116. [[CrossRef](#)]
11. Khansari, E.; Zarre, S.; Alizadeh, K.; Attar, F.; Aghabeigi, F.; Salmaki, Y. Pollen morphology of *Campanula* (Campanulaceae) and allied genera in Iran with special focus on its systematic implication. *Flora-Morphol. Distrib. Funct. Ecol. Plants* **2012**, *207*, 203–211. [[CrossRef](#)]
12. Bell, K.L.; Fowler, J.; Burgess, K.S.; Dobbs, E.K.; Gruenewald, D.; Lawley, B.; Morozumi, C.; Brosi, B.J. Applying pollen DNA metabarcoding to the study of plant–pollinator interactions1. *Appl. Plant Sci.* **2017**, *5*, 1600124. [[CrossRef](#)]
13. Galimberti, A.; De Mattia, F.; Bruni, I.; Scaccabarozzi, D.; Sandionigi, A.; Barbutto, M.; Casiraghi, M.; Labra, M. A DNA Barcoding Approach to Characterize Pollen Collected by Honeybees. *PLoS ONE* **2014**, *9*, e109363. [[CrossRef](#)]
14. Bruni, I.; Galimberti, A.; Caridi, L.; Scaccabarozzi, D.; De Mattia, F.; Casiraghi, M.; Labra, M. A DNA barcoding approach to identify plant species in multiflower honey. *Food Chem.* **2015**, *170*, 308–315. [[CrossRef](#)] [[PubMed](#)]
15. Liu, L.; Li, Y.; Li, S.; Hu, N.; He, Y.; Pong, R.; Lin, D.; Lu, L.; Law, M. Comparison of Next-Generation Sequencing Systems. *J. Biomed. Biotechnol.* **2012**, *2012*, 1–11. [[CrossRef](#)]
16. Taberlet, P.; Coissac, E.; Pompanon, F.; Brochmann, C.; Willerslev, E. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol. Ecol.* **2012**, *21*, 2045–2050. [[CrossRef](#)]
17. Biella, P.; Tommasi, N.; Akter, A.; Guzzetti, L.; Klecka, J.; Sandionigi, A.; Labra, M.; Galimberti, A. Foraging strategies are maintained despite workforce reduction: A multidisciplinary survey on the pollen collected by a social pollinator. *PLoS ONE* **2019**, *14*, e0224037. [[CrossRef](#)]
18. McFrederick, Q.S.; Rehan, S.M. Characterization of pollen and bacterial community composition in brood provisions of a small carpenter bee. *Mol. Ecol.* **2016**, *25*, 2302–2311. [[CrossRef](#)]
19. Richardson, R.T.; Lin, C.H.; Sponsler, D.B.; Quijia, J.O.; Goodell, K.; Johnson, R.M. Application of ITS2 metabarcoding to determine the provenance of pollen collected by honey bees in an agroecosystem. *Appl. Plant Sci.* **2015**, *3*, 1400066. [[CrossRef](#)] [[PubMed](#)]
20. Niemeyer, B.; Epp, L.S.; Stoof-Leichsenring, K.R.; Pestryakova, L.; Herzschuh, U. A comparison of sedimentary DNA and pollen from lake sediments in recording vegetation composition at the Siberian treeline. *Mol. Ecol. Resour.* **2017**, *17*, e46–e62. [[CrossRef](#)] [[PubMed](#)]
21. Alsos, I.G.; Lammers, Y.; Yoccoz, N.G.; Jørgensen, T.; Sjögren, P.; Gielly, L.; Edwards, M.E. Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. *PLoS ONE* **2018**, *13*, e0195403. [[CrossRef](#)] [[PubMed](#)]
22. Ezegbogu, M.O. Identifying the scene of a crime through pollen analysis. *Sci. Justice* **2021**, *61*, 205–213. [[CrossRef](#)]
23. Churko, J.M.; Mantalas, G.L.; Snyder, M.P.; Wu, J.C. Overview of High Throughput Sequencing Technologies to Elucidate Molecular Pathways in Cardiovascular Diseases. *Circ. Res.* **2013**, *112*, 1613–1623. [[CrossRef](#)] [[PubMed](#)]
