HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Species misidentification in ecological studies: incidence and importance from the ecologists' point of view

Shengyu Wang
Supervised by Katja Rönkä and Sanja Hakala
Master's Programme in
Ecology and Evolutionary Biology
Faculty of Biological and Environmental Sciences
University of Helsinki
May 2021

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Abstract

Faculty: Faculty of Biological and Environmental Sciences

Degree programme: Master's Programme in Ecology and Evolutionary Biology

Study track: Ecology and evolutionary biology

Author: Shengyu Wang

Title: Species misidentification in ecological studies: incidence and importance from the ecologists' point of view

Level: Master's thesis

Month and year: May 2021

Number of pages: 34 + appendices 26 pp.

Keywords: ecology, identification errors, life science, methodology, questionnaire survey, science studies, species misclassification

Where deposited: HELDA - Digital Repository of the University of Helsinki
Additional information:

Abstract:

Natural scientists study a wide variety of species, but whether they have identified all studied samples correctly to species is rarely evaluated. Species misidentification in empirical research can cause significant losses of money, information, and time, and contribute to false results. Thus, I study the abundance of species misidentification and ecologists' perceptions of such mistakes through a web survey targeting researchers from scientific institutes around the globe (including universities, research societies and museums) who completed their doctoral degree in any ecology-related field of science. I received 117 responses with either work or educational background from 30 countries. I found that species misidentification widely existed in respondents' research: almost 70% of the respondents noticed species misidentification in their own research, while the estimated proportion of existing studies with species misidentification was 34% (95% CI: 28% - 40%). Although misidentification was mainly found during specimen collection, specimen handling and data analysis, misidentifications in reporting stages (writing, revision and after publishing) could persist until publication. Moreover, according to respondents, reviewers seldom comment about species identification methods or their accuracy, which may affect respondents' (both leading and not leading a research team) low reporting frequency about the possibility of misidentification. Expert checking, training students, and DNA barcoding are the most prevalent approaches to ensure identification accuracy among respondents. My results imply that species misidentification might be widespread in existing ecological research. Although the problem of species misidentification is widely recognized, such an issue seldom be appropriately handled by respondents. To increase the accuracy of species

identification and maintain academic integrity, I suggest that researchers need to focus more on the study species (e.g., sampling process, identification method, and accuracy) when writing and reviewing papers. Furthermore, I appeal for guidelines about reporting species identification methods and their accuracy in papers, as well as research on education about identification skills in universities, as these two topics may constrain the precision of species identification.

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Table of Contents

1. INTRODUCTION

1.1 Ecology and ecological research

Ecology is the study of the relationships between living organisms and their surrounding environment (including non-biotic environment and biotic environment) (Agarwal, 2008). The scope of ecological research covers various levels of organization spanning from micro-level (e.g., cells) interactions to macro-scale (e.g., biosphere) phenomena. Ecology is one of the primary divisions of biology, equivalent to physiology, genetics, and biochemistry. It can also be subdivided into many sub-disciplines to facilitate understanding and discussion within the field of study. Topics of interest include population ecology, landscape ecology, ecophysiology, community ecology, and conservation science.

The many specialties within ecological research provide us with essential information to better understand the world around us, helping us to improve our environment, conserve endangered species, and protect human health. For example, during the 1960s and 1970s, ecologists linked algal blooms to nutrient enrichment caused by human activities (such as agriculture, industry and sewage treatment), enabling citizens to take the necessary measures to restore lakes and streams in their communities — many of which were once again in popular recreation use for fishing and swimming (Schindler, 1974). Moreover, with the support of ecological research, we have brought countless endangered species back from the brink of extinction or stabilized their population, such as giant panda (*Ailuropoda melanoleuca*) (Swaisgood et al., 2016). A notable example where ecology made essential contributions to public health is Lyme disease, a tick-transmitted infection with potential severe pathogenicity. Ecologists found the connections among acorn, mice, deer and tick in the state of New York, which influence the incidence of Lyme disease and, thus, are able to predict the possibility of infection and notify the public when protective measures are needed (Ostfeld et al., 2006). Modern ecology has become an increasingly rigorous science, guiding us to use natural resources in sustainable ways which provide future generations with a salubrious environment.

1.2 Definition of species

In ecology, species is one of eight major taxonomic ranks (i.e., domain, kingdom, phylum, class, order, family, genus, species) as well as the most basic unit. It is also an essential component forming biodiversity, and thus species is indispensable and a vital factor that needs always to be considered in ecological studies. The concept of species is diverse both historically and contemporarily. For early taxonomists like Linnaeus, species were simply groups of organisms that are visually distinguishable compared to other organisms. Nevertheless, with the deepening of the understanding of organisms, this criterion became no longer sufficient. Ernst Mayr (1942) formalized the idea of biological species concept (BSC), defining "species are groups of actually or potentially interbreeding populations, which are reproductively isolated from other such groups." Additionally, the phylogenetic species concept (PSC) is widely preferred, which emphasizes species as the outcome of evolution: a phylogenetic species is "an irreducible (basal) cluster of organisms diagnosably different from other such clusters, and within which there is a parental pattern of ancestry and descent." (Cracraft, 1989) Nonetheless, all definitions share the fundamental idea that species are segments of lineages at the population level of biological organization (de Queiroz, 2005).

Taxonomy is the biological discipline of discovering, describing, classifying, and naming species and other taxa (Tancoigne et al., 2011). Taxonomists classify organisms into taxa and assign them a taxonomic rank. Groups of a certain rank can be aggregated to form a more comprehensive group of higher rank, thus formulating a taxonomic hierarchy. Taxonomy is the foundation of species identification, the process of assigning an organism into a pre-existing taxon (Dallwitz, 1992). Thus, although it is not necessary for researchers who identify species to be taxonomists, the most professional expert in species identification is usually regarded as taxonomist with a university degree, trained with classical taxonomy, molecular taxonomy, and cyber taxonomy (standardized electronic tools for sharing and accessing information) (Wheeler, 2008). The vast array of disciplines, including ecology, natural resource management, medical science even governance, rely heavily on taxonomic information (Cracraft, 2002). Nevertheless, the "taxonomic impediment" has been widely acknowledged, which is the gaps in taxonomic knowledge, the lack of taxonomists and curators, and the impact these deficits have on our ability to manage, use, and conserve the benefits of biodiversity (Giangrande, 2003; Cao et al., 2016).

1.3 Methods of identifying species

Numerous methods have been developed to identify species, including molecular techniques, morphological analysis, audio recording, chemical identification and biogeographic approaches. Morphological measurements might be the most common means of identification, widely applied to all kinds of organisms. It is a systematic analysis of organisms themselves (including specimen), their tracks, pictures or videos, according to previous description and identification key (an identification tool based on gross-morphology visible features leading to a correct identification) of the studied species. For example, the shape, size, and complexity of male genitalia are considered to be one of the most essential and useful species-diagnostic features in insect systematics (Tuxen, 1970). Similarly, leaves and flowers are the most widely used among different organs of the plant since leaves are usually flat and accessible almost all year round, while flowers have a high distinguishing capacity (Nguyen et al., 2017).

Modern taxonomy increasingly relies on molecular techniques, where nucleotide variation in the genome has been recognized as the optimum approach examining inheritable differences (Morgan & Blair, 1998). Characters in DNA sequence can solve the problem of inheritable variation within species, which extensively exist in morphological traits (Hillis, 1987). There has been continuous progress within molecular approaches, the most popular being DNA barcoding, an affordable and straightforward species identification technique (Čandek & Kuntner, 2015) using short DNA fragments from specific genes. After collecting samples, DNA from each sample is extracted, amplified and sequenced, and can then be used to identify an organism to a species by comparing with the reference library (e.g., GenBank, http://www.ncbi.nlm.nih.gov; EMBL, http://www.ebi.ac.uk/embl; DDBJ, http://www.ddbj.nig.ac.jp) of DNA fragments from all sequenced species (Valentini et al., 2009), just as a supermarket salesman scans the black stripes of UPC barcodes to identify commodities in stock according to their database. GenBank is an eminent genetic sequence database supporting various phylogenetic and population genetic studies, such as examining genetic distance among species (Johns & Avise, 1998). There have even been requests for DNA taxonomy, assigning species primarily based on sequences in such databases (Tautz et al., 2003).

Metabolome, the profile of the small-molecule chemicals which participate in metabolic reactions required for sustaining life activities (Harrigan & Goodacre, 2003), is one of the manifestations of the genetic differences which distinguish one species from another. As the final downstream product of the genome, the small-molecule profile ultimately reflects an organism's genes and thus could also be applied for genomic fingerprinting and assessing genetic relatedness between species (Musah et al., 2015). Fatty acid profiles measured by DART (one of the most common mass spectrometric methods enabling rapid determination of sample composition) for different bacterial species have been shown to be distinct and reproducible (Pierce et al., 2007).

Methods of identifying species are not limited to these, many other approaches are also applied extensively. For example, audio recording has been applied to identify many vocal species, like frogs (Yuan & Ramli, 2012), birds and whales (Nanni et al., 2017). In practice, biogeographic approaches are applied to estimate a species' occurrence possibility in a specific region based on its known geographical distribution and environmental requirements (Ancillotto et al., 2020). In many circumstances, only species known to live in a particular area are considered when identifying a species. Innovative technologies are also rising. A study has shown that automated plant identification is sufficiently capable of several routine tasks, providing promising tools for autonomous ecological surveillance systems (Bonnet et al., 2018).

When identifying species, scientists comprehensively consider the feasibility, accuracy, difficulty and cost among different identification methods. Hence, it is crucial to study the identification schemes they usually use and concerns they might have.

1.4 Consequences of species misidentification

False identification in ecological studies can cause profound practical consequences. If the samples under study are not correctly identified, it could affect our understanding of the ecosystem, public health, and environmental management programs, wasting a lot of resources and effort. For example, three species of the coral *Montastraea annularis* with different growth rates, oxygen isotopic ratios and metabolisms are commonly utilized as indicators of environmental degradation and global climate variation (Knowlton et al., 1992; Knowlton & Jackson, 1994). However, when these species and their biology are not adequately recognized, the presumed environmental signal is confounded (Knowlton et al., 1992; Knowlton & Jackson, 1994). Hence, the failure of identifying species affects our understanding of the ecosystem and the environmental factors affecting it.

Minor misidentification may profoundly impact public health, wasting large amounts of money, even human lives. For example, because of the large number of potentially involved species in malaria transmission, the vector system in Southeast Asia is a complex study topic (Van Bortel et al., 2001). However, the classification of *Anopheles* usually has low accuracy as it primarily relied on intuitive taxonomic descriptions of a limited number of morphological similarities (Foley et al., 1998). Consequently, a nonmalaria-vector *Anopheles* species was wrongly identified as a malaria vector for a long time, which significantly hampered the evaluation of the disease and follow-up preventive control practices (Van Bortel et al., 2001).

Misidentification originated in environmental management programs could also cause grave consequences. A clear example of this error comes from a restoration project, transplanting cordgrass

Spartina foliosa on the West Coast of the US. This project team mistakenly chose *Spartina densiflora*, believing it is *Spartina foliosa*, the only *Spartina* species described for the region by then. When the mistakenly chosen species, *Spartina densiflora*, was correctly recognized about 30 years later (Kittelson & Boyd, 1997), repeated transplant had led to the proliferation of this erroneous species, massively displacing native organisms and changing regional landscapes of the West Coast (Daehler & Strong, 1996). At this stage, eradicating *Spartina densiflora* would be remarkably expensive with tremendous uncertainties.

Furthermore, different studies in biological disciplines, even in other areas, share and assimilate results, facilitating the propagation of misidentification among disciplines, as illustrated in the cases mentioned above. Therefore, considering the severe consequences that species misidentification could cause, it is particularly vital to explore the abundance and reasons for this issue.

1.5 Ecologists as a study system

There are a number of factors affecting scientists' methods used to identify species and opinions regarding species misidentification. For example, their education experience, which includes the number and quality of species identification courses scientists had, the hierarchical structure of academia (Elkins & Keller, 2003; Keller, 2017) in which research team leaders and those who are in charge of the identifications might be more influential, funding (Boero, 2010), which may limit the development of related studies and methods that researchers can use.

The publishing process may also have practical significance in monitoring the accuracy of species identification. Peer review is a process of evaluating an author's academic work, research or ideas by one or more people who are experts with similar competencies. In the peer-review process, the author's manuscript is usually forwarded to initial checks assessing its suitability (Ware, 2008). For example, manuscripts with errors, data deficiency or insufficient discussion would be declined. The peer-review process in academia is crucial for assessing research's originality, significance, and quality (which include species identification method and its accuracy).

1.6 Aim of the study

The topics mentioned above are of vital importance to ecological research. However, no study systematically evaluates ecologists' view of species misidentification. Thus, I study respondents estimated incidence of species misidentification at different stages in the course of ecological research and their opinions about this issue. Methods mentioned by the respondents to ensure correct species identification are also summarized and discussed. This study has practical significance for improving the accuracy and efficiency of species identification by revealing and influencing ecologists' views regarding species misidentification.

2. MATERIALS AND METHODS

2.1 Study design

This is a web-based survey gathering information from a subset of ecologists to study their opinions on the occurrence and importance of species misidentification. The survey was designed using Webropol 3.0. The people answering the questionnaire are referred to as "respondents" (Fink, 2003b). The whole questionnaire with all questions and answer options can be found in Appendix 1. The summary of survey data can be found in Appendix 2.

The survey collected information about ecology researchers' background and opinions on species misidentification. I asked how common researchers think misidentification is in the whole field of ecology and in their own research to estimate the abundance of species misidentification. Furthermore, if they had detected misidentifications in their work, I asked at what stage the potential misidentification are noticed to study the incidence of misidentifications throughout the publication process. Moreover, to study researchers' attitude generally, I asked how often reviewers ask or comment about their species identification methods or accuracy, and their reporting frequency of misidentification or its possibility. The frequency to discuss and consider, and how problematic they consider species misidentification in relation to their own research are also investigated to study the degree of attention to such an issue among researchers.

Background information was collected voluntarily and anonymously (see section 2.3 below for ethical considerations), including the number of peer-reviewed journal articles respondents have published, their role in the research group (leading or not a research group), the fields of ecological science and taxa they mainly work with, and the number of species they usually focus on. There is also a set of free-text questions designed to encourage the respondents to specify their background or share their experience and opinions with species misidentification, including "Did you fix the error(s), and if yes, how?", "Do you use some specific methods to ensure that your or your students' species identifications are correct?" and "Do you have some final comments, anecdotes, or stories regarding the topic or any of the questions in this survey?"

2.2 Targeted respondents

The study targeted researchers from scientific institutes around the globe who completed their doctoral degree in any ecology-related field of science. In this study, I only targeted researchers who had already received their doctoral degrees since senior researchers have more experience and are more likely to have encountered species misidentification. I also tracked the incidence of misidentifications throughout the publication process, in which students have not been as involved as senior researchers, and therefore respondents who have multiple publications would provide more data.

The target population of this study are researchers from all taxa and all fields in ecological studies. Taxa in the questionnaire are divided into 8 categories: “Archaea”, “Bacteria”, “Single-celled eukaryotes”, “Fungi”, “Nonvascular plants”, “Vascular plants”, “Invertebrate animals” and “Vertebrate animals”. In addition, I listed 15 common areas (e.g., community ecology, taxonomy, and evolution biology) in the field of ecology, asking the respondents to choose the most relevant options. They were also encouraged to specify their study taxa and research fields in the subsequent free-text questions.

2.3 Ethical considerations

This survey was anonymous. I did not collect any personal data from the respondents. Giving background information at the end of the survey was voluntary, and I did not use it to identify respondents or their publications. I did not treat the data in any way that would jeopardize respondents’ anonymity. However, I will publish the anonymized data to promote transparency of research. The respondents’ background information will be published only in a summary format and not as records of individual respondents. Furthermore, I may exclude some data from publication if the content might reveal respondents’ identity (such as answers in free-text format). The University of Helsinki owns the data, and Dr. Sanja Hakala is the holder of the data.

