

## RESEARCH ARTICLE

## Using DNA barcoding to improve invasive pest identification at U.S. ports-of-entry

Mary J. L. Madden<sup>1</sup>, Robert G. Young<sup>1\*</sup>, John W. Brown<sup>2</sup>, Scott E. Miller<sup>2</sup>, Andrew J. Frewin<sup>1</sup>, Robert H. Hanner<sup>1</sup>**1** Department of Integrated Biology, University of Guelph, Guelph, Ontario, Canada, **2** Entomology Department, National Museum of Natural History, Smithsonian Institution, Washington, D.C., United States of America\* [ryoung04@uoguelph.ca](mailto:ryoung04@uoguelph.ca)

## Abstract

Interception of potential invasive species at ports-of-entry is essential for effective biosecurity and biosurveillance programs. However, taxonomic assessment of the immature stages of most arthropods is challenging; characters for identification are often dependent on adult morphology and reproductive structures. This study aims to strengthen the identification of such specimens through DNA barcoding, with a focus on microlepidoptera. A sample of 241 primarily immature microlepidoptera specimens intercepted at U.S. ports-of-entry from 2007 to 2011 were selected for analysis. From this sample, 201 COI-5P sequences were generated and analyzed for concordance between morphology-based and DNA-based identifications. The retrospective analysis of the data over 10 years (2009 to 2019) using the Barcode of Life Data (BOLD) system demonstrates the importance of establishing and growing DNA barcode reference libraries for use in specimen identification. Additionally, analysis of specimen identification using public data (43.3% specimens identified) vs. non-public data (78.6% specimens identified) highlights the need to encourage researchers to make data publicly accessible. DNA barcoding surpassed morphological identification with 42.3% (public) and 66.7% (non-public) of the sampled specimens achieving a species-level identification, compared to 38.3% species-level identification by morphology. Whilst DNA barcoding was not able to identify all specimens in our dataset, its incorporation into border security programs as an adjunct to morphological identification can provide secondary lines of evidence and lower taxonomic resolution in many cases. Furthermore, with increased globalization, database records need to be clearly annotated for suspected specimen origin versus interception location.

## OPEN ACCESS

**Citation:** Madden MJL, Young RG, Brown JW, Miller SE, Frewin AJ, Hanner RH (2019) Using DNA barcoding to improve invasive pest identification at U.S. ports-of-entry. PLoS ONE 14(9): e0222291. <https://doi.org/10.1371/journal.pone.0222291>

**Editor:** Massimo Labra, Università degli Studi di Milano-Bicocca, ITALY

**Received:** May 23, 2019

**Accepted:** August 26, 2019

**Published:** September 17, 2019

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** Data are currently available on the BOLD system under the dataset DS-ITLP. The DOI for this data is [doi.org/10.5883/DS-ITLP](https://doi.org/10.5883/DS-ITLP).

**Funding:** This project was funded in part through Growing Forward, a federal provincial-territorial initiative with funding awarded to RH. The Agricultural Adaptation Council assists in the delivery of Growing Forward programs in Ontario. Additional funding was provided by Flowers Canada (Ontario) Inc. and the Ontario Ministry of Agriculture and Food, and Rural Affairs (OMAFRA)

## Introduction

Invasive insects are of global concern, causing long-term environmental impacts including reduced ecosystem stability and loss of native species [1]. Direct damage to commodities and management activities focused on invasive insect pests cost the global economy billions of dollars each year [2–4]. With the ever-increasing global movement of commodities and

to RH. The sponsors did not play a role in the study design, data collection and analysis, decision to publish, or preparation of this manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

intensifying effects of global warming, there is little doubt that the frequency of insect invasions will increase [5–8]. To reduce the establishment of invasive species, border inspection programs intercept and identify potential invasive pests and are an integral part of effective biosecurity policies. Data from border inspection programs inform policies regarding import restrictions and guidelines and can trigger biosecurity actions such as trade sanctions and commodity decontamination [9,10].

Border inspection data can also be used in pathway analysis of invasive species, increasing regulators' ability to understand common routes of insect invasions [10]. However, Kenis *et al.* 2007 [11] concluded that only 10% of the invasive species with confirmed populations in non-native ranges were detected at ports-of-entry prior to their establishment, clearly demonstrating a need to improve biosecurity and biosurveillance policies. Strategies to identify introduction pathways through specimen detection and identification have been developed for several countries including the United Kingdom [12,13], Australia [14], and Finland [15]. Ultimately, the effectiveness of such programs relies on the ability to intercept specimens in transit and quickly identify them to a taxonomic rank necessary to inform biosecurity actions and policies.

Microlepidoptera are a paraphyletic grade of Lepidoptera, generally described as 'small moths' [16]. Identifications of microlepidoptera from taxonomic ranks of family to species, relies on morphological characters of the adult, primarily those of the genitalia [17, 18]. Because most microlepidoptera are intercepted as larvae, they pose a special challenge for accurate and timely identifications [3, 19]. Although critical evaluations on the reproducibility of morphology-based taxon assignments are limited, subjective interpretations occur, resulting in inconsistent taxonomic assignment of specimens by different investigators, especially at lower taxonomic ranks [17, 20].

Many border-intercepted specimens are identified at higher ranks (e.g., family or subfamily) due to the challenges of classifying specimens of immature life stages [21]. This is illustrated by the fact that only 40% of the specimens in a USDA interception dataset, received species-level identifications. While taxonomic identifications above species are useful in some investigations, regulatory management of invasive biological organisms typically requires species-specific information such as life history, larval hosts, parasitoids and dispersal ability [22, 23].

A potential solution to the challenge of morphology-based identifications for border interception programs is the integration of DNA barcoding. DNA barcoding is a standardized molecular identification method with numerous applications that has been used extensively to identify immature life stages of animals [24–31]. The Barcode of Life Data Systems (BOLD) is a publicly accessible domain providing a reference library and analytical capabilities for DNA barcode projects [32]. One strength of the BOLD system is in the association of records with metadata including chromatogram files, the location of voucher specimens, and the presence and location of long-term storage of DNA or tissue [33, 34]. These data are only useful when they are publicly accessible to external users, allowing for verification of completeness and accuracy of records. Additionally, BOLD periodically mines sequences from GenBank, increasing the number of sequences available for data analysis [35, 36].

An important feature of the BOLD platform is the barcode index number (BIN). BINs are molecular operational taxonomic units [37] (MOTU) generated by the Refined Single Linkage (RESL) algorithm [38], based on available BOLD data. BINs provide interim taxonomic identifications or suspected species classification, based on a molecular barcode [38]. The BIN framework is most suitable for microlepidoptera identification as it has been shown to be in general agreement with morphology-based species-level identifications [38, 39]. Furthermore, when the RESL algorithm was being developed it was extensively tested against a large

lepidoptera dataset [38]. The capability of DNA barcoding to aid in the identification of specimens can be measured by its success in providing taxonomic resolution equivalent to or better than that achieved by traditional methods. This can include, specimens that can be identified to the species level or specimens that are grouped into an interim taxonomic framework (i.e., BINs).