24. Bell, K.L.; Burgess, K.S.; Botsch, J.C.; Dobbs, E.; Read, T.D.; Brosi, B.J. Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Mol. Ecol.* **2019**, *28*, 431–455. [[CrossRef](#)]
25. Blaxter, M.; Mann, J.D.; Chapman, T.; Thomas, F.; Whitton, C.; Floyd, R.; Abebe, E. Defining operational taxonomic units using DNA barcode data. *Philos. Trans. R. Soc. B Biol. Sci.* **2005**, *360*, 1935–1943. [[CrossRef](#)] [[PubMed](#)]
26. Callahan, B.J.; McMurdie, P.; Holmes, S.P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **2017**, *11*, 2639–2643. [[CrossRef](#)]
27. Alberdi, A.; Aizpurua, O.; Gilbert, M.T.; Bohmann, K. Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods Ecol. Evol.* **2018**, *9*, 134–147. [[CrossRef](#)]
28. Ficetola, G.F.; Taberlet, P.; Coissac, E. How to limit false positives in environmental DNA and metabarcoding? *Mol. Ecol. Resour.* **2016**, *16*, 604–607. [[CrossRef](#)]
29. Clare, E.L.; Chain, F.J.; Littlefair, J.E.; Cristescu, M.E. The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data. *Genome* **2016**, *59*, 981–990. [[CrossRef](#)]
30. Cao, Q.; Sun, X.; Rajesh, K.; Chalasani, N.; Gelow, K.; Katz, B.; Shah, V.H.; Sanyal, A.J.; Smirnova, E. Effects of Rare Microbiome Taxa Filtering on Statistical Analysis. *Front. Microbiol.* **2021**, *11*, 607325. [[CrossRef](#)]



31. Pornon, A.; Baksay, S.; Escaravage, N.; Burrus, M.; Andalo, C. Pollinator specialization increases with a decrease in a mass-flowering plant in networks inferred from DNA metabarcoding. *Ecol. Evol.* **2019**, *9*, 13650–13662. [[CrossRef](#)]
32. Wilson, R.S.; Keller, A.; Shapcott, A.; Leonhardt, S.D.; Sickel, W.; Hardwick, J.L.; Heard, T.A.; Kaluza, B.F.; Wallace, H.M. Many small rather than few large sources identified in long-term bee pollen diets in agroecosystems. *Agric. Ecosyst. Environ.* **2021**, *310*, 107296. [[CrossRef](#)]
33. Baksay, S.; Pornon, A.; Burrus, M.; Mariette, J.; Andalo, C.; Escaravage, N. Experimental quantification of pollen with DNA metabarcoding using ITS1 and trnL. *Sci. Rep.* **2020**, *10*, 1–9. [[CrossRef](#)]
34. Bänsch, S.; Tschardt, T.; Wünschiers, R.; Netter, L.; Brenig, B.; Gabriel, D.; Westphal, C. Using ITS2 metabarcoding and mi-croscopy to analyse shifts in pollen diets of honey bees and bumble bees along a mass-flowering crop gradient. *Mol. Ecol.* **2020**, *29*, 5003–5018. [[CrossRef](#)] [[PubMed](#)]
35. Bell, K.L.; Loeffler, V.M.; Brosi, B.J. An rbcL reference library to aid in the identification of plant species mixtures by DNA metabarcoding. *Appl. Plant Sci.* **2017**, *5*, 1600110. [[CrossRef](#)] [[PubMed](#)]
36. Chiara, B.; Francesco, C.; Fulvio, B.; Paola, M.; Annalisa, G.; Stefania, S.; Luigi, A.P.; Simone, P. Exploring the botanical composition of polyfloral and monofloral honeys through DNA metabarcoding. *Food Control.* **2021**, *128*, 108175. [[CrossRef](#)]
37. Danner, N.; Keller, A.; Härtel, S.; Steffan-Dewenter, I. Honey bee foraging ecology: Season but not landscape diversity shapes the amount and diversity of collected pollen. *PLoS ONE* **2017**, *12*, e0183716. [[CrossRef](#)] [[PubMed](#)]