2.4 Dissemination

The survey was open from 29 April 2020 to 7 August 2020, a total of 100 days, distributing worldwide through Twitter and email. I created a Twitter account for this project (@MIS_identify), in which advertisement Twitters were posted regularly with related hashtags (e.g., #Misidentification, #speciesidentification, #Ecology, and #Survey) to spread the survey. I also used Twitter to tag (@) ecology societies, ecology-related university faculties and departments, and natural museums so that they are more likely to notice and participate in the survey. Additionally, I contacted ecological societies by sending emails or filling the contact forms to ask them to join and spread the survey. I also sent distribution emails to ecology-related university faculties and departments, natural museums, and ecology researchers (listed in Appendix 3).

I used "snowball sampling", where existing respondents are encouraged to spread the survey further to their acquaintances (Fink, 2003a). This technique is usually used in nonprobability sampling (which does not guarantee that all eligible units have an equal chance of being included in a sample) and when a population listing is unavailable and cannot be compiled (Fink, 2003a). Newly identified members name others, and the sample group increases like a rolling snowball.

2.5 Statistical analyses

This research is fundamentally an exploratory study where I inspect the incidence of species misidentification throughout publication processes and researchers' opinions about species misidentification. The results are analyzed and plotted in RStudio (Version 1.4.1103). To calculate the mean value and confidence interval of respondents estimated proportion of published studies with misidentified species in all ecological fields of science, I applied t.test() with "conf.level = 0.95". I used the function ggballoonplot() in ggplot2-based (Wickham, 2016) "ggpubr" package (Kassambara, 2020) to visualize the stages where respondents noticed misidentification, building a contingency table formed by research stages and misidentification rates.

I explored the relationship between researchers' opinions and background in my data with the correlation test, where Kendall rank correlation coefficients (also commonly referred to as Kendall's τ coefficient), which is a non-parametric test measuring the strength of dependence between two variables, were calculated. I used the cor() function to calculate the correlation coefficients between survey questions, in which the "use" is "complete.obs", which means missing values are handled by casewise deletion (and if there are no complete cases, that gives an error). Furthermore, I used the corrplot() function in the "corrplot" package (Wei & Simko, 2017) to build a correlation matrix visualizing the correlation coefficients.

Additionally, to analyze whether being a research team leader or not (i.e. researcher' experience) affects how problematic they consider species misidentification in relation to their own research, I fitted a linear model (Gaussian distribution) using the lm() function of "stats" package (Core Team,

2017) in R. To analyze whether being a research team leader or not affects the reporting frequency about species misidentification, I used a generalized linear model with negative binomial distribution using the glm.nb() function of "MASS" package (Venables & Ripley, 2002). Respondents' views are substituted by numbers (e.g., "1" represents "Never", "2" represents "Rarely") in the data set to be able to use these distributions, which fit the data sufficiently well.

3. RESULTS

3.1 Data acquisition

I received 117 respondents currently working in 27 countries with a doctoral background in 25 countries. 53% of the respondents are leading their own research group (refer to Appendix 2 figure). The respondents are from various fields of ecological studies (Fig. 1a) and studying a broad spectrum of taxa (Fig. 1b).

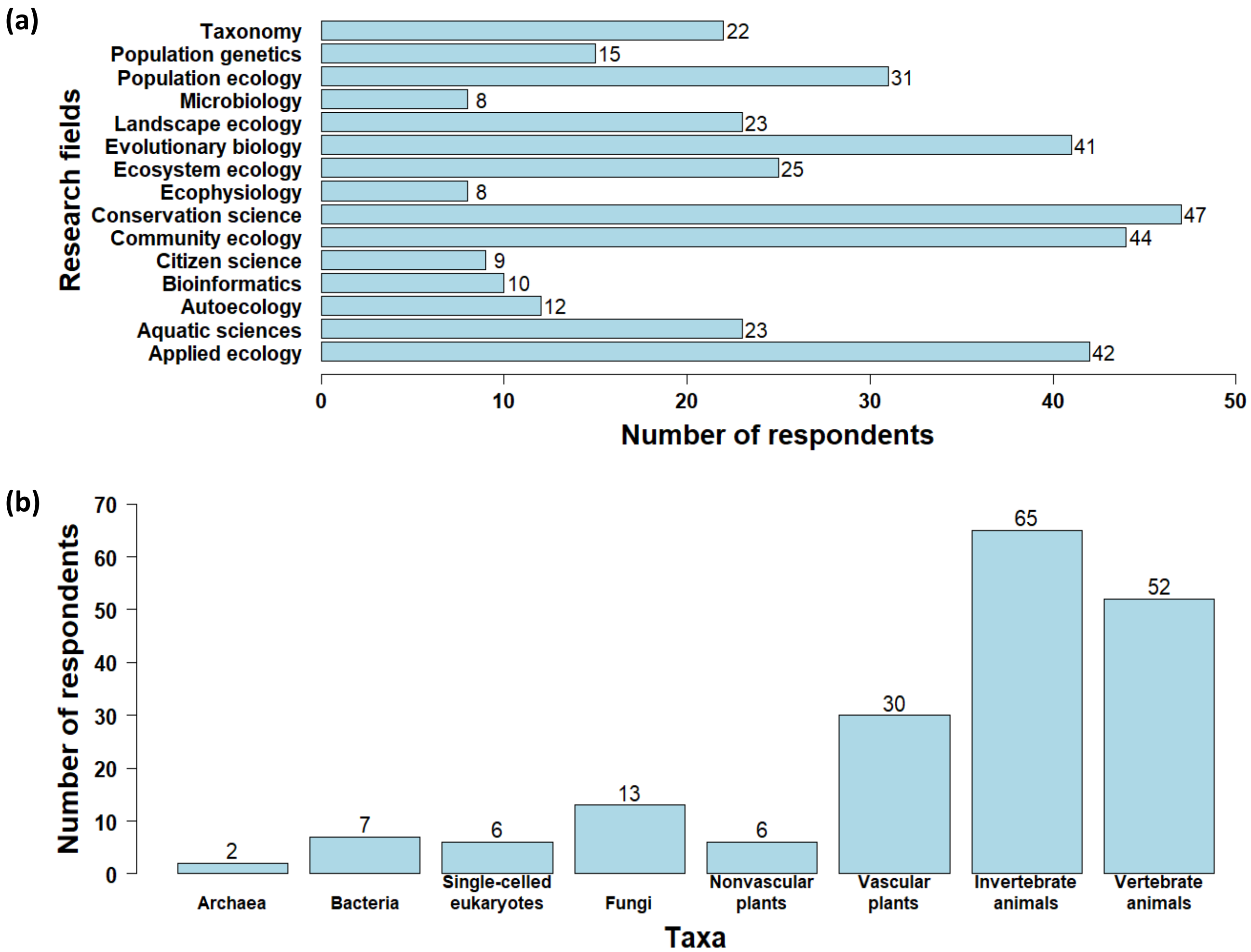


Figure 1 Background of respondents. (a) Respondents' fields of ecological related science (n = 117). **(b)** Respondents' main research taxa (n = 117). "n" indicates the number of respondents. The number next to and above the bars indicate the numbers of respondents who selected each option.

3.2 How respondents estimate the abundance of species misidentification and the stages of the ecological study at which misidentification are noticed

Considering all research stages from data collection to published papers, 68.4% of the respondents noticed species misidentification in their research (Fig. 2). Those who have never noticed species misidentification in their own work or do not remember constitute 31.6%.

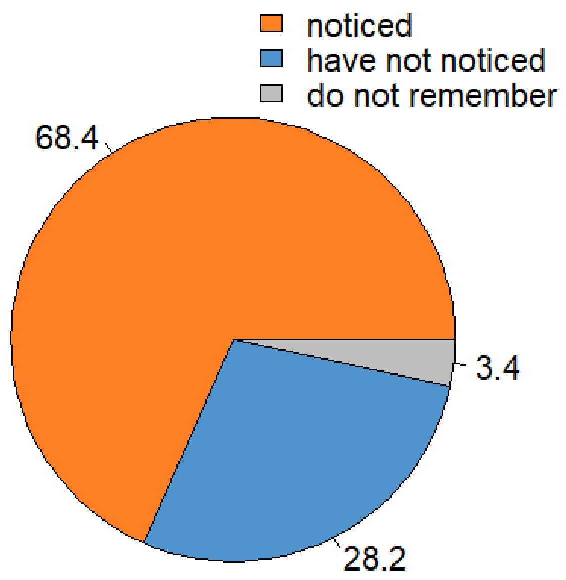


Figure 2 Percentage of respondents who have noticed species misidentification in their research considering all stages of research from data collection to published papers (n = 117).

Most respondents (33.3%) replied, “I don’t know”, when they were asked to estimate the proportion of published studies with species misidentification (Fig. 3). The second most common response was the lowest proportion, [0 - 10%], taking up 17.1%. With some fluctuation, there is a downward trend from the interval [21% - 30%] to the interval [91% - 100%], which constitute 13.7% to 0.9% of all respondents, respectively. Excluding those who answered “I don’t know”, respondents estimated the proportion of existing studies with species misidentification to be 34% (95% confidence interval, 28% - 40%).

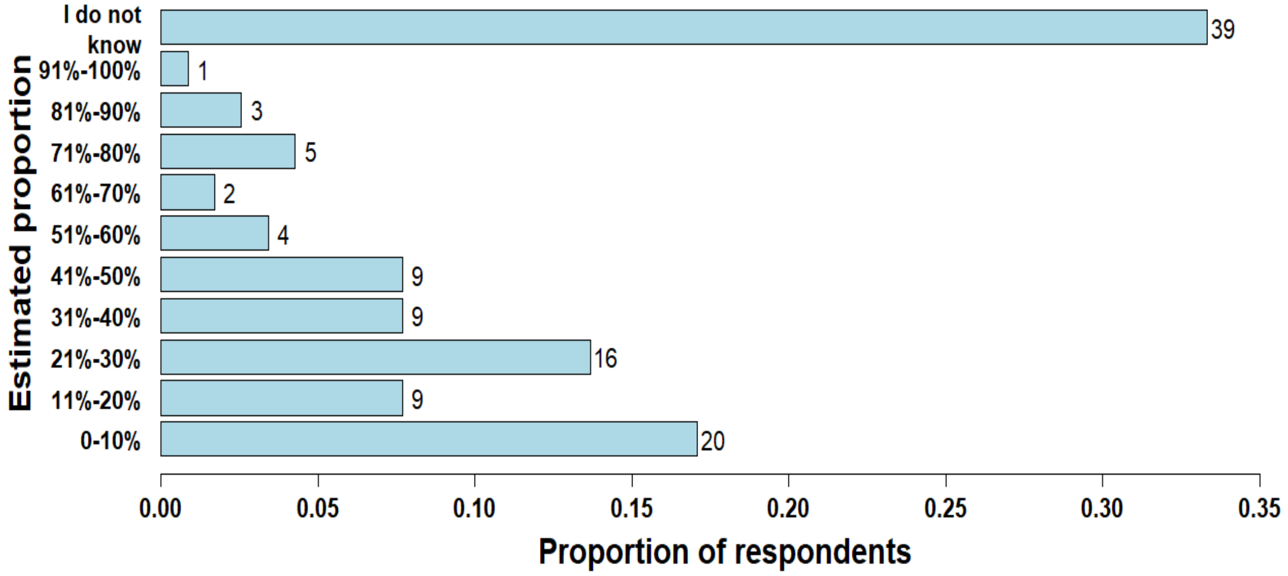


Figure 3 Respondents' estimated proportion of published studies with misidentified species in all ecological fields of science (n = 117). Continuous data was divided with 10% intervals for the figure.

Among those who have noticed species misidentification in their own research (n = 80, Fig. 2), I also studied the research stages where misidentification is noticed. As shown in Figure 4, 0% is the most selected misidentification rate for almost all stages, followed by 20%. Respondents reported that they encountered misidentification mainly during specimen collection (e.g., field and greenhouse), specimen handling (e.g., laboratory work, microscoping and video analysis) and data analysis, whereas they reported fewer misidentifications in later stages (writing, revision and after publishing).

A few respondents seldom correct misidentification for published data. When they were asked whether and how they fix misidentification (optional free-text question), many of them (n ≈ 6, the symbol "≈" indicates that data was collected from free-text questions) share a similar idea: "When detected during data analysis or revision they have been considered and corrected one way or another. When already published, in general, no re-analyses have been done, and no action is taken." (quote from one of the six) Some other respondents said that they only fix misidentification in follow-up papers (n ≈ 4). Many respondents (n ≈ 6) simply replied "no", which means even if they have noticed a mistake, they did not take any action.

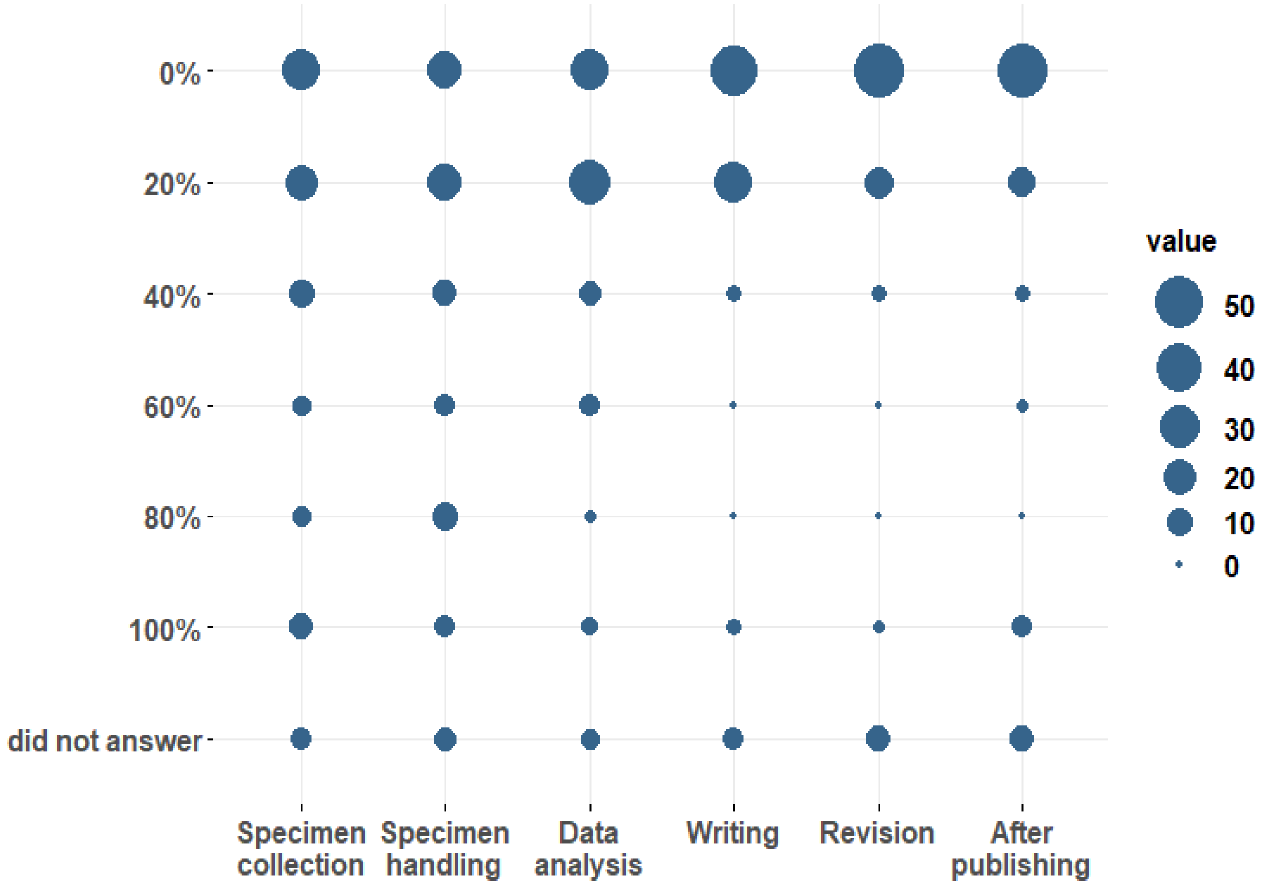


Figure 4 The stage of the ecological study or peer review process at which misidentification is noticed (n = 80). Each column and row show different scientific study stages and misidentification frequencies estimated by the respondents, respectively. Sizes of circles (values) indicate the number of respondents: big circles represent large numbers of respondents, while little circles account for small numbers of respondents.