To test the use of DNA barcoding in the identification of insect specimens, we examined a set of predominantly immature microlepidoptera from the superfamilies Tortricoidea and Gelechioidea. Both Tortricoidea and Gelechioidea are diverse, containing numerous regulated and economically important species, many of which are represented in BOLD [25]. The first objective of this study is to examine the concordance between morphology-based and DNA-based identifications for intercepted microlepidoptera. The second objective is to develop a framework for the use of DNA barcoding and BIN interim taxonomy with respect to border identification protocols for intercepted insect specimens.

We obtained DNA barcodes from 201 microlepidoptera specimens intercepted by USDA Animal and Plant Health Inspection Service Plant Protection and Quarantine (USDA-APHIS-PPQ) inspectors at U.S. ports-of-entry and identified by the USDA Systematic Entomology Laboratory (USDA-SEL). Morphological identifications were then compared to the DNA barcode-based identifications and BIN assignments given to each specimen by BOLD. Over time, the number of specimens in the interception dataset matching a sequence in the BOLD database increased. However, a lack of data movement from private to public records was observed. Through these analyses we demonstrate a need for continued population of molecular libraries with good quality, public data to increase the strength of DNA barcode identifications.

## Materials and methods

### Sequencing

USDA APHIS-PPQ personnel routinely inspect incoming commodities at U.S. ports-of-entry (including Puerto Rico) for plant and animal pests. Unidentified insect specimens were preserved in 75% alcohol (immatures) or pinned (adults) and submitted to the USDA-SEL for identification. Interceptions from commercial commodities are labelled “urgent” and are sent by overnight courier to USDA-SEL for rapid turnaround (i.e., identification the same day the specimen(s) is received by USDA-SEL). Non-critical, non-commercial interceptions labelled “routines,” are saved at the port and periodically sent to USDA-SEL for evaluation.

We selected a sample of microlepidoptera specimens (241) intercepted by APHIS-PPQ consisting mainly of larvae, but with a few adult specimens. The specimens were identified by USDA-SEL staff located in the Entomology Department of the National Museum of Natural History at the Smithsonian Institution in Washington, DC. USDA-SEL staff rely on origin of interception, host plant or other association, and historical records to augment morphology when identifying samples. The collection was assembled with an effort to include a broad range of frequently intercepted microlepidopteran families: Tineidae, Blastobasidae, Cosmopterigidae, Gelechiidae, Oecophoridae, Tortricidae, Plutellidae, Pterophoridae, and Sesiidae, and a few specimens of unknown familial affinity. Nonetheless, Tortricoidea and Gelechioidea dominated the samples because larvae of these two superfamilies are the most commonly intercepted and submitted samples of microlepidoptera.

For adult specimens a single hind leg was used as a source of genomic DNA. For immature specimens (larvae or pupae) > 5 mm in length, DNA was obtained from a 2–3 mm<sup>2</sup> piece of tissue removed from the lateral side of the specimen with flame-sterilized forceps and scissors, while specimen < 5 mm in length had their entire bodies homogenized to extract DNA.

Genomic DNA was extracted following the alkaline lysis DNA extraction XytXtract Insect (ANDE) (Xytogen; Perth, Australia) kit using manufacturer recommended protocols [40] and stored at -20°C prior to analysis.

All PCR reactions were conducted in a total volume of 12.5µL, as described by Wilson et al. (2012) [41] as modified from Ivanova et al. (2009) [42]. PCR amplification was completed using primer pairs LepF1, 5'-ATTCAACCAATCATAAAGATAT-3'; and LepR1, 5'-TAAACTTCTGGATGTCCAAAAA-3' [43]. Reactions that failed to produce sufficient DNA products for sequencing underwent a second amplification using primer pairs mLepF1/LepR1, and LepF1/mLepR1 as described by Wilson et al. (2012) [41]. The following cycling conditions were used for all primer pairs: an initial denaturation at 94°C for 1min, followed by 5 cycles of 94°C for 40s, 45°C for 40s, 72°C for 1min, followed by 35 cycles of 94°C for 40s, 51°C for 40s, 72°C for 1min, and a final extension at 72°C for 5min [44]. PCR products were visualized on a 2% agarose gel pre-stained with SYBR® Safe DNA gel stain (Life Technologies™). PCR products were sent to the Advanced Analysis Centre at the University of Guelph (Guelph, Ontario, Canada) for sequencing. Sequencing was performed on an Applied Biosystems® 3730 DNA Analyzer. Specimen metadata, photographs, trace-files, and DNA barcode sequences were deposited in BOLD project ITLP. Complete specimen data for these individuals can be found on BOLD under the project ITLP ([doi.org/10.5883/DS-ITLP](https://doi.org/10.5883/DS-ITLP)). All specimens in project ITLP were intercepted in the USA. Data on assumed countries of origin was provided by APHIS-PPQ staff, based on the best information available.

### Identification using BOLD

Specimens with sequence data were analyzed using a molecular barcode approach and BOLD. To observe the growth of data in BOLD and how it can improve the ability of the system to identify specimens a comparative past and present analysis was completed. The retrospective analysis was conducted using accumulated data for the years 2009 through 2015. The current BOLD system data, as of the date this study was also used (i.e., 2019). Specimen identifications were conducted using the BOLD Species Level Barcode Record (SLBR) and the Public Record Barcode Database (PRBD) options. The BOLD system defines the SLBR option as “every COI barcode record with a species level identification and a minimum sequence length of 500bp. This includes many species represented by only one or two specimens as well as all species with interim taxonomy.” and the PRBD option as “all published COI records from BOLD and GenBank with a minimum sequence length of 500bp. This library is a collection of records from the published projects section of BOLD.”

### Concordance of identifications

The number of specimens that could be matched to a BOLD record was recorded for both the PRBD and SLBR options, across the archived and current databases (2009–2015, 2019). Concordance was measured based on taxonomic rankings obtained from BOLD. Each specimen identification was put into one of five categories for concordance: 1) a DNA identification having a match (>98%) with a lower-level taxonomic name than the morphological identification, 2) a concordant identification having a match (>98%) to a record with identical species name as morphological identification, or a lower level morphological identification having a match (>98%) to a record with a higher level taxonomic name but consistent with lower level morphological identification, 3) an interim identification having a match (>98%) to a record with interim identification only (BIN) or match to a record (<98%) with the same Linnaean name or match (>98%) to a record with a Linnaean species name sister to the target species, 4) no match available using the BOLD system and the given dataset, 5) the final

category was where a discordant identification result indicated a mismatch to a record in the dataset where the morphological identification and the taxonomic name associated with a record in BOLD do not agree and do not reflect a lower or higher taxonomic rank with respect to the morphological identification (>98%). A further analysis using the concordance data was conducted specifically using results from the 2019 database. This analysis was completed to determine how many specimens were identified to the species level using the current SLBR and PRBD database options versus how many were identified to the species level using morphological methods alone.