38. De Vere, N.; Jones, L.E.; Gilmore, T.; Moscrop, J.; Lowe, A.; Smith, D.; Hegarty, M.J.; Creer, S.; Ford, C.R. Using DNA metabarcoding to investigate honey bee foraging reveals limited flower use despite high floral availability. *Sci. Rep.* **2017**, *7*, 42838. [[CrossRef](#)]
39. Elliott, B.; Wilson, R.; Shapcott, A.; Keller, A.; Newis, R.; Cannizzaro, C.; Burwell, C.; Smith, T.; Leonhardt, S.D.; Kämper, W.; et al. Pollen diets and niche overlap of honey bees and native bees in protected areas. *Basic Appl. Ecol.* **2021**, *50*, 169–180. [[CrossRef](#)]
40. Fahimee, J.; Badrulisham, A.S.; Zulidzham, M.S.; Reward, N.F.; Muzammil, N.; Jajuli, R.; Md-Zain, B.M.; Yaakop, S. Metabarcoding in Diet Assessment of *Heterotrigona itama* Based on trnL Marker towards Domestication Program. *Insects* **2021**, *12*, 205. [[CrossRef](#)] [[PubMed](#)]
41. Galliot, J.N.; Brunel, D.; Bérard, A.; Chauveau, A.; Blanchetête, A.; Lanore, L.; Farruggia, A. Investigating a flower-insect forager network in a mountain grassland community using pollen DNA barcoding. *J. Insect Conserv.* **2017**, *21*, 827–837. [[CrossRef](#)]
42. Gous, A.; Swanevelder, D.Z.; Eardley, C.D.; Willows-Munro, S. Plant–pollinator interactions over time: Pollen metabarcoding from bees in a historic collection. *Evol. Appl.* **2019**, *12*, 187–197. [[CrossRef](#)]
43. Gous, A.; Eardley, C.D.; Johnson, S.D.; Swanevelder, D.Z.; Willows-Munro, S. Floral hosts of leaf-cutter bees (Megachilidae) in a biodiversity hotspot revealed by pollen DNA metabarcoding of historic specimens. *PLoS ONE* **2021**, *16*, e0244973. [[CrossRef](#)]
44. Hawkins, J.; de Vere, N.; Griffith, A.; Ford, C.R.; Allainguillaume, J.; Hegarty, M.J.; Baillie, L.; Adams-Groom, B. Using DNA metabarcoding to identify the floral composition of honey: A new tool for investigating honey bee foraging preferences. *PLoS ONE* **2015**, *10*, e0134735. [[CrossRef](#)]
45. Jones, L.; Brennan, G.L.; Lowe, A.; Creer, S.; Ford, C.R.; De Vere, N. Shifts in honeybee foraging reveal historical changes in floral resources. *Commun. Biol.* **2021**, *4*, 1–10. [[CrossRef](#)]
46. Khansaritoreh, E.; Salmaki, Y.; Ramezani, E.; Azirani, T.A.; Keller, A.; Neumann, K.; Alizadeh, K.; Zarre, S.; Beckh, G.; Behling, H. Employing DNA metabarcoding to determine the geographical origin of honey. *Heliyon* **2020**, *6*, e05596. [[CrossRef](#)]
47. Leidenfrost, R.M.; Bänsch, S.; Prudnikow, L.; Brenig, B.; Westphal, C.; Wünschiers, R. Analyzing the Dietary Diary of Bumble Bee. *Front. Plant Sci.* **2020**, *11*, 287. [[CrossRef](#)]
48. Lucas, A.; Bodger, O.; Brosi, B.; Ford, C.R.; Forman, D.W.; Greig, C.; Hegarty, M.; Neyland, P.J.; de Vere, N. Generalisation, and specialisation in hoverfly (Syrphidae) grassland pollen transport networks revealed by DNA metabarcoding. *J. Anim. Ecol.* **2018**, *87*, 1008–1021. [[CrossRef](#)] [[PubMed](#)]
49. Lucas, A.; Bodger, O.; Brosi, B.J.; Ford, C.R.; Forman, D.W.; Greig, C.; Hegarty, M.; Jones, L.; Neyland, P.J.; de Vere, N. Floral resource partitioning by individuals within generalised hoverfly pollination networks revealed by DNA metabarcoding. *Sci. Rep.* **2018**, *8*, 1–11. [[CrossRef](#)] [[PubMed](#)]