3.3 Researchers' attitude regarding species identification

When it comes to the peer review process, 45.3% of the respondents reported that they have never been asked or commented on their species identification methods or their accuracy, followed by those who were asked rarely and sometimes, taking up 37.6% and 20.0%, respectively. None were asked every time (Fig. 5).

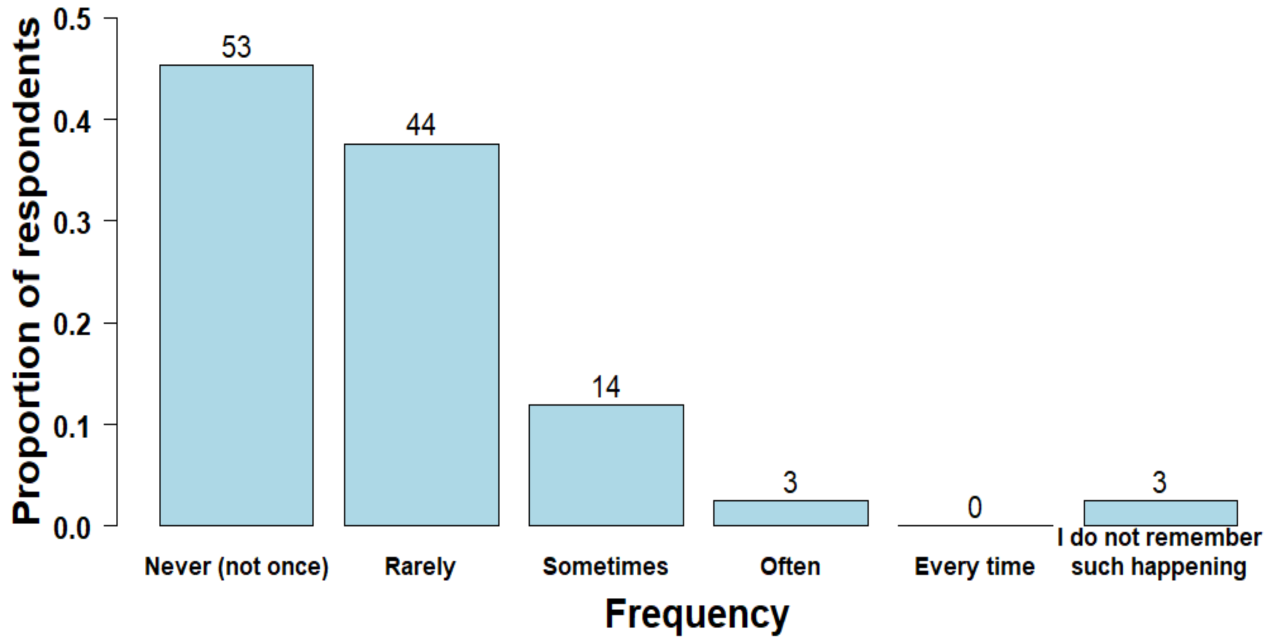


Figure 5 The reported frequency of reviewers having asked or commented about species identification methods or their accuracy in respondents' research (n = 117).

There is a similar trend among respondents when asked about reporting the possibility of species misidentification in papers. More than 35% of the respondents believe that reporting the possibility of species misidentification for their study is not necessary. Approximately 25% of them rarely report the possibility, while around 13% have never thought about that. In contrast, those who report every time or usually constitute 4.27% and 10.26%, respectively (Fig. 8).

I further constructed a correlation matrix to explore the relationships among survey questions (Fig. 6). The main questions (Q2 - Q5) generally correlate with each other, which means respondents who were asked more frequently about identification methods by reviewers tend to consider and discuss more about species misidentification, and think such topic more problematic. Reporting about species misidentification was correlated with considering, discussing the topic more, and being asked more frequently about identification methods by reviewers. Moreover, those who have had more misidentification in their own work (Q7) seem to think it is more common in the whole field (Q1).

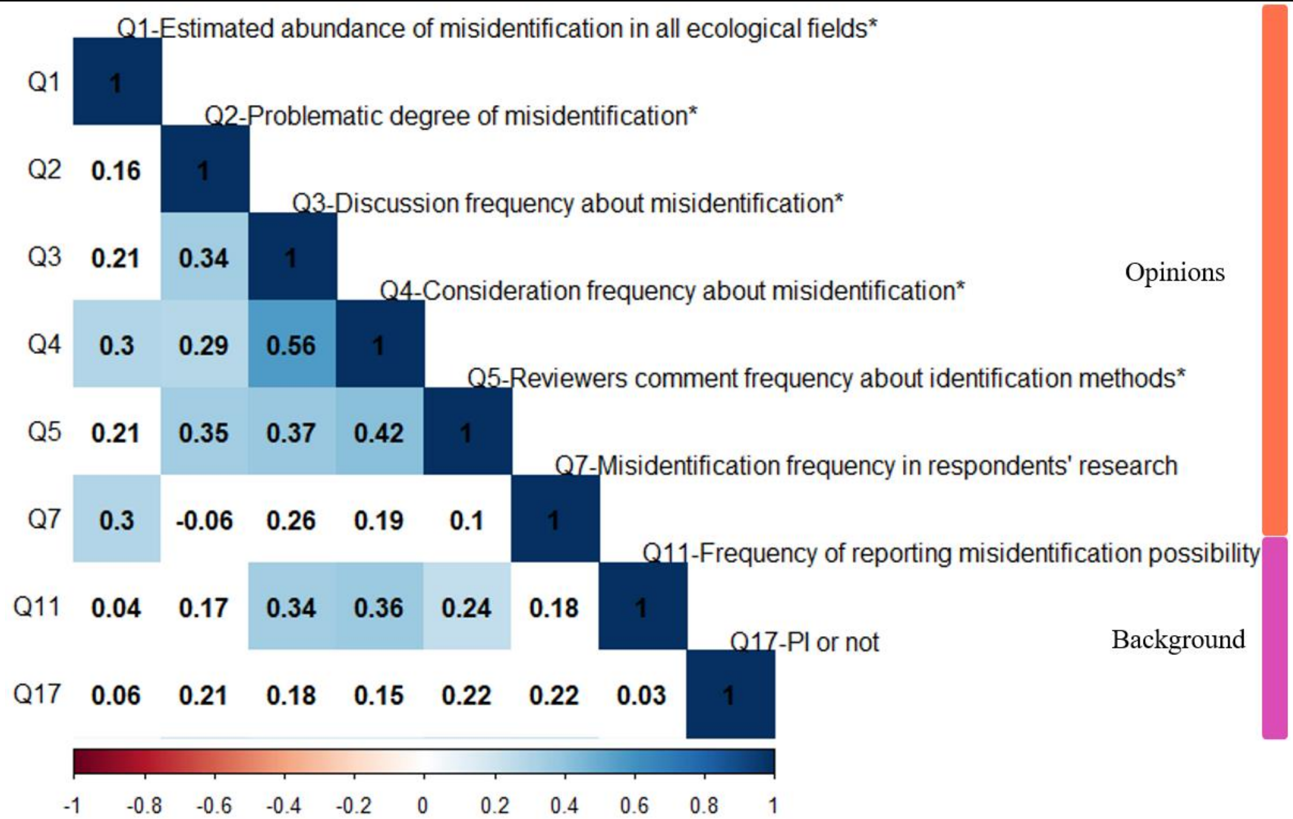


Figure 6 Correlation matrix of each pair of questions. Correlation coefficients are shown on each square and colored according to the value. Blue indicates positive correlations, while red indicates negative correlations. The darker the color, the stronger the correlation. Insignificant correlations (p > 0.01) are left blank. I only chose questions explaining the major factors influencing researchers' opinions about species misidentification. The letter "Q" followed by a number means the question number in the questionnaire, e.g., "Q1" represents Question 1. Compulsory questions are marked with an asterisk (*). The whole questionnaire with all questions and answer options can be found in Appendix 1.

Q1 was about respondents estimated proportion of species misidentification in all ecological fields of science.
Q2 was about how problematic respondents consider the issue of species misidentification in their own work.
Q3 was about discussion frequency about species misidentification with colleagues or students.
Q4 was about researchers considering species identification problems when reading other papers.
Q5 was about reviewers ask or comment about species identification methods or their accuracy.
Q7 was about the species misidentification frequency in respondents' research.
Q11 was about the reporting frequency about the possibility of species misidentification in papers.
Q17 was about respondents' role in their research group.

Figures 7 and 8 compare researchers who are leading or not a research group regarding their views and reporting frequency about species misidentification. In general, principal investigators (PIs) have more extended research experience and more publications than non-PI's, I thus expect that they would report the misidentification possibility more frequently. However, according to my respondents, research group leaders consider species misidentification more problematic than non-PIs do (T-test: TS = 2.54, se = 0.219, p = 0.012), but they do not report the possibility of misidentification more frequently than those who do not lead a research team (T-test: TS = 0.167, se = 0.135, p = 0.867).

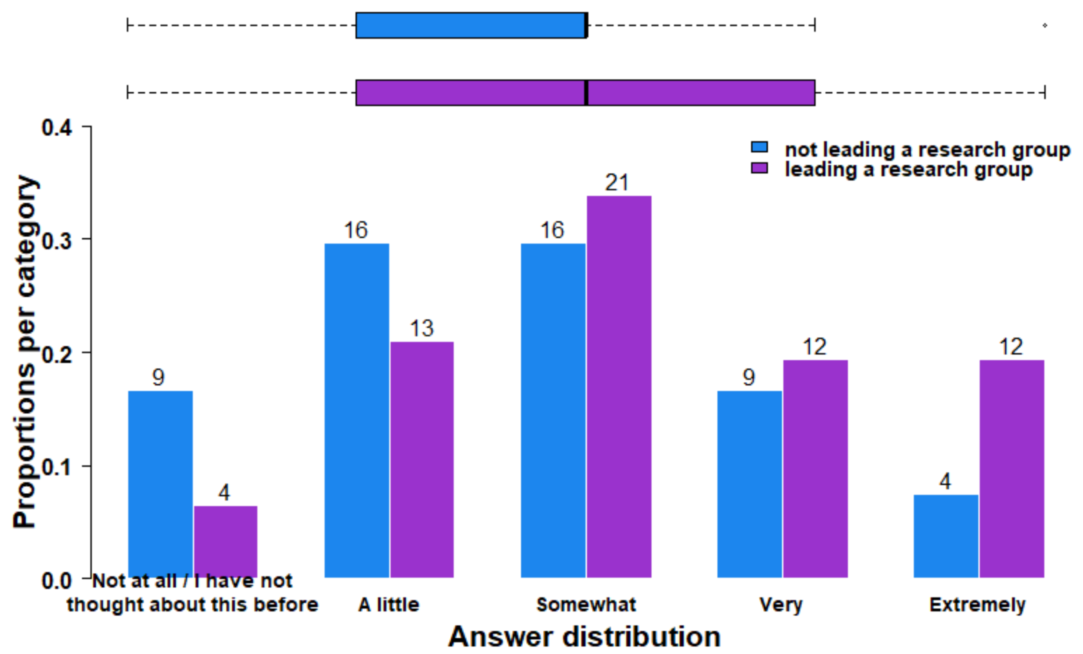


Figure 7 How problematic respondents consider species misidentification in relation to their own research (n = 117).

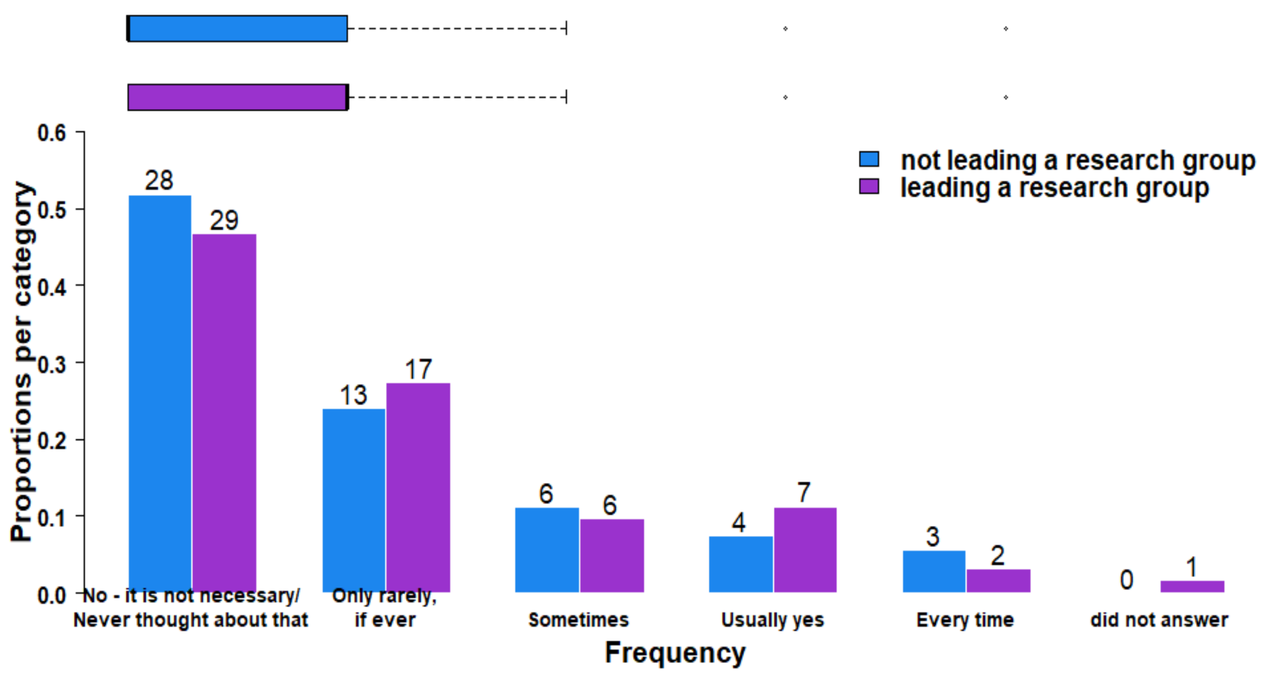


Figure 8 Frequency of reporting the possibility of species misidentification (n = 117).

3.4 Methods of ensuring species identification accuracy

When respondents were asked about their methods to ensure the accuracy of their or their students' species identifications (optional free-text question), the most frequent answers were expert checking (n ≈ 15), training students (n ≈ 14) and DNA barcoding (n ≈ 11). Notably, when they were encouraged to leave some final comments, anecdotes, or stories regarding the topic of species misidentification or ask questions about the survey (optional free-text question), some respondents reported their concerns with "training students" (n ≈ 4), whereas none of them had concerns or negative experience about DNA barcoding.

4. DISCUSSION

4.1 The estimated abundance of species misidentification

Based on my results (section 3.2), the proportion of respondents who noticed species misidentification in their own research is moderately high (n = 80). The mean estimated proportion of existing studies with species misidentification lies in the interval of 34% ± 6%, which is also a fairly high percentage. Both of these results imply that species misidentification is a relatively common problem in ecological research.