### BIN analysis

Using the 2019 SLBR database, specimens associated BINs and the records they contained evaluated. BINs were assigned to one of two categories based on the sequence's associated taxonomic identification: 1) concordant where sequences within the BIN matched to an identical taxonomic rank or gave higher level results that were consistent with lower level taxonomic assignments or 2) discordant where taxonomic naming of sequences within this BIN did not agree.

## Results

### Sequencing

Of the 241 specimens examined, full length DNA barcodes ( $\geq 648$ bp) were obtained for 188 individuals. A single primer pair LepF1/LepR1 was used to generate barcodes for 173 individuals, while the remaining individuals were generated with a combination of mLepF1/LepR1 and LepF1/mLepR1. DNA Barcode fragments ( $\leq 407$ bp) were obtained from an additional 13 individuals, bringing the total number of specimens with sequence data to 201.

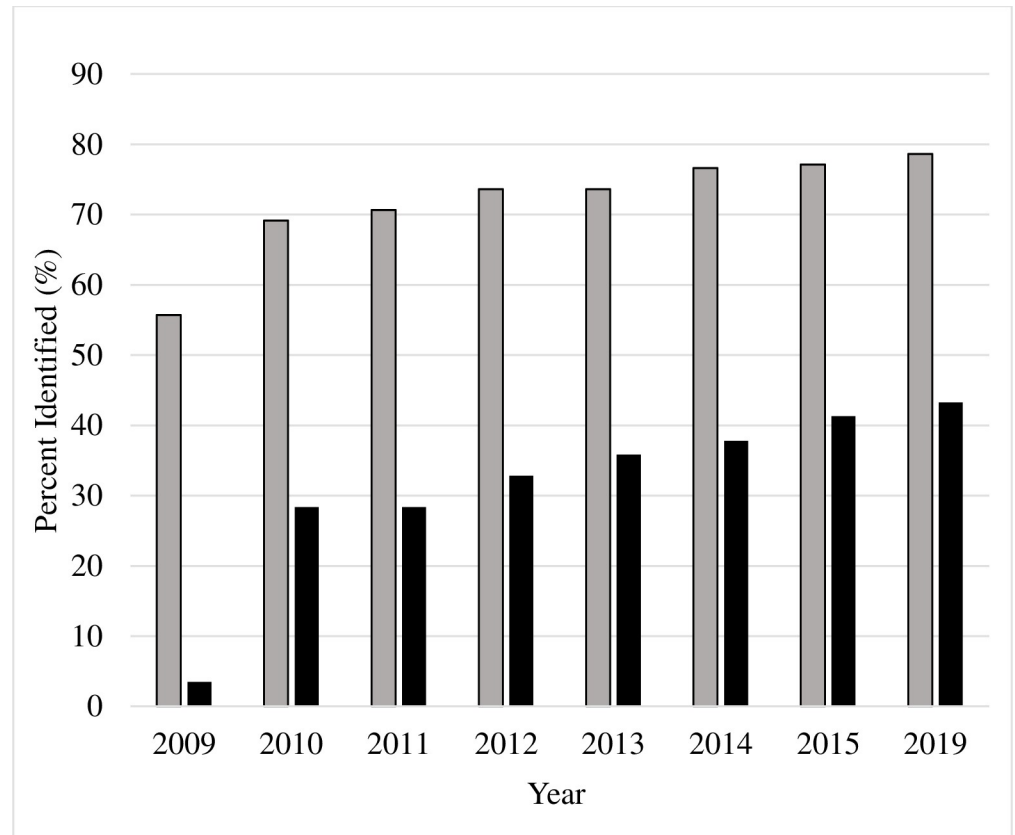
### Identification using BOLD

For the 201 sequences, using both the SLBR (non-public) and PRDB (public) options, the number of BOLD identifications increased over time (Fig 1). Identifications by SLBR was consistently higher than identification by PRDB. Specimen identification by the 2019 PRDB option (43.2%) was unable to surpass the identification level achieved by the SLBR option in 2009 (55%). In both cases, the highest taxonomic identification was at the level of order, and the lowest taxonomic identification was at the level of species.

### Concordance of identifications

The results of the concordance analysis are shown in Fig 2. Category 5 (specimen's sequence did not match to a barcode record in BOLD), represented by the light-blue shaded region of Fig 2, decreased over time for both SLBR and PRDB. In the final 2019 analysis, PRDB had more specimens that matched this category as compared to SLBD identification with 56.7% and 10.9% specimens assigned to this category, respectively. For categories 1–3, all increased over time as reference libraries were being actively built. Additionally, over time there was an increase in the number of specimens assigned to category 5 (mismatch to a record in the dataset where the morphological identification and the taxonomic name associated with a record in the system do not agree and do not reflect a lower or higher taxonomic rank with respect to the morphological identification (>98%)). The increase over time in the number of specimens in category 5 was higher for the SLBD (4.0% to 7.9%) than the PRDB (0.0% to 2.0%).

A more in-depth analysis of concordance for species-level identification by morphology as compared to DNA barcoding using the 2019 PRBD and SLBR results was conducted. DNA barcoding was able to provide species-level identifications for 42.3% (85/201) of specimens



**Fig 1. Percent identification for 201 specimen increases for SLBR and PRBD as BOLD increases in age, specimen identification by Species Level Barcode Record (SLBR = grey) and Public Record Barcode Database (PRBD = black).**

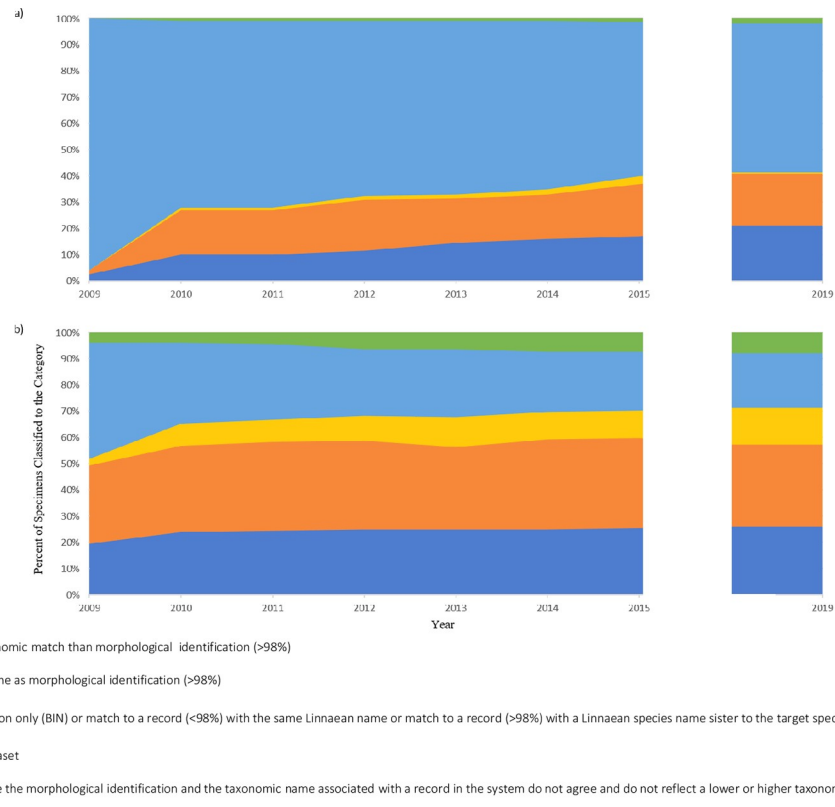
<https://doi.org/10.1371/journal.pone.0222291.g001>