50. Lucek, K.; Galli, A.; Gurten, S.; Hohmann, N.; Maccagni, A.; Patsiou, T.; Willi, Y. Metabarcoding of honey to assess differences in plant-pollinator interactions between urban and non-urban sites. *Apidologie* **2019**, *50*, 317–329. [[CrossRef](#)]
51. Macgregor, C.J.; Kitson, J.J.N.; Fox, R.; Hahn, C.; Lunt, D.H.; Pocock, M.J.O.; Evans, D.M. Construction, validation, and application of nocturnal pollen transport networks in an agro-ecosystem: A comparison using light microscopy and DNA metabarcoding. *Ecol. Entomol.* **2019**, *44*, 17–29. [[CrossRef](#)]
52. Nürnberger, F.; Keller, A.; Härtel, S.; Steffan-Dewenter, I. Honey bee waggle dance communication increases diversity of pollen diets in intensively managed agricultural landscapes. *Mol. Ecol.* **2019**, *28*, 3602–3611. [[CrossRef](#)]
53. Peel, N.; Dicks, L.; Clark, M.; Heavens, D.; Percival-Alwyn, L.; Cooper, C.; Davies, R.G.; Leggett, R.; Yu, D.W. Semi-quantitative characterisation of mixed pollen samples using MinION sequencing and Reverse Metagenomics (RevMet). *Methods Ecol. Evol.* **2019**, *10*, 1690–1701. [[CrossRef](#)]
54. Piko, J.; Keller, A.; Geppert, C.; Batáry, P.; Tschardt, T.; Westphal, C.; Hass, A.L. Effects of three flower field types on bumblebees and their pollen diets. *Basic Appl. Ecol.* **2021**, *52*, 95–108. [[CrossRef](#)]
55. Pornon, A.; Escaravage, N.; Burrus, M.; Holota, H.; Khimoun, A.; Mariette, J.; Pellizzari, C.; Iribar, A.; Etienne, R.; Taberlet, P.; et al. Using metabarcoding to reveal and quantify plant-pollinator interactions. *Sci. Rep.* **2016**, *6*, 27282. [[CrossRef](#)] [[PubMed](#)]

56. Pornon, A.; Andalo, C.; Burrus, M.; Escaravage, N. DNA metabarcoding data unveils invisible pollination networks. *Sci. Rep.* **2017**, *7*, 1–11. [CrossRef]
57. Potter, C.; de Vere, N.; Jones, L.E.; Ford, C.R.; Hegarty, M.J.; Hodder, K.H.; Diaz, A.; Franklin, E.L. Pollen metabarcoding reveals broad and species-specific resource use by urban bees. *PeerJ* **2019**, *7*, e5999. [CrossRef] [PubMed]
58. Richardson, R.T.; Lin, C.H.; Quijia, J.O.; Riusech, N.S.; Goodell, K.; Johnson, R.M. Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach. *Appl. Plant Sci.* **2015**, *3*, 1500043. [CrossRef] [PubMed]
59. Richardson, R.T.; Curtis, H.R.; Matcham, E.G.; Lin, C.-H.; Suresh, S.; Sponsler, D.B.; Hearon, L.E.; Johnson, R.M. Quantitative multi-locus metabarcoding and waggle dance interpretation reveal honey bee spring foraging patterns in Midwest agroecosystems. *Mol. Ecol.* **2019**, *28*, 686–697. [CrossRef] [PubMed]
60. Richardson, R.T.; Eaton, T.D.; Lin, C.H.; Cherry, G.; Johnson, R.M.; Sponsler, D.B. Application of plant metabarcoding to identify diverse honeybee pollen forage along an urban–agricultural gradient. *Mol. Ecol.* **2021**, *30*, 310–323. [CrossRef] [PubMed]
61. Sickel, W.; Ankenbrand, M.J.; Grimmer, G.; Holzschuh, A.; Härtel, S.; Lanzen, J.; Steffan-Dewenter, I.; Keller, A. Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach. *BMC Ecol.* **2015**, *15*, 1–9. [CrossRef] [PubMed]