Some respondents rarely revise misidentification for a published study (section 3.2). This is probably because revising species identification problems encountered in a published study is much more challenging than in the early research stages. Misidentification noticed in the early stages of research is sometimes effortless to be adjusted. Nevertheless, it might be the opposite for the reporting stages like writing, revision, and publishing, since modifying the species identity may require changing the data, analysis methods and conclusion, even submitting an erratum or retracting the paper. In addition, researchers' promotion and tenure are closely associated with the quantity and quality of their publications (Ali et al., 1996). Therefore, researchers might be more likely to publish new articles in reputable journals than revise mistakes in a published article.

Even if researchers do dedicate themselves to revising the erroneous data, it may not be straightforward. For instance, concerns related to the quality of data in GenBank seem rather prominent among researchers (Harris, 2003): sequence errors in GenBank are widely prevalent (Karlin et al., 2001; Forster, 2003), which overturned research conclusions drawn from these original erroneous data (Hagelberg et al., 2000). These findings revealed a deficit of decent editing in GenBank. The quality control of the raw data submitted to GenBank may entirely depend on the researchers who produced it. These problems might be ongoing as many of my respondents had a similar experience with erroneous sequences in GenBank (n ≈ 4): "it is astonishing to see so many authors find that their identifications are mistaken, but they fail to change them in GenBank." My respondent reported that some incorrect IDs in the GenBank dataset were provided with citation history leading to discovering the erroneous IDs and subsequent correction. Nevertheless, corrections have never been down in GenBank, which further exposed a deficiency in GenBank when revising erroneous data.

Thus, although respondents encountered misidentification primarily in specimen collection, specimen handling, and data analysis, misidentification noticed during writing, revision, and after publishing would still require considerable effort from researchers correcting them.

4.2 Researchers' perceptions and responsibilities

Individual researchers, journals, and evaluation systems must be taken into account in order to establish rigorous and vibrant academia. In the academic community, it is the researchers' responsibility to ensure the accuracy of their data, which undoubtedly includes identifying their study species correctly. By default, researchers abide by the integrity principle, the data in experiments and papers are authentic, and we thus hold a trusting attitude towards published papers. In this way, later researchers can safely use these data and conclusions, and continue to study them, so that science can continue to progress iteratively. If there is a problem with species identity, we cannot trust any inferences build on this premise. This means that basic scientific research, which consumes much energy, will have to pay a higher cost because of misidentification (Bortolus, 2008).

Among researchers, the principal investigator's (PIs) leadership is particularly crucial for a research team (Keller, 2017). It is PIs' duty to effectively recruit, coordinate, and motivate knowledge professionals from a diverse range of backgrounds and disciplines (Elkins & Keller, 2003). Among the respondents, PIs have more concern than non-PIs (Fig. 7), probably because they are responsible for avoiding potential mistakes. Interestingly, they report in a similar manner as non-PIs (Fig. 8), which presumably because researchers do not realize to report misidentification possibility or worry that it would affect the credibility of their articles. Moreover, PIs and non-PIs regularly collaborate to publish articles (Kyvik & Reymert, 2017), which could be another explanation for their comparable frequency in reporting misidentification. The deeper reasons for this phenomenon are worth exploring in future research. Although the way of leading a research team can be various, I would suggest that

PIs should take more responsibility guiding other team members to identify their study species correctly and provide more details about identification methods and accuracy when publishing.

Besides researchers having to be responsible for the accuracy of species identification, journals and their evaluation systems also have a responsibility to monitor the quality of their publications (Carraway, 2009). In the process of peer review, the accuracy of species identification in manuscripts is supposed to be assessed. Subsequently, suppose a journal publishes an article with misidentified species, the corresponding editor and reviewers for the publication should take responsibility to find out the reason, urge the author to revise the mistake and take precautions so that we can maintain academic preciseness. My results (Fig. 6) imply that researchers tend to consider and discuss more the topic of misidentification and report more frequently about the possibility of species misidentification when reviewers ask or comment more on their identification methods or accuracy. Nevertheless, respondents reported that reviewers seldom do so (Fig. 5), which may limit the identification of potential errors and succeeding modification. Hence, the peer review process is crucial in ascertaining the accuracy of species identification and can assist scientists to filter out imprecise identification results.

4.3 Species identification accuracy

Investigating researchers' methods of ensuring correct species identification and exploring why they use these methods are beneficial to improve the efficiency and accuracy of species identification. Expert checking ($n \approx 15$), training students ($n \approx 14$) and DNA barcoding ($n \approx 11$) ranked in the top when respondents were asked to describe methods ensuring their or their students' species identifications accuracy.

4.3.1 Dependence on proficient

My respondents profoundly rely on expert checking ($n \approx 15$) for their identification results. However, for many respondents, the definition of "expert" simply is someone familiar with the species they want to identify, which agrees with a previous study, stating that taxonomy has increasingly relied on professionals in other fields who become part-time taxonomist when needed (Godfray & Knapp, 2004). The need for experienced colleagues is probably due to the low availability of taxonomists, whose funding and position have subsided dramatically (Boero, 2010). Since taxonomy is mistakenly recognized as a descriptive science with no intellectual challenges nor experimentation (Felsenstein, 2004; Padial et al., 2010), and since editors in journals with high impact factor reluctant to publish taxonomic research which thus limit tenure and promotion (Werner, 2009), taxonomy attracts diminishing young researchers, which further contribute to “taxonomic impediment” (Giangrande, 2003; Cao et al., 2016). Some scientists hold a negative attitude towards professional taxonomy, where a significant resurgence in financial support might be unrealistic (Agnarsson & Kuntner, 2007; Packer et al., 2009). I would argue that several approaches still can be made to increase taxonomists’ voice in species identification. The measures may include citing the original author(s) of a formally named organism, and inviting taxonomists as co-authors when the conclusions of the paper are solely dependent on the accuracy of the study taxon (Wägele et al., 2011). I further appeal that academia view taxonomy as crucial and rigorous science which would promote the classification and description of species, and provide substantial assistance for species identification.

4.3.2 Education about species identification skills

Student identification ($n \approx 14$) is also playing an indispensable role among my respondents, in which the usual procedure is to recommend handbooks, determination keys, or checklists before the students begin their project and personally double-check or spot check their performance. However, available studies about education in species identification mainly focus on elementary education, such as children’s interests and student-teachers’ views about species identification (Randler, 2008; Palmberg et al., 2015; Skarstein & Skarstein, 2020). No research is about species identification teaching in university and its role in scientific research based on my knowledge. Some respondents expressed their worry that ecology students rarely learn sufficient species identification skills or that the relevant courses come too late in their education. “It is alarming to see a biogeographer who has never identified a plant but can produce great maps and analyses in R and python to get a taxonomy position, while those who know the importance of this discipline as fundamental research get overlooked.” These experiences imply that the training of species identification skills may be defective in the education system related to ecology. In most cases, it might be unrealistic to know beforehand which taxa an ecology student will be working with in the future. What should prepare for students is to teach them basic principles of identification skills (e.g., the use of proper identification keys and online databases (Prakash et al., 2017)) as well as essential characteristics of major taxa in relevant identification courses. Supervisors or taxonomic experts can then give project-specific training before collecting or identifying samples so that the outcomes based on student identification would be more accurate, and the processes of sample collection and handling would be more effective. In the future, research on the education system of species identification in higher education and its effects on identification accuracy is still needed.

4.3.3 Use of DNA barcoding

DNA barcoding (n ≈ 11) is another essential strategy applied to ensure identification accuracy, especially when identification based on morphology is challenging. Indeed, identifying species based on morphology alone could be insufficient in some circumstance since closely related species might have high-morphological similarities: homoplasy (Monis, 1999), phenotypic plasticity (Galazzo et al., 2002), and a lack of conserved structures (Jousson & Bartoli, 2001). On the other hand, DNA barcoding is widely recognized as a novel alternative, providing accurate, rapid and automatable species identifications (Hebert & Gregory, 2005). It links different life-history stages and improves the accuracy and efficiency of field studies involving different taxa with high morphological identification difficulty (Janzen, 2004; Moritz & Cicero, 2004). It also has been shown to bring substantial benefits to various areas such as biosecurity (Ficetola et al., 2008), conservation (Francis et al., 2010), biodiversity assessment (Gotelli, 2004) and paleoecology (Kuch et al., 2002; Willerslev et al., 2003).

Despite the extensive benefits that DNA barcoding can bring to a diverse range of biological disciplines, shortcomings in this method are also of concern. For example, up to 17% of species misidentification in barcoding studies has been observed in cypraeid marine gastropods if the reference database is not comprehensive (Meyer & Paulay, 2005). Errors in GenBank, which I discussed in section 4.1, could also constrain the accuracy of DNA barcoding. Additionally, for environmental samples inspecting DNA from dead animals or dead parts of plants, it can be complex to amplify DNA fragments long enough (usually >500 bp (Hebert et al., 2003)) for barcoding (Deagle et al., 2006). Nevertheless, none of my respondents had negative report regarding DNA barcoding, which proved the reliability of this method to a certain extent. I believe that DNA barcoding will become an increasingly reliable approach in species identification, attracting more ecologists. The current limitation would be remedied through the availability of reliable databases, which are now under construction (Valentini et al., 2009), and the use of shorter barcoding markers (Meusnier et al., 2008).

4.4 Reporting identification accuracy

In this study, I found that researchers' scheme treating species identification might be concerning. A previous study has shown that even in top international peer-reviewed journals published from 2005 to 2007 in ecological disciplines, more than half of articles did not indicate how they identified the organisms in their research (Bortolus, 2008). Moreover, gray literature (Mahood et al., 2014), including technical reports and ecological theses were used to base species identification more commonly than specialized taxonomic literature. This indicates that ecologists seldom identify or report their study taxa appropriately. Even to this day, according to the respondents, nearly half of them have never been asked or commented about their species identification methods or accuracy by reviewers. The frequency of reporting the possibility of species misidentification is also fairly low (both for PIs and non-PIs). Most respondents believe that stating the likelihood of species misidentification for their study is unnecessary, indicating either that they are studying relatively readily recognizable species and are confident with their identification results or they merely do not know how to report the possibility of species misidentification. In some circumstances, identifying species is indeed straightforward. "It's pretty hard to get it wrong," according to a respondent who works on European honeybees. There are also many similar cases, in which the study species are grown in a laboratory or greenhouse, the chance of identifying a species incorrectly may indeed be trivial, and thus when reporting in papers, researchers could mention where the strain original stocked, and how the study materials were sampled and stored (e.g., Ning et al., 2017). Nevertheless, ecology papers, especially those related to genetics, mainly focus on sample treatment and data analysis in their "materials and methods" section but rarely mention the sample itself, such as breeding stock (Bortolus, 2008). On the other hand, I checked the Author Guidelines of prominent journals in the field of ecology, such as Ecology, Ecology Letters, Journal of Ecology, Nature Ecology and Evolution, and The American Naturalist, but instructions about how to report the possibility of misidentification are missing. I would suggest that researchers should at least indicate the identification method they used and the corresponding accuracy if applicable (e.g., Yuan & Ramli, 2012). Assigning the study species to higher classification ranks, such as genus or family (e.g., Ko et al., 2013), or listing several potential identities that the study species could be, may also be a wise alternative of dealing with uncertainties, based on my respondents. Still, instruction guiding researchers on reporting the species identification accuracy is urgently needed to improve transparency, reliability and repeatability (Ihle et al., 2017).

4.5 Representativeness of the data

The respondents are from all continents (except Antarctica), with either study or work background of 30 countries (Fig. 9), providing excellent resources to investigate the patterns and opinions about species misidentification. However, continent and country distribution are highly unbalanced due to my connections, local education development, and language, which means my results represent all regions unequally. Overall, most of the respondents are from western countries, where most ecological research is conducted. More respondents from those less studied regions like Asia, South America and Africa could have altered my result and increased its scalability. The estimation would also have higher representativeness of opinions of ecologists worldwide. For example, if I would have received more answers from tropical areas, respondents estimated proportion of species misidentification might be higher, considering the higher species diversity but less available research, scholar and funding in the tropics (New, 1995).

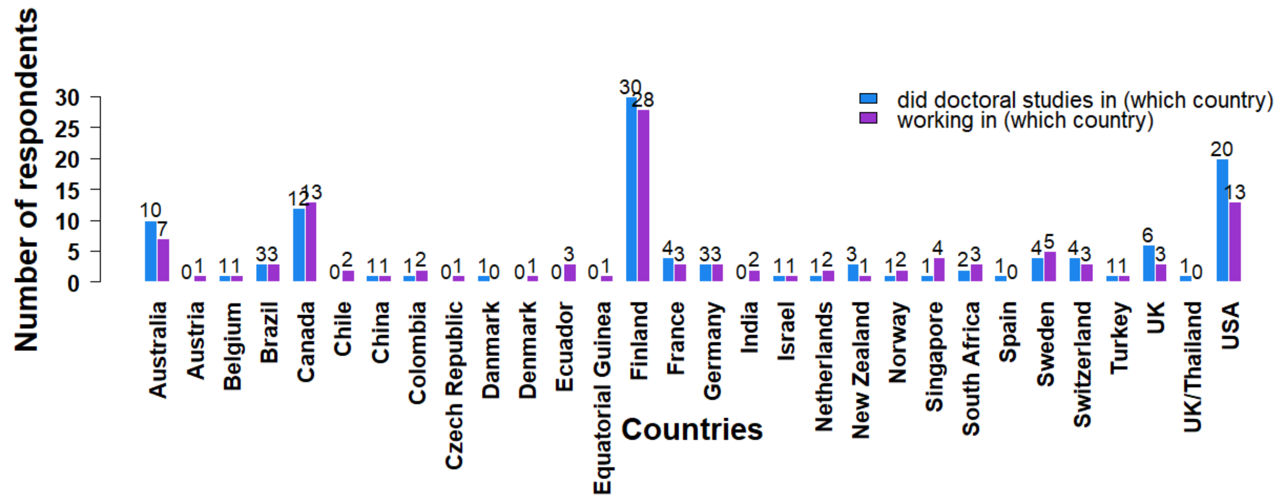


Figure 9 The countries where respondents completed their doctoral degrees and the countries where they worked.

Distributing the survey through personal contact is one of the leading advertising approaches. Dr. Katja Rönkä, Dr. Sanja Hakala, Dr. Kaisa Raatikainen, and I invited researchers we know to join the study and encouraged them to spread the survey further to their colleagues. Since 3/4 of survey initiators being Finnish, the survey is most widely distributed in Finland. Approximately 1/4 of the respondents are from Finland, the foremost source of information.

Additionally, since the survey is voluntary, researchers who consider species misidentification significant might be more likely to answer, whereas those who view this topic as inconsequential to their work would probably ignore the survey. Consequently, there is a possibility that my results may be more representative of those researchers who care about this problem.

Researchers in ecological societies, universities, and museums are my primary target group, but the development and distribution of those institutions are highly unbalanced globally: North America and Europe have more prestigious scientific institutions, educating innumerable elite students and attracting researchers to study and work there, while Latin America, Africa, and Asia have relatively fewer such organizations. Thus, it is plausible to have more respondents from the USA, Canada, Australia, and European countries. Although my data does not comprise the whole world equally, it does cover many areas and institutions where most ecological research is carried out (Martin et al., 2012). I also have some respondents from Africa, Asia and South America, so my data can represent the ecologists' opinions from these regions to some extent, and thus it is not wholly biased.

The degree of similarity between languages could also affect the distribution of respondents. The survey has only the English version, and all means of distribution are in English. Consequently, my questionnaire was more prevalent among native speakers of English and other alphabetic languages. In contrast, it was likely to be ignored by native speakers of other languages, like Chinese, Korean and Japanese, which limits the spread of the survey in these groups. Despite the equal effort to distribute in developed countries with the advanced education system like Japan, much fewer respondents were from those regions.