using the PRBD option, and 66.7% (134/201) using the SLBR option using the 2019 dataset. This is greater than the number of specimens identified to the species level by morphology which identified only 38.3% (77/201) of the specimens to the species level. To take a closer look at concordance, the number of specimens identified to the species level by both morphology and DNA barcoding were examined. Of the 38.3% (77/201) of specimens identified to species level using morphology, 55.8% (43/77) and 100% (77/77) were also identified to the species level using the PRBD and SLBR options, respectively. Focusing on PRBD, 93.0% (40/43) gave a concordant ID, 4.65% (2/43) gave an interim-ID and 2.3% (1/43) gave a discordant ID. In contrast, using the SLBR option, 77.9% (60/77) gave a concordant ID, 14.3% (11/77) gave an interim identification and 2.6% (2/77) gave a discordant ID.

### BIN analysis

Finally, for the total set of 201 specimens, barcode analysis using the BIN system resulted in 92 BINs, 34 of which were newly established in BOLD. Upon examination, 40.2% (37/92) contained a single sequence and thus were excluded from further BIN analysis because with only a single sequence in the BIN, variation of the records corresponding to morphological identification was not possible. The remaining 59.8% (55/92) of the BINs had multiple sequences and were analyzed based on the morphological identification agreement for the records. The results indicate that in 55 BINs with multiple sequences, 85.5% (47/55) were concordant and





**Fig 2. Development of concordance between the PRDB and SLBR with morphological identification over the course of 10 years, analysis of PRDB = Fig 2a, analysis of SLBR = Fig 2b.**

<https://doi.org/10.1371/journal.pone.0222291.g002>

14.5% (8/55) were discordant. Furthermore, it was found that 61 of the BINs were based on barcode data from 10 or fewer records.

### Discussion

To demonstrate the improved utility of BOLD to identify specimens, a retrospective analysis from the years 2009 to 2015 was conducted. This was followed by an analysis of current 2019 BOLD system. To highlight a lack of data movement, we contrasted specimen identifications using public (PRDB) and non-public (SLBR) data. Concordance between morphological and molecular identifications was also studied, and the results support DNA barcoding as a valuable method for specimen identification. In addition, analysis of the BINs used in our dataset demonstrated the need to increase BOLD data for two reasons: 1) to increase our confidence in BINs that contain 10 or fewer specimens; and 2) to resolve potential species complexes demonstrated by morphological uncertainty within a single BIN. We find that the use of DNA barcoding is advantageous in identifying specimens at ports-of-entry and suggest that increased efforts are needed to continue to populate public DNA barcode record databases.

An inevitable increase in insect invasions due to globalization is expected in the coming decades, especially in developed countries that already experience a high number of invasion events [45,46]. To slow this trend, an increase in our knowledge of potential invasion pathways, and the effective storage of interception data, are necessary. In our dataset, the country of collection of the specimen was classified as “Exception Quarantine Capture” in the BOLD system. Putative origin (of the commodity) and actual point of interception (port-of-entry) are

described in the BOLD “Collection Notes.” These data may provide regulatory agencies with information regarding the pathway of the commodity, thereby identifying potential points at which insect stowaways may have been acquired [45]. The use of collection localities for intercepted specimens such as the port-of-entry (e.g., “Miami, Florida” for a commodity that originated in Ghana) will leave out key information that may be valuable for the development of regulatory protocols concerning pathways of invasive species [46]. With increased globalization, clear annotation of data regarding origin versus point of interception is critical. Labelling intercepted specimens as “Exception-Quarantine Capture” provides more accuracy in the associated data and should be implemented as best practice for all future datasets concerning intercepted specimens.

As reference barcode sequences continue to be generated, bioinformatic re-examination of molecular DNA barcode data can help to evaluate progress in cataloging global biodiversity for use in molecular DNA identifications. In this study BOLD archives were used to compare specimen identifications from the years 2009–2015 to current data in 2019. Our results indicate an increase in BOLD’s ability to identify microlepidoptera specimens, for both the SLBR and PRBD data sets (Fig 1). From the years 2015 to 2019 BOLD growth for our dataset of interest slowed. These years generated a small increase in identified specimens, with the number rising by only 1.5% (SLBR) and 2.0% (PRBD). Furthermore, for the current 2019 BOLD system, identifications were achieved for only 43.3% (PRBD) and 78.6% (SLBR) of the 201 specimens in the dataset. This indicates a need for further reference library development.

The continued development of molecular libraries, like BOLD, is made possible through the support of diverse user groups with different mandates and research emphasis including biodiversity, taxonomy, and forensic applications (e.g., targeted species identifications) [33]. Continued support for discovery-based research is needed, particularly for biodiverse countries and taxa [47]. It is the initiation of cooperative regional programs for the barcoding of intercepted specimens that simultaneously allow for the confirmation of morphological identification while bolstering sequence data [48].

The data on BOLD has varying levels of access, including datasets with records that are public (PRDB) and not public (SLBR). As seen in Fig 1, the number of specimens identified by the SLBR, which includes non-public data, is much greater than those identified by the PRDB. Looking at the 2019 PRDB results only 43.3% (87/201) of specimens received an identification by BOLD. This is less than the number of specimens identified in 2009 by SLBR which was 55.7% (112/201). It is startling to see that the 2019 public database is still unable to surpass the 2009 non-public database in terms of specimen identification for our dataset. This could indicate that some data has not been made public for over 10 years. For non-public data, BOLD users do have the ability to request access to information from anonymous projects (refer to [www.boldsystems.org](http://www.boldsystems.org)). This gives users the potential to increase their datasets however, access may not be granted, and the process can be lengthy depending on the response time of the data-owner. Access to public data on BOLD also supplies users with metadata such as chromatogram files, the location of voucher specimens, and the presence and location of long-term storage of DNA or tissue [33, 34]. BOLD records that are public and complete with metadata are what allow for linkages between intercepted larvae and vouchered adults [33]. Alternatively, for species that are yet to be morphologically described, including a molecular approach has the potential to integrate taxonomic data, allowing specimens of different life-stages to be connected [49, 50]. It should be expected that the occurrence of linkages between specimens of two different life-stages would increase as more biodiversity-focused barcoding-library building projects are conducted [51–53].

Public access to BOLD records is also important for data verification. For DNA-based specimen identification in biosecurity programs, only records with public and complete metadata



should be used. This allows for record validation by a content expert for issues that arise during the identification process. Furthermore, publicly accessible data encourages harmonization between governments, potentially leading to more efficient conflict resolution regarding regulatory blocks and other trade-related issues. This transparency may not be possible for other DNA barcode reference libraries such as Genbank [54], which historically have not supported sequence-specimen metadata associations. The adoption of systems to evaluate and/or rank the quality of specimen records [55] will be a necessary prerequisite in the widespread adaptation of DNA barcoding for regulatory applications.