62. Simanonok, M.P.; Otto, C.R.V.; Cornman, R.S.; Iwanowicz, D.D.; Strange, J.P.; Smith, T.A. A century of pollen foraging by the endangered rusty patched bumble bee (*Bombus affinis*): Inferences from molecular sequencing of museum specimens. *Biodivers. Conserv.* **2021**, *30*, 123–137. [CrossRef]
63. Smart, M.D.; Cornman, R.S.; Iwanowicz, D.D.; McDermott-Kubeczko, M.; Pettis, J.S.; Spivak, M.S.; Otto, C. A Comparison of Honey Bee-Collected Pollen from Working Agricultural Lands Using Light Microscopy and ITS Metabarcoding. *Environ. Entomol.* **2017**, *46*, 38–49. [CrossRef] [PubMed]
64. Suchan, T.; Talavera, G.; Sáez, L.; Ronikier, M.; Vila, R. Pollen metabarcoding as a tool for tracking long-distance insect migrations. *Mol. Ecol. Resour.* **2019**, *19*, 149–162. [CrossRef]
65. Swenson, S.J.; Gemeinholzer, B. Testing the effect of pollen exine rupture on metabarcoding with Illumina sequencing. *PLoS ONE* **2021**, *16*, e0245611. [CrossRef]
66. Tanaka, K.; Nozaki, A.; Nakadai, H.; Shiwa, Y.; Shimizu-Kadota, M. Using pollen DNA metabarcoding to profile nectar sources of urban beekeeping in Kōtō-ku, Tokyo. *BMC Res. Notes* **2020**, *13*, 1–6. [CrossRef]
67. Tommasi, N.; Biella, P.; Guzzetti, L.; Lasway, J.V.; Njovu, H.K.; Tapparo, A.; Agostinetto, G.; Peters, M.K.; Steffan-Dewenter, I.; Labra, M. Impact of land use intensification and local features on plants and pollinators in Sub-Saharan smallholder farms. *Agric. Ecosyst. Environ.* **2021**, in press. [CrossRef]
68. Tremblay É., D.; Duceppe, M.O.; Thurston, G.B.; Gagnon, M.C.; Côté, M.J.; Bilodeau, G.J. High-resolution biomonitoring of plant pathogens and plant species using metabarcoding of pollen pellet contents collected from a honey bee hive. *Environ. DNA* **2019**, *1*, 155–175. [CrossRef]
69. Vaudo, A.D.; Biddinger, D.J.; Sickel, W.; Keller, A.; López-Urbe, M.M. Introduced bees (*Osmia cornifrons*) collect pollen from both coevolved and novel host-plant species within their family-level phylogenetic preferences. *R. Soc. Open Sci.* **2020**, *7*, 200225. [CrossRef]
70. Metz, C.E. Basic principles of ROC analysis. *Semin. Nucl. Med.* **1978**, *8*, 283–298. [CrossRef]
71. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M.J. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **2011**, *12*, 77. [CrossRef] [PubMed]
72. Youden, W.J. Index for rating diagnostic tests. *Cancer* **1950**, *3*, 32–35. [CrossRef]
73. Anderson, M.J. A new method for non-parametric multivariate analysis of variance. *Austral. Ecol.* **2001**, *26*, 32–46.
74. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **2003**, *14*, 927–930. [CrossRef]
75. Biella, P.; Akter, A.; Ollerton, J.; Nielsen, A.; Klecka, J. An empirical attack tolerance test alters the structure and species richness of plant–pollinator networks. *Funct. Ecol.* **2020**, *34*, 2246–2258. [CrossRef]
76. Dormann, C.F.; Gruber, B.; Fründ, J. Introducing the bipartite package: Analysing ecological networks. *Interaction* **2008**, *8*, 8–11.
77. Doulier, G.; Stouffer, D. Rnetcarto: Fast Network Modularity and Roles Computation by Simulated Annealing, 2015. R Package, Version 0.2.4. Available online: <https://cran.r-project.org/web/packages/rnetcarto/rnetcarto.pdf> (accessed on 7 September 2021).