5. CONCLUSIONS

According to respondents' estimation, there is a relatively high proportion of species misidentification in ecological research. Although concerns about species misidentification widely exist among respondents, it seems rare for them to report identification methods and accuracy in papers. In general, reviewers seldom ask or comment about authors' species identification methods or accuracy. Expert checking, training students and DNA barcoding are of great importance in ensuring species identification accuracy among my respondents. In addition, this study found that guidelines about how to report species misidentification are deficient. Research on the education system about identification skills in universities is also urgently needed. I call for future research to explore these two topics.

ACKNOWLEDGEMENTS

I am very grateful to my supervisors, Dr. Katja Rönkä (University of Helsinki, Finland) and Dr. Sanja Hakala (University of Fribourg, Switzerland), for their support throughout this project. Only under their guidance can I complete such an excellent thesis project and improved a lot. I thank Dr. Kaisa Raatikainen (University of Jyväskylä, Finland) for her technical support. She gave me numerous valuable suggestions for survey design and data analysis. I also appreciate those who participated,

advertised, and commented on my survey. It is the information they provided made this research possible.

REFERENCES

Agarwal, S. K. (2008). *Fundamentals of ecology*. APH Publishing.

Agnarsson, I., & Kuntner, M. (2007). Taxonomy in a changing world: seeking solutions for a science in crisis. *Systematic Biology, 56*(3), 531-539.

Ali, S. N., Young, H. C., & Ali, N. M. (1996). Determining the quality of publications and research for tenure or promotion decisions: a preliminary checklist to assist. *Library Review,*

Ancillotto, L., Bosso, L., Smeraldo, S., Mori, E., Mazza, G., Herkt, M., Galimberti, A., Ramazzotti, F., & Russo, D. (2020). An African bat in Europe, Plecotus gaisleri: Biogeographic and ecological insights from molecular taxonomy and Species Distribution Models. *Ecology and Evolution, 10*(12), 5785-5800.

Boero, F. (2010). The study of species in the era of biodiversity: a tale of stupidity. *Diversity, 2*(1), 115-126.

Bonnet, P., Goëau, H., Hang, S. T., Lasseck, M., Šulc, M., Malécot, V., Jauzein, P., Melet, J., You, C., & Joly, A. (2018). Plant identification: experts vs. machines in the era of deep learning. *Multimedia tools and applications for environmental & biodiversity informatics* (pp. 131-149). Springer.

Bortolus, A. (2008). Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *AMBIO: A Journal of the Human Environment, 37*(2), 114-118.

Čandek, K., & Kuntner, M. (2015). DNA barcoding gap: reliable species identification over morphological and geographical scales. *Molecular Ecology Resources, 15*(2), 268-277.

Cao, X., Liu, J., Chen, J., Zheng, G., Kuntner, M., & Agnarsson, I. (2016). Rapid dissemination of taxonomic discoveries based on DNA barcoding and morphology. *Scientific Reports, 6*(1), 1-13.

Carraway, L. N. (2009). Ethics for and responsibilities of authors, reviewers and editors in science. *The American Midland Naturalist, 161*(1), 146-164.

Core Team, R. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing. *Vienna, Austria: URL Https://Www.R-Project.Org/.*

Cracraft, J. (1989). Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. *Speciation and its Consequences, 28*, 59.

Cracraft, J. (2002). The seven great questions of systematic biology: an essential foundation for conservation and the sustainable use of biodiversity. *Annals of the Missouri Botanical Garden,* , 127-144.

Daehler, C. C., & Strong, D. R. (1996). Status, prediction and prevention of introduced cordgrass Spartina spp. invasions in Pacific estuaries, USA. *Biological Conservation, 78*(1-2), 51-58.

Dallwitz, M. J. (1992). A comparison of matrix-based taxonomic identification systems with rule-based systems. Paper presented at the *Proceedings of IFAC Workshop on Expert Systems in Agriculture*, 215-218.

de Queiroz, K. (2005). A unified concept of species and its consequences for the future of taxonomy. *Proceedings of the California Academy of Sciences,*

Deagle, B. E., Eveson, J. P., & Jarman, S. N. (2006). Quantification of damage in DNA recovered from highly degraded samples–a case study on DNA in faeces. *Frontiers in Zoology, 3*(1), 1-10.

Elkins, T., & Keller, R. T. (2003). Leadership in research and development organizations: A literature review and conceptual framework. *The Leadership Quarterly, 14*(4-5), 587-606.

Felsenstein, J. (2004). A digression on history and philosophy. *Inferring Phylogenetics.Sinauer,* , 123-146.

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters, 4*(4), 423-425.

Fink, A. (2003a). Sound survey design and sampling. *The Survey Handbook,* , 31-45.

Fink, A. (2003b). What is a survey? When do you use one. *Fink A.the Survey Handbook.USA,*

Foley, D. H., Bryan, J. H., Yeates, D., & Saul, A. (1998). Evolution and Systematics ofAnopheles: Insights from a Molecular Phylogeny of Australasian Mosquitoes. *Molecular Phylogenetics and Evolution, 9*(2), 262-275.

Forster, P. (2003). To err is human. *Annals of Human Genetics, 67*(1), 2-4.

Francis, C. M., Borisenko, A. V., Ivanova, N. V., Eger, J. L., Lim, B. K., Guillén-Servent, A., Kruskop, S. V., Mackie, I., & Hebert, P. D. (2010). The role of DNA barcodes in understanding and conservation of mammal diversity in Southeast Asia. *PloS One, 5*(9), e12575.

Galazzo, D. E., Dayanandan, S., Marcogliese, D. J., & McLaughlin, J. D. (2002). Molecular systematics of some North American species of Diplostomum (Digenea) based on rDNA-sequence data and comparisons with European congeners. *Canadian Journal of Zoology, 80*(12), 2207-2217.

Giangrande, A. (2003). Biodiversity, conservation, and the ‘Taxonomic impediment’. *Aquatic Conservation: Marine and Freshwater Ecosystems, 13*(5), 451-459.

Godfray, H., & Knapp, S. (2004). Introduction. Taxonomy for the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences, 359*(1444), 559.

Gotelli, N. J. (2004). A taxonomic wish–list for community ecology. *Philosophical Transactions of the Royal Society of London.Series B: Biological Sciences, 359*(1444), 585-597.

Hagelberg, E., Goldman, N., Lio, P., Whelan, S., Schiefenhövel, W., Clegg, J. B., & Bowden, D. K. (2000). Evidence for mitochondrial DNA recombination in a human population of island Melanesia: correction. *Proceedings of the Royal Society of London.Series B: Biological Sciences, 267*(1452), 1595-1596.

Harrigan, G. G., & Goodacre, R. (2003). *Metabolic profiling: its role in biomarker discovery and gene function analysis: its role in biomarker discovery and gene function analysis*. Springer Science & Business Media.

Harris, D. J. (2003). Can you bank on GenBank? *Trends in Ecology & Evolution, 18*(7), 317-319.

Hebert, P. D., & Gregory, T. R. (2005). The promise of DNA barcoding for taxonomy. *Systematic Biology, 54*(5), 852-859.

Hebert, P. D., Ratnasingham, S., & De Waard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London.Series B: Biological Sciences, 270*(suppl_1), S96-S99.

Hillis, D. M. (1987). Molecular versus morphological approaches to systematics. *Annual Review of Ecology and Systematics, 18*(1), 23-42.

Ihle, M., Winney, I. S., Krystalli, A., & Croucher, M. (2017). Striving for transparent and credible research: practical guidelines for behavioral ecologists. *Behavioral Ecology, 28*(2), 348-354.

Janzen, D. H. (2004). Now is the time. *Philosophical Transactions of the Royal Society of London.Series B: Biological Sciences, 359*(1444), 731-732.

Johns, G. C., & Avise, J. C. (1998). A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene. *Molecular Biology and Evolution, 15*(11), 1481-1490.

Jousson, O., & Bartoli, P. (2001). Molecules, morphology and morphometrics of Cainocreadium labracis and Cainocreadium dentecis n. sp.(Digenea: Opecoelidae) parasitic in marine fishes. *International Journal for Parasitology, 31*(7), 706-714.

Karlin, S., Bergman, A., & Gentles, A. J. (2001). Annotation of the Drosophila genome. *Nature, 411*(6835), 259-260.

Kassambara, A. (2020). *Ggpubr:'ggplot2'based Publication Ready Plots (R Package Version 0.4.0).*

Keller, R. T. (2017). A longitudinal study of the individual characteristics of effective R&D project team leaders. *R&D Management, 47*(5), 741-754.

Kittelson, P. M., & Boyd, M. J. (1997). Mechanisms of expansion for an introduced species of cordgrass, Spartina densiflora, in Humboldt Bay, California. *Estuaries, 20*(4), 770-778.

Knowlton, N., & Jackson, J. B. (1994). New taxonomy and niche partitioning on coral reefs: jack of all trades or master of some? *Trends in Ecology & Evolution, 9*(1), 7-9.

Knowlton, N., Weil, E., Weigt, L. A., & Guzmán, H. M. (1992). Sibling species in Montastraea annularis, coral bleaching, and the coral climate record. *Science, 255*(5042), 330-333.

Ko, H., Wang, Y., Chiu, T., Lee, M., Leu, M., Chang, K., Chen, W., & Shao, K. (2013). Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS One, 8*(1), e53451.

Kuch, M., Rohland, N., Betancourt, J. L., Latorre, C., Steppan, S., & Poinar, H. N. (2002). Molecular analysis of a 11 700-year-old rodent midden from the Atacama Desert, Chile. *Molecular Ecology, 11*(5), 913-924.

Kyvik, S., & Reymert, I. (2017). Research collaboration in groups and networks: differences across academic fields. *Scientometrics, 113*(2), 951-967.

Mahood, Q., Van Eerd, D., & Irvin, E. (2014). Searching for grey literature for systematic reviews: challenges and benefits. *Research Synthesis Methods, 5*(3), 221-234.

Martin, L. J., Blossey, B., & Ellis, E. (2012). Mapping where ecologists work: biases in the global distribution of terrestrial ecological observations. *Frontiers in Ecology and the Environment, 10*(4), 195-201.

Mayr, E. (1942). Systematics and the origin of species Columbia University Press. *New York, 334*

Meusnier, I., Singer, G. A., Landry, J., Hickey, D. A., Hebert, P. D., & Hajibabaei, M. (2008). A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics, 9*(1), 1-4.

Meyer, C. P., & Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol, 3*(12), e422.

Monis, P. T. (1999). Invited review the importance of systematics in parasitological research. *International Journal for Parasitology, 29*(3), 381-388.

Morgan, J., & Blair, D. (1998). Mitochondrial ND1 gene sequences used to identify echinostome isolates from Australia and New Zealand. *International Journal for Parasitology, 28*(3), 493-502.

Moritz, C., & Cicero, C. (2004). DNA barcoding: promise and pitfalls. *PLoS Biol, 2*(10), e354.

Musah, R. A., Espinoza, E. O., Cody, R. B., Lesiak, A. D., Christensen, E. D., Moore, H. E., Maleknia, S., & Drijfhout, F. P. (2015). A high throughput ambient mass spectrometric approach to species identification and classification from chemical fingerprint signatures. *Scientific Reports, 5*(1), 1-16.

Nanni, L., Aguiar, R. L., Costa, Y. M., Brahnam, S., Silla Jr, C. N., Brattin, R. L., & Zhao, Z. (2017). Bird and whale species identification using sound images. *IET Computer Vision, 12*(2), 178-184.

New, T. R. (1995). *Introduction to invertebrate conservation biology*

Nguyen, T. T., Le, T., Vu, H., Nguyen, H., & Hoang, V. (2017). A combination of deep learning and hand-designed feature for plant identification based on leaf and flower images. Paper presented at the *Asian Conference on Intelligent Information and Database Systems,* 223-233.

Ning, K., Chen, S., Huang, H., Jiang, J., Yuan, H., & Li, H. (2017). Molecular characterization and expression analysis of the SPL gene family with BpSPL9 transgenic lines found to confer tolerance to abiotic stress in Betula platyphylla Suk. *Plant Cell, Tissue and Organ Culture (PCTOC), 130*(3), 469-481.

Ostfeld, R. S., Canham, C. D., Oggenfuss, K., Winchcombe, R. J., & Keesing, F. (2006). Climate, deer, rodents, and acorns as determinants of variation in Lyme-disease risk. *PLoS Biology, 4*(6), e145.

Packer, L., Grixti, J. C., Roughley, R. E., & Hanner, R. (2009). The status of taxonomy in Canada and the impact of DNA barcoding. *Canadian Journal of Zoology, 87*(12), 1097-1110.

Padial, J. M., Miralles, A., De la Riva, I., & Vences, M. (2010). The integrative future of taxonomy. *Frontiers in Zoology, 7*(1), 1-14.

Palmberg, I., Berg, I., Jeronen, E., Kärkkäinen, S., Norrgård-Sillanpää, P., Persson, C., Vilkonis, R., & Yli-Panula, E. (2015). Nordic–Baltic student teachers’ identification of and interest in plant and animal species: The importance of species identification and biodiversity for sustainable development. *Journal of Science Teacher Education, 26*(6), 549-571.

Pierce, C. Y., Barr, J. R., Cody, R. B., Massung, R. F., Woolfitt, A. R., Moura, H., Thompson, H. A., & Fernandez, F. M. (2007). Ambient generation of fatty acid methyl ester ions from bacterial whole cells by direct analysis in real time (DART) mass spectrometry. *Chemical Communications,* (8), 807-809.

Prakash, P. Y., Irinyi, L., Halliday, C., Chen, S., Robert, V., & Meyer, W. (2017). Online databases for taxonomy and identification of pathogenic fungi and proposal for a cloud-based dynamic data network platform. *Journal of Clinical Microbiology, 55*(4), 1011-1024.

Randler, C. (2008). Teaching species identification–A prerequisite for learning biodiversity and understanding ecology. *Eurasia Journal of Mathematics, Science and Technology Education, 4*(3), 223-231.

Schindler, D. W. (1974). Eutrophication and recovery in experimental lakes: implications for lake management. *Science, 184*(4139), 897-899.

Skarstein, T. H., & Skarstein, F. (2020). Curious children and knowledgeable adults–early childhood student-teachers’ species identification skills and their views on the importance of species knowledge. *International Journal of Science Education, 42*(2), 310-328.

Swaisgood, R., Wang, D., & Wei, F. (2016). Ailuropoda melanoleuca. *The IUCN Red List of Threatened Species, 2016*, e. T712A102080907.

Tancoigne, E., Bole, C., Sigogneau, A., & Dubois, A. (2011). Insights from Zootaxa on potential trends in zoological taxonomic activity. *Frontiers in Zoology, 8*(1), 1-13.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H., & Vogler, A. P. (2003). A plea for DNA taxonomy. *Trends in Ecology & Evolution, 18*(2), 70-74.

Tuxen, S. L. (1970). Taxonomist's glossary of genitalia in insects.

Valentini, A., Pompanon, F., & Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in Ecology & Evolution, 24*(2), 110-117.

Van Bortel, W., Harbach, R. E., Trung, H. D., Roelants, P., Backeljau, T., & Coosemans, M. (2001). Confirmation of Anopheles varuna in Vietnam, previously misidentified and mistargeted as the malaria vector Anopheles minimus. *The American Journal of Tropical Medicine and Hygiene, 65*(6), 729-732.

Venables, W. N., & Ripley, B. D. (2002). Modern applied statistics (Fourth S., editor) New York.

Wägele, H., Klussmann-Kolb, A., Kuhlmann, M., Haszprunar, G., Lindberg, D., Koch, A., & Wägele, J. W. (2011). The taxonomist-an endangered race. A practical proposal for its survival. *Frontiers in Zoology, 8*(1), 1-7.