The number of specimens which were matched to a record in the BOLD system has increased across the years analyzed (Fig 1). While it is true that we could expect an increase in the number of identifications due to an increase in the number of reference sequences in the database, this is not necessarily true for concordance between DNA barcoding and morphology [56–58]. Our results from Fig 2 also highlight the need for continued sampling, data acquisition, and storage in the public datasets given the large percentage of our dataset specimens which could not be identified (Fig 2 category 4), had interim identification based on molecular data (Fig 2 category 3), or could only be placed to a barcode record with a higher taxonomic rank (Fig 2 category 2). There is also need for increased data management of existing records, such as records with corresponding molecular evidence but having differing morphological taxonomic assignments (Fig 2 category 5).

There is continued need for new public data submitted to BOLD, but there is also need for the movement of existing data from the SLDB to PRDB. This movement of data will provide researchers with access to metadata necessary when evaluating the quality of a record and to verify matches. The addition of data and efforts to further curate accessible data will result in a greater number of records providing reliable sequence representation for species. In the context of invasive pest management, identifications relevant to inform biosecurity decisions must be at the species level. Obtaining specimen identifications to the species level is desired, but given the magnitude of global arthropod biodiversity, this is still challenging [49, 47]. Using the 2019 BOLD data, molecular identification surpassed morphological species assignments (made in 2009–2015) in the number of specimens that received a species-level identification; morphology 38.3% (77/201), PRDB 42.3% (85/201), and SLBR 66.7% (134/201) (Fig 2 category 1).

While DNA barcoding may not be able to provide species level identifications for all specimens at this time, its use as a complementary identification technique is recommended. Fig 2 categories 1 and 2 demonstrate that DNA barcoding is able to place a specimen to a species or taxonomic rank consistent with morphological identification for more than 30% of our cases using the PRDB 2019. These data show how the use of both DNA barcoding and morphological identifications in conjunction can provide secondary evidence to support morphological identifications and, in the category 1 cases, increase taxonomic resolution. Both outcomes are important in biosecurity as unidentifiable tissue and difficult to identify specimens (such as microlepidoptera) are often intercepted.

DNA barcoding also has value in highlighting potential misidentifications by morphological methods [59]. For this microlepidoptera dataset, DNA barcode species-level identifications lacked concordance with morphological species-level identifications for 2.3% (PRDB 2019) and 2.6% (SLBR 2019) of specimens (Fig 2 category 5). The use of record metadata, including the location of voucher specimens for re-evaluation where appropriate, is essential when scrutinizing these data to make informed decisions regarding clarification of identifications and curation of records in the database. While the cases present in this study, may represent a morphological misidentification, particularly considering the difficult nature of microlepidoptera identifications, these cases are difficult to verify due to the lack of public barcode records.

Future efforts to populate the BOLD system with adult expertly identified microlepidoptera specimens is necessary to fully investigate these problematic records.

DNA barcoding can also provide interim identifications through molecular operational taxonomic units (MOTU) which, when molecular libraries are further populated, can retroactively provide identifications [49, 47]. In this dataset, intercepted specimens fell into 92 BINS. Of these, 34 BINS were new to BOLD. This is not surprising because outside of Costa Rica, there are a limited number of records of microlepidoptera with sequence data for species from the Neotropics, the origin of many of the specimens in the intercepted dataset [51]. Variation in global sample collections likely make specimen identifications biased based on region of origin. Specimen interceptions entering North America with origins from Europe and Australia are more likely to have BOLD records due to greater research efforts on microlepidoptera from these regions. In contrast, interceptions from the Neotropics, Africa, and Asia are less likely to be identified to a low taxonomic level owing to the paucity of sequences in BOLD, or their absence altogether. Due to the global nature of tracking and identifying insects of concern, a collaboration of nations, particularly those connected by shared borders or trade routes, is necessary to further build molecular barcode libraries for use in biosecurity applications [47, 60].

Nearly 15% of the BIN's identified in our study were classified as discordant BINs (taxonomic naming of sequences within a BIN did not agree) by BOLD. This discordance may be the result of species complexes (i.e., closely related species that are not easily separated by morphology and/or barcodes) or may be the result of misidentified specimens. For species complexes, resolution relies on the associated metadata and may require additional molecular techniques, morphometry, ecology, or morphological characters from all life stages [61]. Unrecognized species complexes may be due to the use of poor taxonomic keys, inadequate sampling, identifier bias (i.e., emphasizing different morphological characters), and/or geographical naming [61]. Upon analysis of the discordant BINs by expert Lepidoptera taxonomists (co-authors Brown and Miller), some BINs were suspected as discordant due to misidentified specimens in the reference library rather than a species complex. For example, one discordant BIN (BOLD: AAP2599) contained three families; Blastobasidae, Gelechiidae and Oecophoridae, but only a single species name; *Calosima albapennella*. This use of Gelechiidae and Oecophoridae in this BIN likely represents misidentifications, with the true family being Blastobasidae. In contrast, another BIN (BOLD: AAA7690) was associated with three species, all from the same genus; *Archips packardiana*, *Archips alberta* and *Archips tsuganus*. This specific example is more likely to represent a species complex and further investigation is required to determine the source of discordance.

In the context of pest-species complexes, a higher-level taxonomic classification (i.e., genus-level identification) may be suitable to enact biosecurity actions. Therefore, clarity of BIN discordance can improve the applicability of DNA barcoding and further support biosecurity decisions. These assessments of discordant BINs required the use of accessible metadata, such as image files, and demonstrates the importance of well documented public records. Additionally, increasing sampling efforts and data accumulation typically helps reveal the cause of disagreement among records in a BIN [49]. In our dataset, 66.3% (61/92) of the BINs were comprised of less than 10 records, indicating a need to improve sampling efforts.

Material collected from biodiversity and ecological surveys provide a unique opportunity to catalogue undescribed species as well as augment distributions and sample haplotype diversity for species already represented in the reference library. It is just as important to further our efforts in obtaining expertly identified material from sources such as museum collections to increase records with species level names in the BOLD system [62]. Together, these efforts can provide a robust data system upon which the identification and subsequent verification of

identifications can be conducted. DNA barcoding studies, such as plant protection and quarantine studies like this one, must endeavor to catalogue DNA barcoding data by archiving query sequences along with associated metadata. As such, it is encouraged that all query sequences be archived in publicly accessible libraries regardless of the reason they were initially generated. Through the use of MOTU approaches like BINs, the task of organizing otherwise unidentifiable material into species-like units is entirely possible on a large scale. Harmonizing these MOTUs in a central database (i.e., BOLD) greatly facilitates their use as interim taxonomy which is accessible by multiple working groups and researchers globally.