78. Bruno, A.; Sandionigi, A.; Bernasconi, M.; Panio, A.; Labra, M.; Casiraghi, M. Changes in the drinking water microbiome: Effects of water treatments along the flow of two drinking water treatment plants in a urbanized area, Milan (Italy). *Front. Microbiol.* **2018**, *9*, 2557. [CrossRef] [PubMed]
79. Frigerio, J.; Agostinetto, G.; Galimberti, A.; De Mattia, F.; Labra, M.; Bruno, A. Tasting the differences: Microbiota analysis of different insect-based novel food. *Food Res. Int.* **2020**, *137*, 109426. [CrossRef]
80. Galimberti, A.; Bruno, A.; Mezzasalma, V.; De Mattia, F.; Bruni, I.; Labra, M. Emerging DNA-based technologies to characterize food ecosystems. *Food Res. Int.* **2015**, *69*, 424–433. [CrossRef]
81. Galimberti, A.; Casiraghi, M.; Bruni, I.; Guzzetti, L.; Cortis, P.; Berterame, N.M.; Labra, M. From DNA barcoding to personalized nutrition: The evolution of food traceability. *Curr. Opin. Food Sci.* **2019**, *28*, 41–48. [CrossRef]
82. Galimberti, A.; Bruno, A.; Agostinetto, G.; Casiraghi, M.; Guzzetti, L.; Labra, M. Fermented food products in the era of globalization: Tradition meets biotechnology innovations. *Curr. Opin. Biotechnol.* **2021**, *70*, 36–41. [CrossRef] [PubMed]

83. Casey, J.M.; Meyer, C.P.; Morat, F.; Brandl, S.J.; Planes, S.; Parravicini, V. Reconstructing hyperdiverse food webs: Gut content metabarcoding as a tool to disentangle trophic interactions on coral reefs. *Methods Ecol. Evol.* **2019**, *10*, 1157–1170. [[CrossRef](#)]
84. Arrizabalaga-Escudero, A.; Clare, E.L.; Salsamendi, E.; Alberdi, A.; Garin, I.; Aihartza, J.; Goiti, U. Assessing niche partitioning of co-occurring sibling bat species by DNA metabarcoding. *Mol. Ecol.* **2018**, *27*, 1273–1283. [[CrossRef](#)] [[PubMed](#)]
85. Ruppert, K.M.; Kline, R.J.; Rahman, S. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Glob. Ecol. Conserv.* **2019**, *17*, e00547. [[CrossRef](#)]
86. Andriollo, T.; Gillet, F.; Michaux, J.R.; Ruedi, M. The menu varies with metabarcoding practices: A case study with the bat *Plecotus auritus*. *PLoS ONE* **2019**, *14*, e0219135. [[CrossRef](#)]
87. Zinger, L.; Bonin, A.; Alsos, I.G.; Bálint, M.; Bik, H.; Boyer, F.; Chariton, A.A.; Creer, S.; Coissac, E.; Deagle, B.E.; et al. DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Mol. Ecol.* **2019**, *28*, 1857–1862. [[CrossRef](#)]
88. Elbrecht, V.; Vamos, E.E.; Meissner, K.; Aroviita, J.; Leese, F. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods Ecol. Evol.* **2017**, *8*, 1265–1275. [[CrossRef](#)]
89. Forup, M.L.; Henson, K.S.; Craze, P.G.; Memmott, J. The restoration of ecological interactions: Plant–pollinator networks on ancient and restored heathlands. *J. Appl. Ecol.* **2008**, *45*, 742–752. [[CrossRef](#)]
90. Fisogni, A.; Quaranta, M.; Grillenzoni, F.-V.; Corvucci, F.; de Manincor, N.; Bogo, G.; Bortolotti, L.; Galloni, M. Pollen load diversity and foraging niche overlap in a pollinator community of the rare *Dictamnus albus* L. *Arthropod-Plant Interact.* **2017**, *12*, 191–200. [[CrossRef](#)]
91. Vianna, M.R.; Da Luz, C.F.P.; Kleinert, A.D.M.P. Interaction networks in a Brazilian cerrado: What changes when you add palynological information to floral visitor data? *Apidologie* **2013**, *45*, 418–430. [[CrossRef](#)]
92. Marchand, P.; Harmon-Threatt, A.; Chapela, I. Testing models of bee foraging behavior through the analysis of pollen loads and floral density data. *Ecol. Model.* **2015**, *313*, 41–49. [[CrossRef](#)]