Ware, M. (2008). *Peer review: benefits, perceptions and alternatives*. Citeseer.

Wei, T., & Simko, V. (2017). *R Package "corrplot": Visualization of a Correlation Matrix (Version 0.84),*

Werner, Y. L. (2009). The aspiration to be good is bad: The "Impact Factor" hurts both science and society. *Int J Sci Soc, 1*, 99-105.

Wheeler, Q. D. (2008). *The new taxonomy*. CRC Press.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis.,(Springer-Verlag: New York, NY, USA.).

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D. A., & Cooper, A. (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science, 300*(5620), 791-795.

Yuan, C. L. T., & Ramli, D. A. (2012). Frog sound identification system for frog species recognition. Paper presented at the *International Conference on Context-Aware Systems and Applications,* 41-50.

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

APPENDICES

Appendix 1: Whole questionnaire with all (24) questions and answer options.

Appendix 2: Summary of survey data collection

Appendix 3: Distribution list

Appendix 1: Whole questionnaire with all (24) questions and answer options.

Compulsory questions are marked with an asterisk (*).

1. Based on your opinion, estimate the proportion of published studies with misidentified species in all ecological fields of science? *

0 %

0% — 100%

☐ I do not know

2. Do you consider species misidentification as problematic in relation to your own research? *

- ○ Not at all (such errors do not change the interpretation of my results)
- ○ A little
- ○ Somewhat
- ○ Very
- ○ Extremely (such errors would have serious impacts on the implementation of my findings)
- ○ I have not thought about this before

3. Have you discussed about species misidentification (or the possibility of it) with your colleagues or students? *

- ○ Never (identification errors are not that relevant to my work)
- ○ Rarely
- ○ Sometimes
- ○ Often
- ○ Regularly (the accuracy of identification is considered in research projects that have species data)
- ○ I do not remember having such discussions

4. While you read or review papers, how often you consider the possibility of species misidentification? *

- Never (I assume that there are no errors, or they do not matter)
- Rarely
- Sometimes
- Often
- Every time (in all occasions when the paper has species data)
- I do not remember thinking about this

5. How often have reviewers asked or commented about your species identification methods or their accuracy? *

- Never (not once)
- Rarely
- Sometimes
- Often
- Every time (when my manuscripts have species data)
- I do not remember such happening

6. Have you noticed species misidentification in your own research? *

Please consider all stages of research from data collection to published papers.

- Yes
- No
- I do not remember

7. How frequent species misidentification has been in your research?

As research projects vary in their size, we wanted to come up with a comparable measure. Thus, we ask you to answer this question in terms of papers produced. E.g., if there is one species misidentified in a data set that is used for several publications, answer according to the number of publications. Please, include also manuscripts in preparation and identification errors that were corrected. How many of your manuscripts/papers have included data with misidentified species (at some point of the project)?

- 1-2 of my manuscripts
- 3-5 of my manuscripts
- 6-15 of my manuscripts
- 16-30 of my manuscripts
- over 30 of my manuscripts
- all my manuscripts

8. At what stage of your study species misidentification(s) was observed? Please indicate the error frequency among different stages.

- If you have not encountered species misidentification in a given stage of research process, choose "0%" for that stage.
- If you have encountered misidentification only in a given stage, choose "100%" for that stage and "0%" for others.
- If you have encountered misidentifications in several stages, choose the percentages according to your experience.

	0%	20%	40%	60%	80%	100%
a) Specimen collection (field, greenhouse etc.)	○	○	○	○	○	○
b) Specimen handling (laboratory work, microscoping, video analysis etc.)	○	○	○	○	○	○
c) Data analysis	○	○	○	○	○	○
d) Writing (reporting stage)	○	○	○	○	○	○
e) Revision	○	○	○	○	○	○
f) After publishing	○	○	○	○	○	○

9. Did you fix the error(s), and if yes, how?

10. How many peer-reviewed journal articles have you published? *

Select

Select
1-10
11-20
21-30
31-40
41-50
51-100
Over 100

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

11. Do you report the possibility of species misidentification, e.g. through error rates, in your papers?

- No - it is not necessary
- Only rarely, if ever
- Sometimes
- Usually yes
- Every time - when there is species data
- I have not ever thought about that!

12. Do you use some specific methods to ensure that your or your students' species identifications are correct?

13. Do you have some final comments, anecdotes, or stories regarding the topic or any of the questions in this survey?

14. Gender

Select

- Select
- Female
- Male
- Other

15. Where are you from?

I did my doctoral studies in (which country)

At this moment, I am working in (which country)

16. How many full years since receiving your doctoral degree?

Select

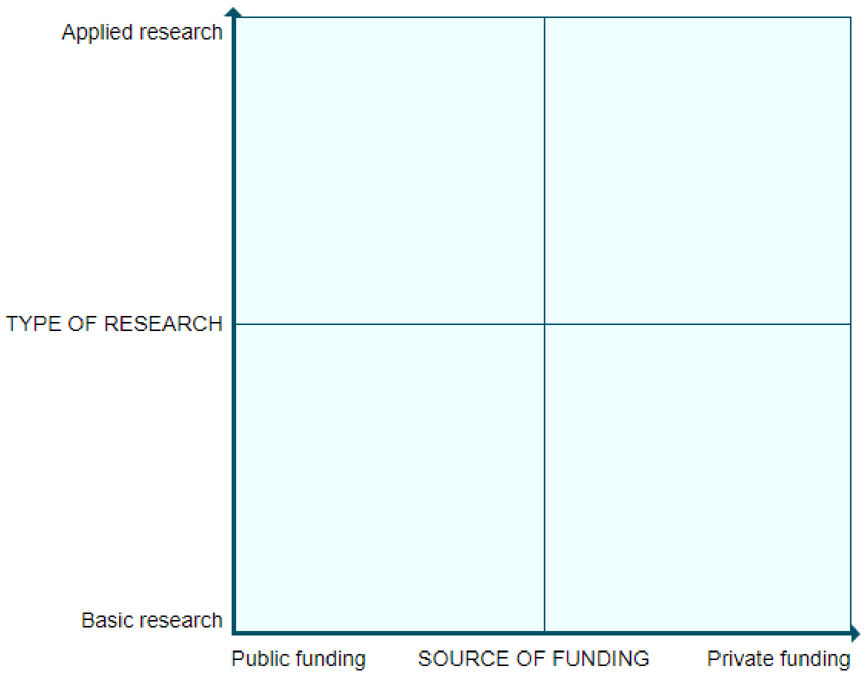
Select
0-4
5-9
10-14
15 or more

17. Are you leading your own research group?

- Yes
- No

18. What kind of research environment do you work in? Click on the graph.

Clear all Clear selected



19. What is your field(s) of science? Choose the most relevant option(s).

- ☐ Applied ecology (agroecology, forest ecology etc.)
- ☐ Aquatic sciences
- ☐ Autoecology (single-species studies)
- ☐ Bioinformatics
- ☐ Citizen science
- ☐ Community ecology
- ☐ Conservation science
- ☐ Ecophysiology
- ☐ Ecosystem ecology
- ☐ Evolutionary biology
- ☐ Landscape ecology
- ☐ Microbiology
- ☐ Population ecology
- ☐ Population genetics
- ☐ Taxonomy

20. Other, what?

My field of research is... ____________________

21. With which of the following taxa do you (mainly) work with?

- ☐ Archaea
- ☐ Bacteria
- ☐ Single-celled eukaryotes
- ☐ Fungi
- ☐ Nonvascular plants
- ☐ Vascular plants
- ☐ Invertebrate animals
- ☐ Vertebrate animals

22. If applicable, please specify the taxa you mainly work with.

The taxa I mainly work with are... [ ]

23. How many species your research usually focuses on?

Select

- Select
- 1
- 2-5
- 6-10
- 11-100
- 100-200
- More than 200

24. Where did you hear about this survey?

- [ ] Social media
- [ ] Group e-mail (via a mailing list)
- [ ] From a friend or a colleague (personal contact)
- [ ] Other

Appendix 2: Summary of survey data collection

Figures numbers corresponding to question numbers. Summary of free-text questions are not shown to protect respondents' background information. Only those which were not included in the Results or Discussion sections are presented. "n" indicates the number of respondents who answered a specific question.

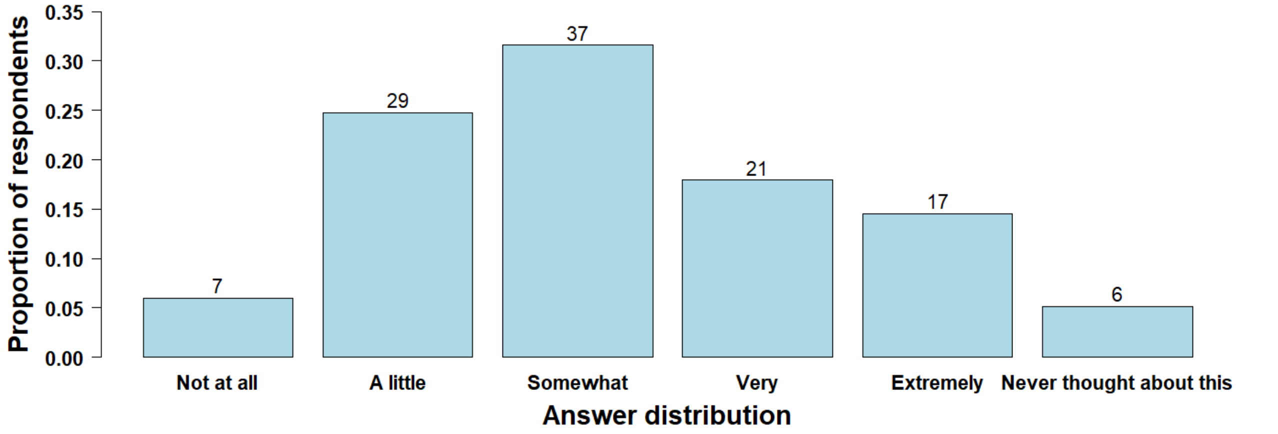


Figure 2 How problematic respondents consider species misidentification in relation to their own research (n = 117).

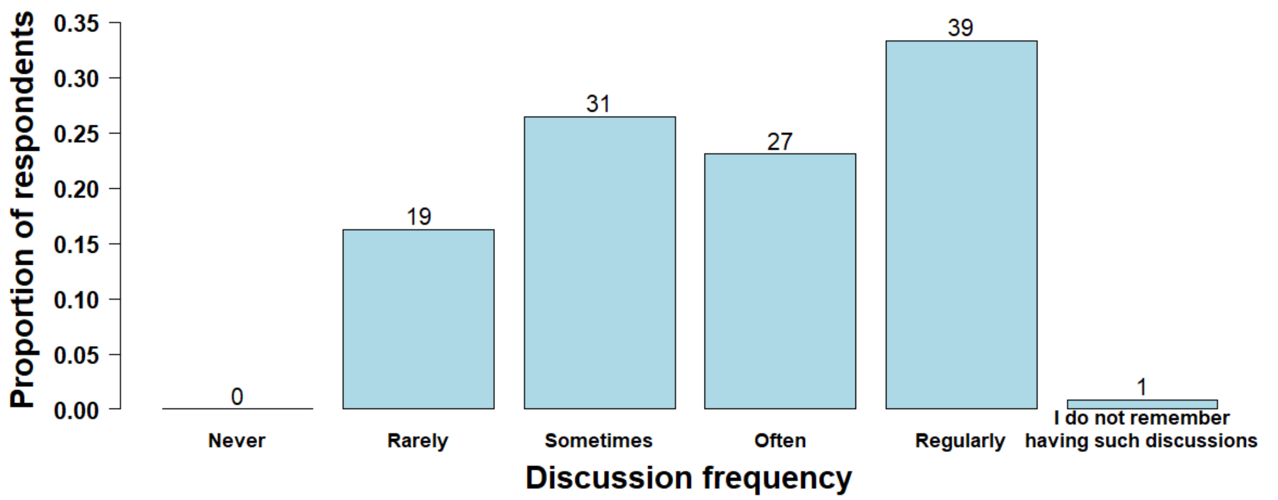


Figure 3 Frequency of discussing species misidentification (or the possibility of it) with colleagues or students (n = 117).

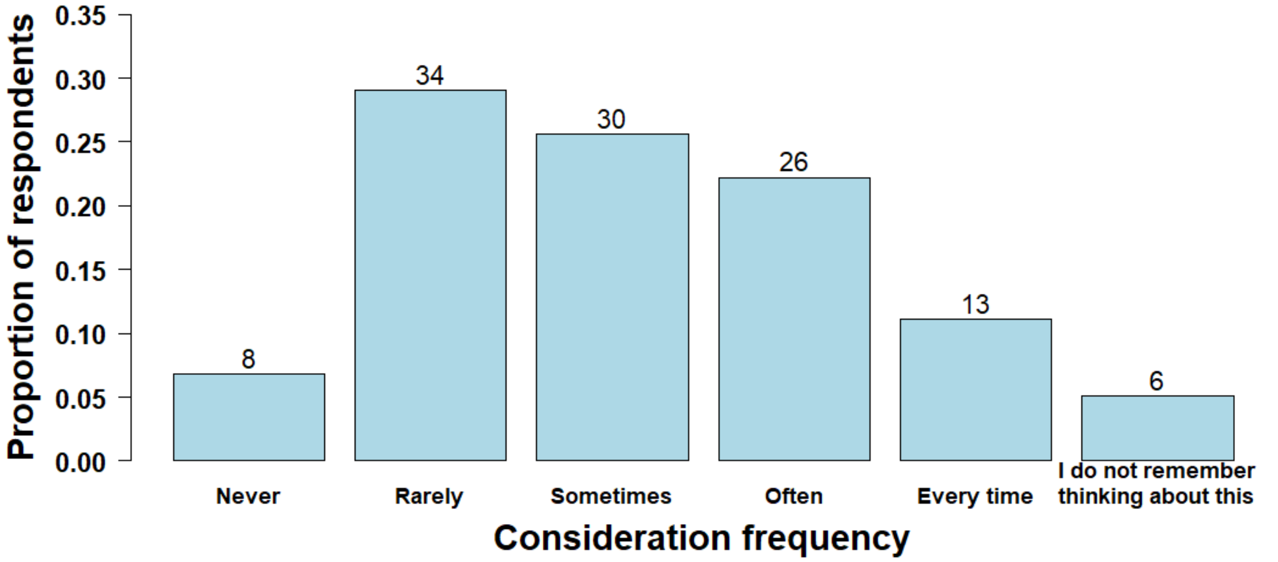


Figure 4 Frequency of considering the possibility of species misidentification while reading or reviewing papers (n = 117).