Unfortunately, 16.6% (40/241) of intercepted specimens in this dataset did not yield sequence data, which was surprising given that most samples were less than one year old and presumably experienced similar storage conditions. Considering that the primers employed in this study have been used extensively for barcoding of Lepidoptera [44, 51, 52, 63–65], it seems likely that failure to amplify sequences from these specimens was a result of DNA quality rather than primer selection. One explanation is that larval samples from APHIS-PPQ are typically submitted in 75% ethanol, which may result in the degradation of DNA over time. In addition to data collection, we would encourage the use of storage methods amenable to molecular data acquisition such as the storage of specimens or tissue in >95% ethanol along with the immediate and consistent storage of ethanol preserved specimens at low temperatures until genetic material is obtained [66]. These efforts can dramatically increase our ability to obtain molecular data from specimens and accomplish many of the advantages discussed above.

## Conclusion

The identification of specimens for biosecurity are time sensitive; i.e., they are most valuable when made within a few hours, not days. Although barcoding may increase the reliability of “routine” identifications (i.e., those submitted without an associated deadline), as of yet, the process is unable to assist in the identification of specimens submitted as “urgent” (i.e., requiring rapid turnaround), where morphological identification remains the best option. Even so, increasing the accuracy of “routine” identifications can help to speed the DNA barcoding of specimens, making future rapid identifications using a molecular approach feasible. This study demonstrates the need for the accumulation of records relevant to biosurveillance globally. Furthermore, recording location data for suspected country of origin and country of interception separately, needs to be standardized as the method of reporting data for border-interceptions. With increased globalization the use of both morphological and molecular identification approaches is necessary if we want to effectively combat the growing number of intercepted specimens at ports-of-entry. A combined approach is essential for building DNA barcoding reference libraries thereby increasing the reliability of identifications which will inform future biosecurity actions.

## Acknowledgments

Continued progress on DNA barcoding Lepidoptera from the Smithsonian collection has been made by collaboration between the University of Guelph and the Smithsonian, with help from many people, including those listed in Levesque-Beaudin et al. (2017). Contributions by JWB were supported by the USDA, ARS, Systematic Entomology Laboratory, Washington, DC.

## Author Contributions

**Conceptualization:** Scott E. Miller, Andrew J. Frewin, Robert H. Hanner.

**Data curation:** Mary J. L. Madden, Robert G. Young, Scott E. Miller.

**Formal analysis:** Mary J. L. Madden, Robert G. Young, Andrew J. Frewin, Robert H. Hanner.

**Funding acquisition:** Robert H. Hanner.

**Investigation:** Mary J. L. Madden, Andrew J. Frewin, Robert H. Hanner.

**Methodology:** John W. Brown, Scott E. Miller.

**Project administration:** Robert H. Hanner.

**Resources:** John W. Brown, Robert H. Hanner.

**Supervision:** Robert H. Hanner.

**Validation:** Mary J. L. Madden, Robert G. Young, John W. Brown, Scott E. Miller.

**Visualization:** Mary J. L. Madden, Robert G. Young.

**Writing – original draft:** Mary J. L. Madden, Robert G. Young, Andrew J. Frewin, Robert H. Hanner.

**Writing – review & editing:** Mary J. L. Madden, Robert G. Young, John W. Brown, Scott E. Miller, Andrew J. Frewin, Robert H. Hanner.

## References

1. Caffrey JM, Baars JR, Barbour JH, Boets P, Boon P, Davenport K, et al. Tackling invasive alien species in Europe: the top 20 issues. *Manage Biol Invasions*. 2014; 5(1): 1–20. <https://doi.org/10.3391/mbi.2014.5.1.01>
2. Oerke EC. Crop losses to pests. *J Agri Sci*. 2006; 144(1): 31–43. <https://doi.org/10.1017/S0021859605005708>
3. Augustin S, Boonham N, De Kogel WJ, Donner P, Faccoli M, Lees DC, et al. A Review of Pest Surveillance Techniques for Detecting Quarantine Pests in Europe. *EPPO Bulletin*. 2012; 42: 515–551. <https://doi.org/10.1111/epp.2600>
4. Oliveira CM, Auad AM, Mendes SM, Frizzas MR. Crop losses and the economic impact of insect pests on Brazilian agriculture. *Crop Prot*. 2014; 56: 50–54. <https://doi.org/10.1016/j.cropro.2013.10.022>
5. Westphal MI, Browne M, MacKinnon K, Noble I. The link between international trade and the global distribution of invasive alien species. *Biol Invasions*. 2008; 10(4): 391–398. <https://doi.org/10.1007/s10530-007-9138-5>
6. Navia D, Ochoa R, Welbourn C, Ferragut F. Adventive eriophyoid mites: a global review of their impact, pathways, prevention and challenges. *Exp Appl Acarol*. 2010; 51(1–3): 225–55. <https://doi.org/10.1007/s10493-009-9327-2> PMID: 19844795
7. Lenda M, Skorka P, Knops JMH, Moron D, Sutherland WJ, Kuszewska K et al. Effect of the internet commerce on dispersal modes of invasive alien species. *Plos One* 2014; 9:7. <https://doi.org/10.1371/journal.pone.0099786> PMID: 24932498
8. Suffert M, Wilstermann A, Petter F, Schrader G, Grousset F. Identification of new pests likely to be introduced into Europe with the fruit trade. *EPPO Bulletin*, 2018; 48(1): 144–154. <https://doi.org/10.1111/epp.12462>
9. Moser WK, Barnard EL, Billings RF, Crocker SJ, Dix M, Gray AN, et al. Impacts of Nonnative Invasive Species on US Forests and Recommendations for Policy and Management. *J Forest*, 2009; 107(6): 320–327. <https://doi.org/10.1093/jof/107.6.320>
10. Bacon SJ, Bacher S, Aebi A. Gaps in Border Controls Are Related to Quarantine Alien Insect Invasions in Europe. *PLoS ONE*, 2012; 7(10). <https://doi.org/10.1371/journal.pone.0047689> PMID: 23112835
11. Kenis M, Rabitsch W, Auger-Rozenberg MA, Roques A. How can alien species inventories and interception data help us prevent insect invasions? *Bull Entomol Res*. 2007; 97(5): 489–502. <https://doi.org/10.1017/S0007485307005184> PMID: 17916267
12. Harwood TD, Xu X, Pautasso M, Jeger MJ, Shaw MW. Epidemiological risk assessment using linked network and grid based modelling: *Phytophthora ramorum* and *Phytophthora kernoviae* in the UK. *Ecol Model*. 2009; 220: 3353–3361. <https://doi.org/10.1016/j.ecolmodel.2009.08.014>