93. Beil, M.; Horn, H.; Schwabe, A. Analysis of pollen loads in a wild bee community (Hymenoptera: Apidae)—A method for elucidating habitat use and foraging distances. *Apidologie* **2008**, *39*, 456–467. [[CrossRef](#)]
94. Bohmann, K.; Gopalakrishnan, S.; Nielsen, M.; Nielsen, L.D.S.B.; Jones, G.; Streicker, D.G.; Gilbert, M.T.P. Using DNA metabarcoding for simultaneous inference of common vampire bat diet and population structure. *Mol. Ecol. Resour.* **2018**, *18*, 1050–1063. [[CrossRef](#)] [[PubMed](#)]
95. Arribas, P.; Andújar, C.; Salces-Castellano, A.; Emerson, B.C.; Vogler, A.P. The limited spatial scale of dispersal in soil arthropods revealed with whole-community haplotype-level metabarcoding. *Mol. Ecol.* **2021**, *30*, 48–61. [[CrossRef](#)]
96. Poos, M.S.; Jackson, D.A. Addressing the removal of rare species in multivariate bioassessments: The impact of methodological choices. *Ecol. Indic.* **2012**, *18*, 82–90. [[CrossRef](#)]
97. Tylianakis, J.M.; Laliberté, E.; Nielsen, A.; Bascompte, J. Conservation of species interaction networks. *Biol. Conserv.* **2010**, *143*, 2270–2279. [[CrossRef](#)]
98. Thébault, E.; Fontaine, C. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science* **2010**, *329*, 853–856. [[CrossRef](#)] [[PubMed](#)]
99. Strydom, T.; Dalla Riva, G.V.; Poisot, T. SVD entropy reveals the high complexity of ecological networks. *EcoEvoRxiv.* **2020**, *2*. Available online: <https://ecoevorxiv.org/q9v85/> (accessed on 7 September 2021).
100. Cirtwill, A.R.; Eklöf, A.; Roslin, T.; Wootton, K.; Gravel, D. A quantitative framework for investigating the reliability of empirical network construction. *Methods Ecol. Evol.* **2019**, *10*, 902–911. [[CrossRef](#)]
101. Alotaibi, S.S.; Sayed, S.M.; Alosaimi, M.; Alharthi, R.; Banjar, A.; Abdulqader, N.; Alhamed, R. Pollen molecular biology: Applications in the forensic palynology and future prospects: A review. *Saudi J. Biol. Sci.* **2020**, *27*, 1185–1190. [[CrossRef](#)]
102. Fan, J.; Upadhye, S.; Worster, A. Understanding receiver operating characteristic (ROC) curves. *CJEM* **2006**, *8*, 19–20. [[CrossRef](#)] [[PubMed](#)]
103. Nutz, S.; Döll, K.; Karlovsky, P. Determination of the LOQ in real-time PCR by receiver operating characteristic curve analysis: Application to qPCR assays for *Fusarium verticillioides* and *F. proliferatum*. *Anal. Bioanal. Chem.* **2011**, *401*, 717–726. [[CrossRef](#)] [[PubMed](#)]
104. Siddique, A.B.; Biella, P.; Unterseher, M.; Albrechtsen, B.R. Mycobiomes of Young Beech Trees Are Distinguished by Organ Rather than by Habitat, and Community Analyses Suggest Competitive Interactions among Twig Fungi. *Front. Microbiol.* **2021**, *12*, 646302. [[CrossRef](#)] [[PubMed](#)]
105. Serrao, N.R.; Reid, S.M.; Wilson, C.C. Establishing detection thresholds for environmental DNA using receiver operator characteristic (ROC) curves. *Conserv. Genet. Resour.* **2018**, *10*, 555–562. [[CrossRef](#)]
106. Ficetola, G.F.; Pansu, J.; Bonin, A.; Coissac, E.; Giguët-Covex, C.; De Barba, M.; Gielly, L.; Lopes, C.M.; Boyer, F.; Pompanon, F.; et al. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Mol. Ecol. Resour.* **2014**, *15*, 543–556. [[CrossRef](#)] [[PubMed](#)]
107. Deagle, B.E.; Thomas, A.C.; McInnes, J.C.; Clarke, L.J.; Vesterinen, E.J.; Clare, E.L.; Kartzinel, T.R.; Eveson, J.P. Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Mol. Ecol.* **2019**, *28*, 391–406. [[CrossRef](#)]