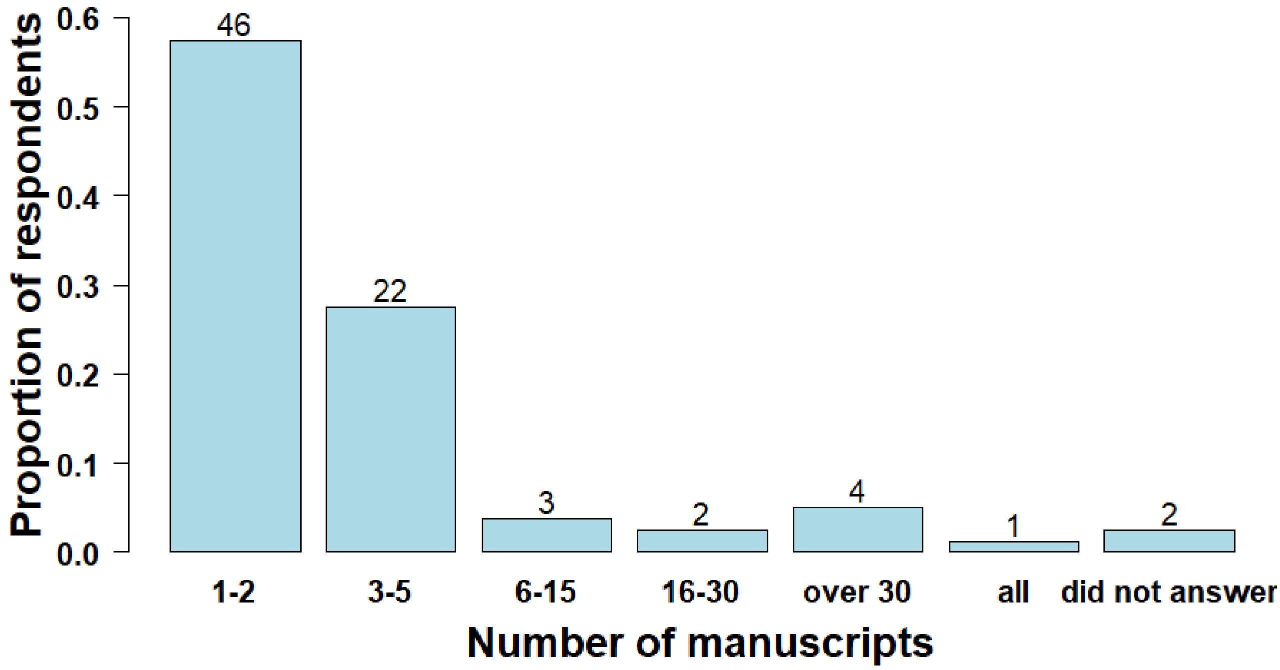


Figure 7 Frequency of species misidentification in terms of papers produced (n = 80). Manuscripts in preparation and identification errors that were later corrected are also included.

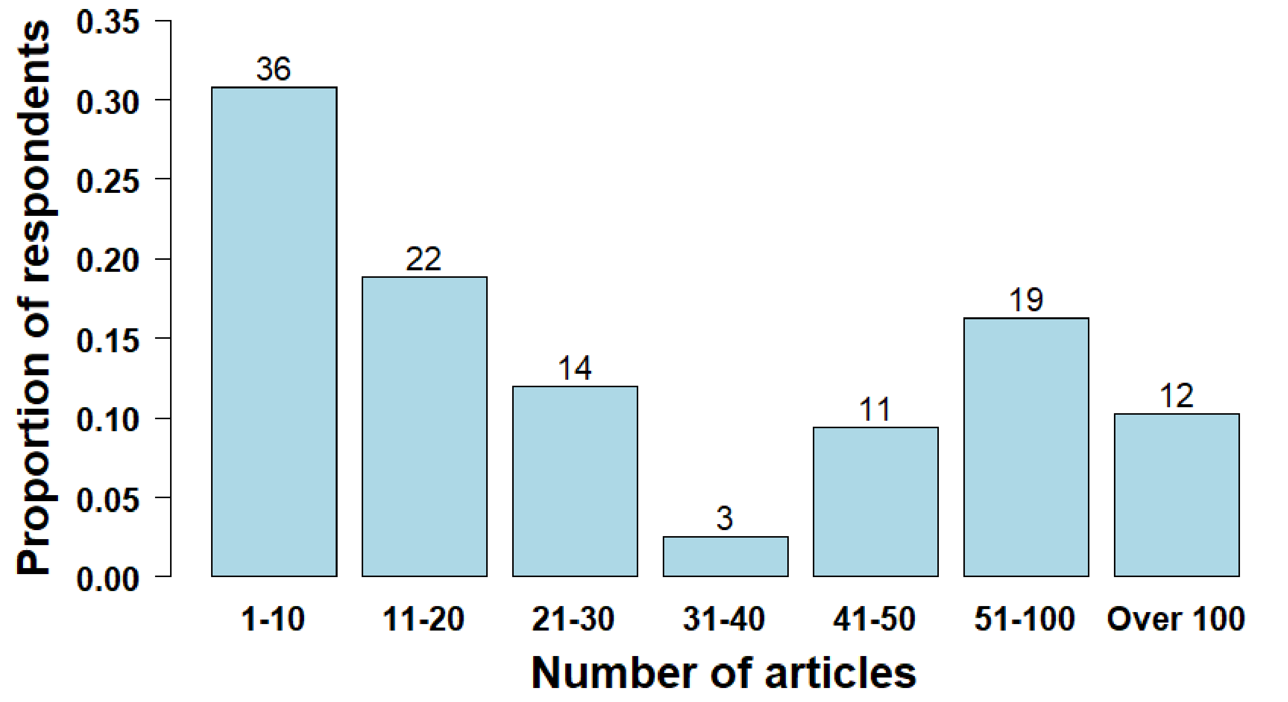


Figure 10 Number of articles respondents have published in peer-reviewed journal (n = 117).

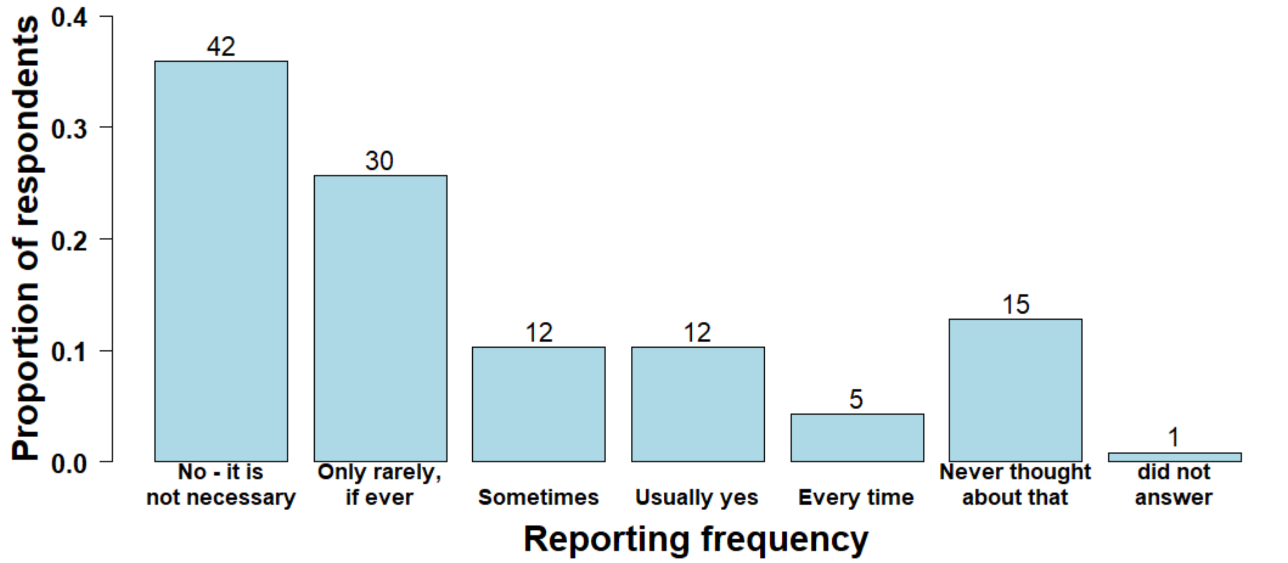


Figure 11 Respondents' estimated frequency of reporting the possibility of species misidentification (n = 117).

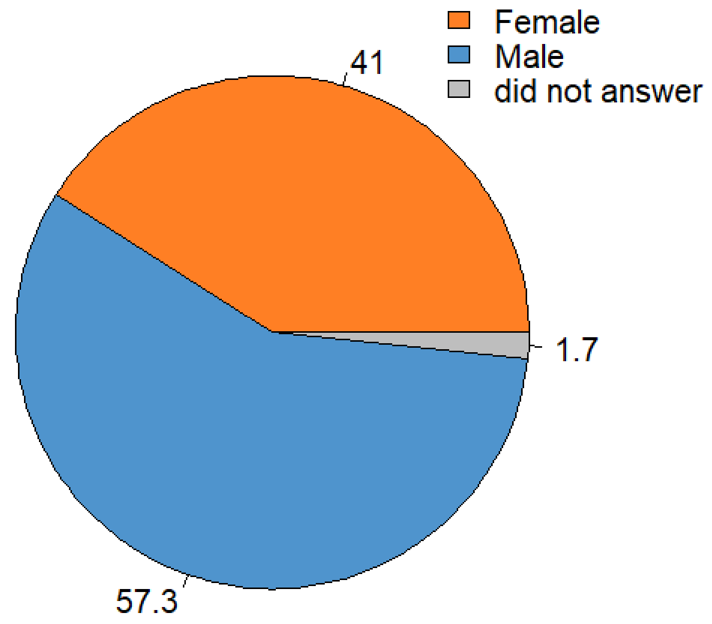


Figure 14 Gender of respondents (n = 117).

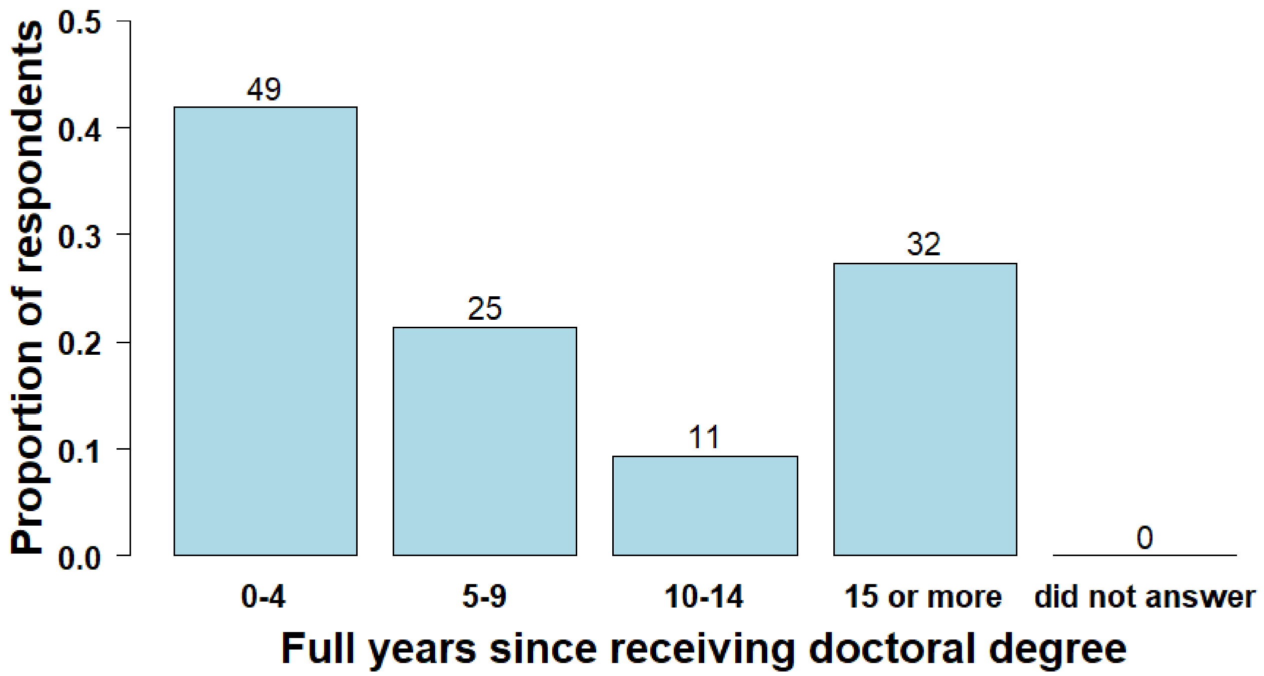


Figure 16 Full years since respondents receiving a doctoral degree (n = 117).

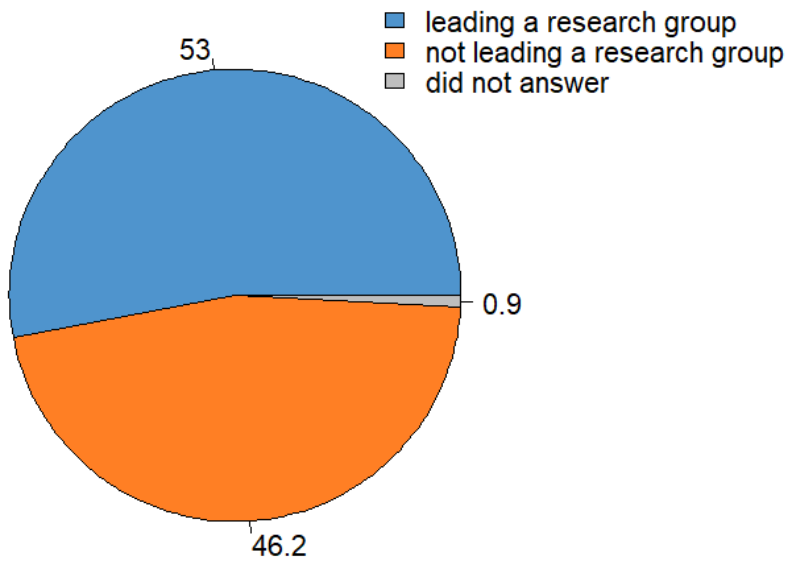


Figure 17 Percentage of respondents who are leading their own research group (n = 117).

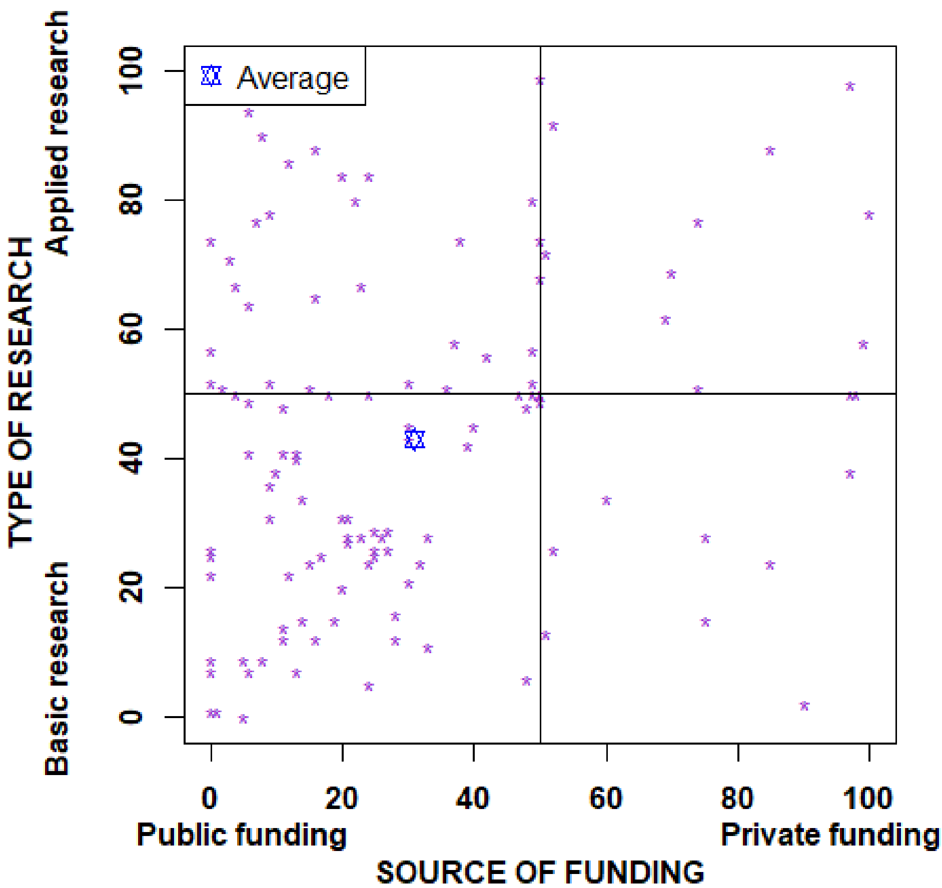


Figure 18 Research environment (n = 117).