13. Moslonka-Lefebvre M, Finley A, Dorigatti I, Dehnen-Schmutz K, Harwood T, Jeger MJ, et al. Networks in plant epidemiology: From genes to landscapes, countries, and continents. *Phytopathology*. 2011; 101: 392–403. <https://doi.org/10.1094/PHYTO-07-10-0192> PMID: 21062110
14. Whittle P, Stoklosa R, Barrett S, Jarrad FC, Majer JD, Martin PAJ, et al. A method for designing complex biosecurity surveillance systems: detecting non-indigenous species of invertebrates on Barrow Island. *Divers and Distrib*. 2013; 19: 629–639. <https://doi.org/10.1111/ddi.12056>
15. Vanninen I, Worner S, Huusela-Veistola E, Tuovinen T, Nissinen A, Saikkonen K. Recorded and potential alien invertebrate pests in Finnish agriculture and horticulture. *Agri Food Sci*. 2011; 20: 96–113. <https://doi.org/10.2137/145960611795163033>
16. Schachat SR. The wing pattern of *Moerarchis Durrant*, 1914 (Lepidoptera: Tineidae) clarifies transitions between predictive models. *R Soc Open Sci*. 2017; 4(3), 1–12. <https://doi.org/10.1098/rsos.161002> PMID: 28405390
17. Sweeney BW, Battle JM, Jackson JK, Dapkey T. Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *J N Am Benthol Soc*. 2011; 30(1): 195–216. <https://doi.org/10.1899/10-016.1>
18. Stehr FW, editors. *Immature insects*. Vol. 1. Dubuque: Kendall Hunt Publishing Co; 1987. np.
19. Gilligan TM, Goldstein PZ, Timm AE, Farris R, Ledezma L, Cunningham AP. Identification of Heliothine (Lepidoptera: Noctuidae) Larvae Intercepted at U.S. Ports of Entry from the New World. *J Econ Entomol*. 2019; 1–13. <https://doi.org/10.1093/jee/toy247>
20. Ko HL, Wang YT, Chiu TS, Lee MA, Leu MY, Chang KZ, et al. Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *Plos One* 2013; 8: <https://doi.org/10.1371/journal.pone.0053451> PMID: 23382845
21. McCullough DG, Work TT, Cavey JF, Liebhold AM, Marshall D. Interceptions of nonindigenous plant pests at US ports of entry and border crossings over a 17-year period. *Biol Invasions*. 2006; 8: 611–630. <https://doi.org/10.1007/s10530-005-1798-4>
22. Worner SP, Gevrey M. Modelling global insect pest species assemblages to determine risk of invasion. *J Appl Ecol*. 2006; 43: 858–867. <https://doi.org/10.1111/j.1365-2664.2006.01202.x>
23. Brasier M. The biosecurity threat to the UK and global environment from international trade in plants. *Plant Pathol*. 2008; 57(5): 792–808. <https://doi.org/10.1111/j.1365-3059.2008.01886.x>
24. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proc R Soc B*. 2003; 270: 313–321. <https://doi.org/10.1098/rspb.2002.2218> PMID: 12614582
25. Frewin A, Scott-Dupree C, Hanner R. DNA barcoding for plant protection: applications and summary of available data for arthropod pests. *CAB Rev*. 2013; 8(18): 1–13. <https://doi.org/10.1079/PAVSNNR20138018>
26. Wilson JRU, Ivey P, Manyama P, Naenni I. A new national unit for invasive species detection, assessment and eradication planning. *S Afr J Sci*. 2013; 109: 33–45. <https://doi.org/10.1590/sajs.2013/20120111>
27. Serrao NR, Steinke D, Hanner RH. Calibrating Snakehead Diversity with DNA Barcodes: Expanding Taxonomic Coverage to Enable Identification of Potential and Established Invasive Species. *Plos One* 2014; 9(6): e99546. <https://doi.org/10.1371/journal.pone.0099546> PMID: 24915194
28. Wilson AD, Schiff NM. Identification of *Sirex noctilio* and Native North American Woodwasp Larvae using DNA Barcode. *J Entomol*. 2010; 7: 60–79. <https://doi.org/10.3923/je.2010.60.79>
29. Brabrand A, Bremnes T, Koestler AG, Marthinsen G, Pavels H, Rindal E, et al. Mass occurrence of bloodsucking blackflies in a regulated river reach: Localization of oviposition habitat of *Simulium truncatum* using DNA barcoding. *River Res Appl*. 2014; 30: 602–608. <https://doi.org/10.1002/rra.2669>
30. Mastrangelo T, Paulo DF, Bergamo LW, Morais EGF, Silva M, Bezerra-Silva G, et al. Detection and genetic diversity of a heliothine invader (Lepidoptera: Noctuidae) from north and northeast of Brazil. *J Econ Entomol*. 2014; 107: 970–980. <https://doi.org/10.1603/ec13403> PMID: 25026655
31. Pramual P, Wongpakam K. Association of black fly (Diptera: Simuliidae) life stages using DNA barcode. *J Asia-Pac Entomol*. 2014; 17: 549–554. <https://doi.org/10.1016/j.aspen.2014.05.006>
32. Ratnasingham S, Hebert PDN. BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Mol Ecol Notes*. 2007; 7: 355–364. <https://doi.org/10.1111/j.1471-8286.2007.01678.x> PMID: 18784790
33. Borisenko AV, Sones JE, Hebert PD. The front-end logistics of DNA barcoding: Challenges and prospects. *Mol Ecol Resour*. 2009; 9: 27–34. <https://doi.org/10.1111/j.1755-0998.2009.02629.x> PMID: 21564961
34. Hanner R. 2009. Data Standards for BARCODE Records in INSDC (BRIs). <https://github.com/SIBarcodeNetwork/SIBarcodeNetwork/blob/master/BARCODE%20Data%20Standards%20v2.4.pdf>