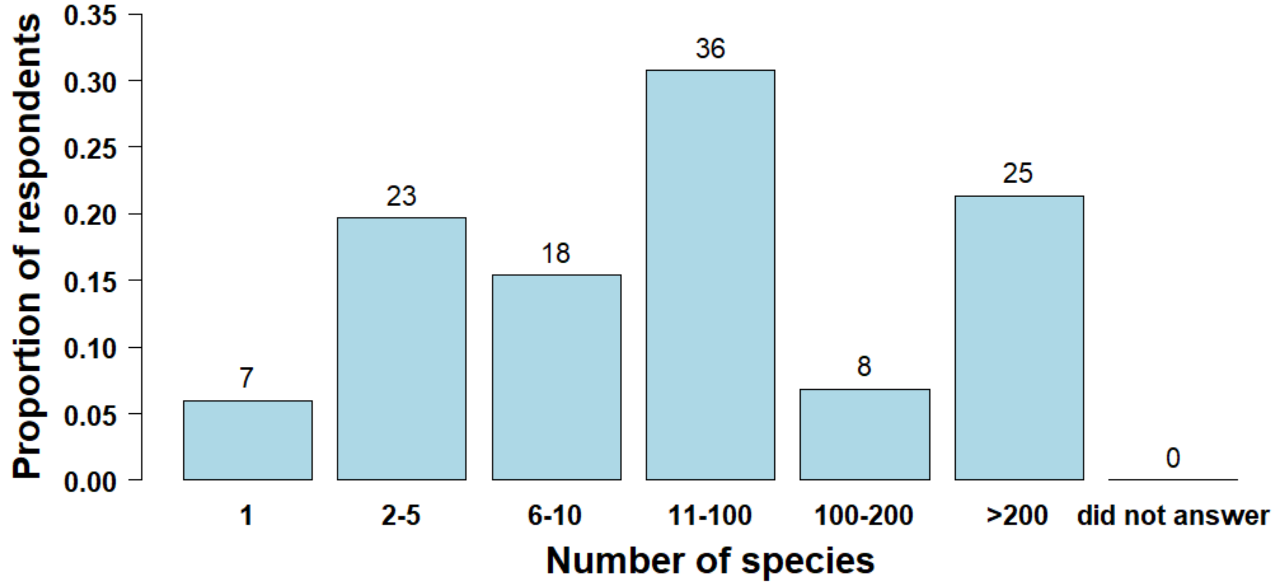


Figure 23 Number of species that respondents usually focus on in their research (n = 117)

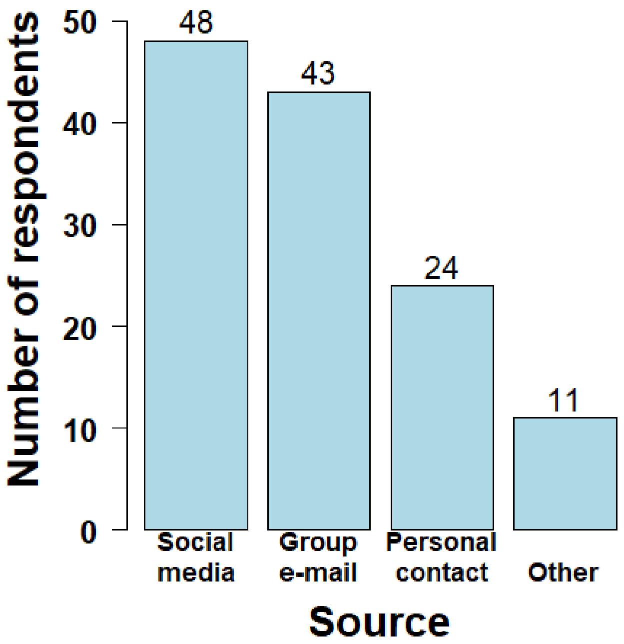


Figure 24 The source where respondents heard about this survey (n = 117)

Appendix 3: Distribution list

Society/ institution	Continent	Country	Field
AEET (Asociacion Espanola de Ecologia Terrestre)	Europe	Spain	terrerstrial ecology
AEHMS (Aquatic Ecosystem Health & Management Society)	International	International	the protection and conservation of global aquatic resources
AOS (American ornithological society)	International	International	ornithology
ASAE (Asociacion Argentina de Ecologia)	South America	Argentina	all (general ecology)
ASFB (Australian Society for Fish Biology)	Oceania	Australia	research, education and management of fish and fisheries
ASLE (Association for the Study of Literature and Environment)	International	International	biodiversity, environmental studies, climate change
ASN (American society of naturalists)	International	International	ecology, evolution, and behavior
AWDF (African Wildlife Defence Force)	Africa	some African countries	wildlife conservation and rainforest conservation
Beacon Institute	North America	USA	estuary and freshwater ecosystems
BEI (Balanced Ecology)	North America	USA	wildlife and habitat conservation, and a sustainable coexistence

BES (British ecological society)	Europe	British	all (general ecology)
biodivcanada	North America	Canada	biodiversity, conservation
BSA (Botanical society of America)	International	International	plant
California State Coastal Conservancy	North America	USA	coastal resources conservation
Cary Institute of Ecosystem Studies	North America	USA	ecosystem studies: disease ecology, freshwater, forests, urban ecology
Census of Marine Life	International	International	diversity (how many different kinds), distribution (where they live), and abundance (how many) of marine life
CICHAZ (Centro de Investigaciones Cientifícas de las Huastecas 'Aguazarca')	North America	Mexico	a range of disciplines in the natural and social sciences from political science to behavioral ecology and physiology.
CIEEM (Chartered Institute of Ecology and Environmental Management)	Europe	UK, Ireland	all (general ecology)
CSEE (Canadian Society for Ecology and Evolution)	North America	Canada	all (general ecology)
DICE (Durrell Institute of Conservation and Ecology)	Europe	UK	conservation, biodiversity

E3G(Third Generation Environmentalism)	Europe	UK, Belgium, Germany	climate change
Ecocity Builders	North America	USA	urban ecology
EEF (European ecological federation)	Europe	European countries	all (general ecology)
ESA (Ecological Society of Australia)	Oceania	Australia	all (general ecology)
ESC (Ecological Society of China)	Asia	China	all (general ecology)
ESEB (European society for evolutionary biology)	Europe	European countries	evolutionary biology
ESJ (Ecological Society of Japan)	Asia	Japan	all (general ecology)
ESK (Ecological Society of Korea)	Asia	Korea	all (general ecology)
EuroNatur Foundation	Europe	European countries	conservation
EWA (Edison Wetlands Association)	North America	USA	conservation
Freeland	International	International	wildlife and ecosystems conservation
FSC (Field Studies Council)	Europe	UK	environmental sciences
GBIF (Global Biodiversity Information Facility)	International	International	biodiversity

GfÖ (Ecological Society of Germany, Austria and Switzerland)	Europe	Germany, Austria and Switzerland	all (general ecology)
GLEON (Global Lake Ecological Observatory Network)	International	International	lake ecology
HES (Hungarian Ecological Society)	Europe	Hungary	all (general ecology)
I.E. (Institute of Ecotechnics)	International	International	ecotechnology, the environment, conservation, and heritage
IALE (International Association of Landscape Ecology)	International	International	landscape ecology
IAPT (International association for plant taxonomy)	International	International	taxonomy and systematics of algae, fungi, and plants
iDigBio (Integrated Digitized Biocollections)	North America	USA	digitized information about vouchered natural history collections
ILTER (International Long Term Ecological Research)	International	International	long-term, site-based ecological and socio-economic research relating to environmental change
INTECOL (International association for ecology)	International	International	all (general ecology)
IRF (International Ranger Federation)	International	International	the practical protection and preservation of all aspects of wild areas, historical and cultural sites

ISBE (International society for behavior ecology)	International	International	behavioral ecology
ISEE (International Society for Ecological Economists)	International	International	ecological economics
ISEM (International Society for Ecological Modelling)	International	International	the application of systems analysis and simulation in ecology and natural resource management
ISME (International Society for Microbial Ecology)	International	International	microbial Ecology
IUBS (International Union of Biological Sciences)	International	International	all (general ecology)
IUSSI (International Union for the Study of Social Insects)	International	International	social insects and other social organisms in the broadest sense
JRBP (Jasper Ridge Biological Preserve)	North America	USA	conservation
Local Futures	International	International	environmental-economic study
LTER (Long Term Ecological Research Network)	North America	USA	all (general ecology)
MBA (The marine biological association)	International	International	marine biological research
MES (Macedonian Ecological Society)	Europe	Macedonia	all (general ecology)
NBN (National Biodiversity Network)	Europe	UK	biodiversity

NCEAS (National Center for Ecological Analysis and Synthesis)	North America	USA	environmental Science
NEON (National Ecological Observatory Network)	North America	USA	ecosystem
NIE (National Institute of Ecology (India)	Asia	India	all (general ecology)
NSO (Nordic society oikos)	Europe	Nordic countries	all (general ecology)
NZES (New Zealand Ecological Society)	Oceania	New Zealand	all (general ecology)
OBFS (Organization of Biological Field Stations)	International	International	all (general ecology)
Ocean Elders	International	International	ocean conservation, ocean’s habitat and wildlife
RCEES (Research Center for Eco-Environmental Sciences)	Asia	China	environmental chemistry, environmental science, and systems ecology
rECOrd (Local Biological Records Centre)	Europe	UK	all (general ecology)
RES (Romanian Ecological Society)	Europe	Romania	all (general ecology)
RMBL (Rocky Mountain Biological Laboratory)	North America	USA	pollination, changing climate and high-altitude ecosystems
S.It.E. (Italian Society of Ecology)	Europe	Italy	all (general ecology)

SAFER (SubAntarctic Foundation for Ecosystems Research)	Oceania	New Zealand	wildlife research and ecological restoration in the Falkland Islands
SCOPE (Scientific Committee on Problems of the Environment)	International	International	environmental science
SETAC (Society of Environmental Toxicology and Chemistry)	International	International	sustainable environmental quality and ecosystem integrity
SFE (Societe Francaise d'Ecologie)	Europe	French	all (general ecology)
SHE (Society for Human Ecology)	International	International	human ecology
SOCECOL (Sociedad de Ecologia de Chile)	South America	Chile	all (general ecology)
SURE (Society for urban ecology)	International	International	urban ecology
Swedish Oikos Society	Europe	Sweden	terrerstrial ecology
The Linnean Society of London	International	International	all (general ecology)
The Reef Ball Foundation	International	International	ocean ecosystem
TWI (The Wetlands Initiative)	North America	USA	wetland restoration ecology
UFZ (Helmholtz Centre for Environmental Research)	Europe	Germany	environmental research- biodiversity, functioning ecosystems, clean water and intact soils

UNDERC (University of Notre Dame Environmental Research Center)	North America	USA	environmental research
Vlaamse Vereniging voor Ecologie	Europe	Netherlands	all (general ecology)
WCFS (World Council of Fisheries Societies)	International	International	fisheries science, conservation and management
Wildlife Conservation Society (including societies of 33 countries)	International	International	wildlife conservation
WWF (World Wide Fund for Nature)	International	International	conservation, climate change

Continent	Country	University	Faculty
Africa	Egypt	Cairo University	Faculty of Agriculture
	Ethiopia	Addis Ababa University	College of Natural Sciences; Department of Animal Science
	South Africa	University of Cape Town	Department of Biological Sciences; Department of Environmental & Geographical Science; Department of Oceanography; all researchers in related fields
		University of Witwatersrand	Departments of Botany, Zoology and Environmental Sciences; School of Molecular & Cell Biology
		Stellenbosch University	Faculty of AgriSciences; Departments of Botany and Zoology; Department of Earth Sciences; Department of Microbiology
	Uganda	Makerere University	College of Agricultural and Environmental Sciences; College of Natural Sciences; College of Veterinary Medicine; Animal Resources & Bio-security
Asia	China	Fudan University	all researchers in related fields
		Peking University	School of Life Sciences; College of urban and environmental sciences
		The Chinese University of Hong Kong	School of Life Sciences

	Tsinghua University	School of Life Sciences; all researchers in related fields
	University of Science and Technology of China	School of Life Sciences
Japan	Osaka University	Department of Biological Sciences; all researchers in related fields
	The University of Tokyo	Department of Biological Sciences; Faculty of Agriculture
	Tohoku University	Faculty of Agriculture; Graduate School of Life Sciences; Graduate School of Environmental Studies
Saudi Arabia	King Abdulaziz University	Biological sciences department; Faculty of Meteorology; Environment and Arid Land Agriculture
Singapore	Nanyang Technological University	School of Biological Sciences; Asian School of the Environment
	National University of Singapore	Department of Biological Sciences; Lee Kong Chian Natural History Museum; all researchers in related fields
South Korea	Pohang University of Science and Technology	Department of Life Sciences; Division of Environmental Science & Engineering; School of Interdisciplinary Bioscience and Bioengineering (I-BIO)
	Seoul National University	College of Natural Sciences; College of Agriculture & Life Sciences

		Yonsei University	College of Life Science and Biotechnology
Europe	Denmark	Aarhus University	Department of Bioscience; Department of Molecular Biology and Genetics
		Technical University of Denmark	National Institute of Aquatic Resources; DTU Environment
		University of Copenhagen	Department of Biology; Department of Plant and Environmental Sciences; Natural History Museum of Denmark; all researchers in related fields
	Germany	University of Munich	all researchers in related fields
	Italy	Sapienza University of Rome	Department of Biology and Biotechnology
		University of Bologna	Agricultural and Food Sciences - DISTAL; Biological, Geological, and Environmental Sciences - BiGeA
		University of Milan	Faculty of Agricultural and Food Sciences; Faculty of Science and Technology
	Russia	M. V. Lomonosov Moscow State University	Faculty of Biology; Faculty of Biotechnology; Faculty of Bioengineering and Bioinformatics
	UK	Imperial College London	Department of Life Sciences; Centre for Environmental Policy

		University of Cambridge	School of the Biological Sciences; Faculty of Biology
		University of Edinburgh	School of Biological Sciences
		University of Manchester	Faculty of Biology, Medicine and Health
		University of Oxford	Department of Plant Sciences; Department of Zoology; Museum of Natural History; Botanic Garden and Arboretum
North America	Canada	McGill University	Faculty of Agricultural and Environmental Sciences; Department of Biology; all researchers in related fields
		University of British Columbia	Department of Biology; all researchers in related fields
		University of Toronto	Department of Ecology and Evolutionary Biology; School of Environment; Department of Cell & Systems Biology; all researchers in related fields
		Dalhousie University, McMaster University, Queen's University, University of Alberta, University of Calgary, University of Manitoba, University of Ottawa, University of Saskatchewan, Western University	all researchers in related fields
	USA	Harvard University	Department of Organismic and Evolutionary Biology

		Stanford University	Earth System Science Department; Department of Biology
		University of California--Berkeley	Department of Plant & Microbial Biology; Department of Environmental Science, Policy, and Management; Museum of Paleontology; Museum of Vertebrate Zoology; Department of Integrative Biology
Oceania	Australia	Australian National University	Research School of Biology; Research School of Earth Sciences
		University of Melbourne	School of BioSciences; School of Ecosystem and Forest Sciences
		University of Sydney	School of Life and Environmental Sciences; all researchers in related fields
	New Zealand	University of Auckland	School of Biological Sciences; School of Environment; Institute of Marine Science
		University of Canterbury	School of Biological Sciences; School of Earth and Environment
		University of Otago	Department of Botany; Department of Marine Science; Department of Zoology
South America	Brazil	Universidade de São Paulo	Museum of Zoology; Oceanographic Institute; Institute of Biology; College of Agriculture
	Chile	Pontificia University Católica de Chile	all researchers in related fields