35. Curry C. J., Gibson J. F., Shokralla S., Hajibabaei M., & Baird D. J. (2018). Identifying North American freshwater invertebrates using DNA barcodes: are existing COI sequence libraries fit for purpose?. *Freshwater Science*, 37(1), 178–189.
36. Porter T. M., & Hajibabaei M. Over 2.5 million COI sequences in GenBank and growing. *PloS one*. 2018; 13(9), e0200177. <https://doi.org/10.1371/journal.pone.0200177> PMID: 30192752
37. Blaxter M.L. The promise of a DNA taxonomy. *Philos.Trans. R. Soc. B Biol. Sci.* 2004; 359(1444): 669–679. <https://doi.org/10.1098/rstb.2003.1447> PMID: 15253352
38. Ratnasingham S, Hebert PDN. A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *Plos One* 2013; 8: <https://doi.org/10.1371/journal.pone.0066213> PMID: 23861743
39. Zahiri RJ, Lafontaine D, Schmidt BC, deWaard JR, Zakharov EV, Hebert PDN. A transcontinental challenge—A test of DNA barcode performance for 1,541 species of Canadian Noctuoidea (Lepidoptera). *Plos One*. 2014; 9(3): e92797. <https://doi.org/10.1371/journal.pone.0092797> PMID: 24667847
40. Castalanelli MA, Severtson DL, Brumley CJ, Szito A, Footitt RG, Grimm M, et al. A rapid non-destructive DNA extraction method for insects and other arthropods. *J Asia-Pac Entomol.* 2012; 13(3): 243–248. <https://doi.org/10.1016/j.aspen.2010.04.003>
41. Wilson JJ. DNA Barcodes for Insects. *Method Mol Biol*, 2012; 858: 17–46. [https://doi.org/10.1007/978-1-61779-591-6\\_3](https://doi.org/10.1007/978-1-61779-591-6_3)
42. Ivanova NV, Borisenko AV, Hebert PDN. Express barcodes: racing from specimen to identification. *Mol Ecol Resour.* 2009; 9(s1). <https://doi.org/10.1111/j.1755-0998.2009.02630.x> PMID: 21564962
43. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc Natl Acad Sci USA.* 2004; 101(41): 14812–14817. <https://doi.org/10.1073/pnas.0406166101> PMID: 15465915
44. Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PD. DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences.* 2006; 103(4):968–71.
45. Hulme PE. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *J Appl Ecol.* 2009; 46(1): 10–18. <https://doi.org/10.1111/j.1365-2664.2008.01600.x>
46. Early R, Bradley BA, Dukes JS, Lawler JJ, Olden JD, Blumenthal DM, et al. Global threats from invasive alien species in the twenty-first century and national response capacities. *Nat Commun.* 2016; 7: 12485. <https://doi.org/10.1038/ncomms12485> PMID: 27549569
47. Watts C, Dopheide A, Holdaway R, Davis C, Wood J, Thornburrow D, et al. DNA metabarcoding as a tool for invertebrate community monitoring: A case study comparison with conventional techniques. *Austral Entomol.* 2019; 8(7): e1000417. <https://doi.org/10.1111/aen.12384>
48. Bonants PJM. Results of the EU Project QBOL, Focusing on DNA Barcoding of Quarantine Organisms, Added to an International Database (Q-Bank) on Identification of Plant Quarantine Pathogens and Relatives. In: Gullino M, Bonants P, editors. *Detection and Diagnostics of Plant Pathogens. Plant Pathology in the 21st Century (Contributions to the 9th International Congress)*. Dordrecht: Springer; 2014. pp. 119–133.
49. Phillips JD, Gillis DJ, Hanner RH. Incomplete estimates of genetic diversity within species: Implications for DNA barcoding. *Ecol Evol*, 2019; 9(5), 2996–3010. <https://doi.org/10.1002/ece3.4757> PMID: 30891232
50. Padiá JM, Miralles A, De la Riva I, Vences M. The integrative future of taxonomy. *Front Zool.* 2010; 7: 1–14. <https://doi.org/10.1186/1742-9994-7-1>
51. Janzen DH, Hajibabaei M, Burns JM, Hallwachs W, Remigio E, Hebert PDN. Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Philos Trans R Soc Lond B Biol Sci.* 2005; 360: 1835–1845. <https://doi.org/10.1098/rstb.2005.1715> PMID: 16214742
52. Janzen DH, Hallwachs W, Harvey DJ, Darrow K, Rougerie R, Hajibabaei M, et al. What happens to the traditional taxonomy when a well-known tropical saturniid moth fauna is DNA barcoded? *Invertebr Syst.* 2012; 26(6): 478–505. <https://doi.org/10.1071/IS12038>
53. Smith AM, Fernández-Triana JL, Eveleigh E, Gómez J, Guclu C, Hallwachs W, et al. DNA barcoding and the taxonomy of Microgastrinae wasps (Hymenoptera, Braconidae): impacts after 8 years and nearly 20 000 sequences. *Mol Ecol Resour* 2012; 13: 168–176. <https://doi.org/10.1111/1755-0998.12038> PMID: 23228011
54. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2005. GenBank. *Nucleic Acids Res*, 2005; 33(Database issue): D34–8. <https://doi.org/10.1093/nar/gki063> PMID: 15608212
55. Costa FO, Landi M, Martins R, Costa MH, Costa ME, Carneiro M, et al. A ranking system for reference libraries of DNA barcodes: Application to marine fish species from Portugal. *Plos One.* 2012; 7(4): <https://doi.org/10.1371/journal.pone.0035858> PMID: 22558244
56. Meyer CP, Paulay G (2005) DNA Barcoding: Error Rates Based on Comprehensive Sampling. *PLOS Biology* 3(12): e422. <https://doi.org/10.1371/journal.pbio.0030422> PMID: 16336051



57. Wiemers M., & Fiedler K. (2007). Does the DNA barcoding gap exist?—a case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in zoology*, 4(1), 8.
58. Puillandre N., Macpherson E., Lambourdière J., Cruaud C., Boisselier-Dubayle M. C., & Samadi S. (2011). Barcoding type specimens helps to identify synonyms and an unnamed new species in *Eumunida* Smith, 1883 (Decapoda: Eumunididae). *Invertebrate Systematics*, 25(4), 322–333.
59. Liu Z, Ci X, Li L, Li H, Conran JG, Li J. DNA barcoding evaluation and implications for phylogenetic relationships in Lauraceae from China. *Plos One*, 2017;12(4). <https://doi.org/10.1371/journal.pone.0175788> PMID: 28414813
60. Vernooy R, Haribabu E, Muller MR, Vogel JH, Hebert PDN, Schindel DE, et al. Barcoding life to conserve biological diversity: Beyond the taxonomic imperative. *PLoS Biol* 2010; 8: e1000417. <https://doi.org/10.1371/journal.pbio.1000417> PMID: 20644709
61. Milić D, Radenković S, Ačanski J, Vujić A. The importance of hidden diversity for insect conservation: a case study in hoverflies (the *Merodon atratus* complex, Syrphidae, Diptera). *J Insect Conserv.* 2019; 23(1): 29–44. <https://doi.org/10.1007/s10841-018-0111-7>
62. Levesque-Beaudin V., Rosati M. E., Silverson N., Warne C. P., Brown A., Telfer A. C., Sobel C. N., Miskie R. N., Miller M. E., Sones J. E., Miller S. E., and de Waard J. R. 2017. Museum harvesting in major natural history collections. *Genome* 60(11):962. <https://doi.org/10.1139/gen-2017-0178>
63. Hebert PDN, deWaard JR, Landry JF. DNA barcodes for 1/1000 of the animal kingdom. *Biol Lett* 2009; 6: 359–362. <https://doi.org/10.1098/rsbl.2009.0848> PMID: 20015856
64. deWaard JR, Hebert PDN, Humble LM. A comprehensive DNA barcode library for the looper moths (Lepidoptera: Geometridae) of British Columbia, Canada. *Plos One*. 2011;6: <https://doi.org/10.1371/journal.pone.0018290> PMID: 21464900
65. Hebert PDN, deWaard JR, Zakharov EV, Prosser SWJ, Sones JE, McKeown JTA, et al. A DNA 'barcode blitz': Rapid digitization and sequencing of a natural history collection. *Plos One* 2013; 8(7). <https://doi.org/10.1371/journal.pone.0068535> PMID: 23874660
66. Prosser S, Martínez-Arce A, Elías-Gutiérrez M. A new set of primers for COI amplification from freshwater microcrustaceans. *Mol. Ecol. Resour.* 2013; 13(6): 1151–1155. <https://doi.org/10.1111/1755-0998.12132> PMID: 23795700.