DOI: https://doi.org/10.3897/arphapreprints.e162421

# From detection to action—a proposed workflow to ensure first reports of alien species from molecular analyses are acted upon

**Laura Fernandez Winzer, Katelyn T. Faulkner, Trudy Paap, John R.U. Wilson**

# From detection to action—a proposed workflow to ensure first reports of alien species from molecular analyses are acted upon

**Laura Fernández Winzer [1,2,3*], Katelyn T. Faulkner[1,4], Trudy Paap[5], John R.U. Wilson[1,2]**

[1] South African National Biodiversity Institute, Kirstenbosch Research Centre, Cape Town, South Africa

[2] Department of Botany and Zoology, Centre for Invasion Biology, Stellenbosch University, Stellenbosch, South Africa

[3] School of Natural Sciences, Macquarie University, Sydney, Australia

[4] Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa

[5] Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

*Corresponding author: laura.f.winzer@gmail.com

## Abstract

Biosecurity is crucial for managing biological invasions, yet communication gaps between detection and action often hinder timely responses. In South Africa, the polyphagous shot hole borer (*Euwallacea fornicatus*) was first recorded through a DNA barcoding project in 2012 but no action was taken until a separate field observation was made in 2017. Here we present a workflow that aims to ensure molecular observations are evaluated and acted upon. We demonstrate its value using South Africa as a case study. 10,084 records from South Africa's 'Barcode of Life Data System' were compared with three datasets: a watch list of species of concern (400 species), a list of high-risk pests detected at-border, but not post-border (218 species); and the Botanical Database of Southern Africa (BODATSA, 68,153 records). Four species from the BOLD were identified as being on the watch list or in the list of detected high-risk pests and four species were flagged as possible additions to the list of alien plants outside of cultivation in South Africa. The method is codified in a simple, generalisable workflow that, we argue, will improve the flow of information from detection to action and so allow biosecurity systems to become more responsive.

**Keywords:** biosecurity, BOLD, checklist, DNA barcoding, EDRR, nativity, prohibited species, species identification, species of concern, taxonomic standardisation

## Introduction

Detecting an alien species soon after it has been introduced is likely to mean more options are available for management (Reaser et al. 2020). At early stages, populations are typically small and geographically limited, making containment and eradication more feasible. As a result, large-scale harmful invasions can potentially be avoided (IPBES 2023). Alien taxa, as defined here, are any taxa that are not native to the region being considered [for this study, any taxa that is not native to the Republic of South Africa (RSA)], that have been introduced by human derived action – intentionally or unintentionally. Nativity was determined using authoritative databases such as Plants of the World Online for plants, supplemented by published sources when needed, or by using species lists that contained only alien taxa (e.g., watch lists). However, the challenges of early detection are universal, affecting ecosystems across the globe. Two aspects of biosecurity frameworks are particularly important: first, surveillance systems need to be improved so alien species introductions are detected; and second a system is needed to flag detections that should be acted upon (Liebhold et al. 2016).

An infamous example from South Africa is the polyphagous shot-hole borer (PSHB), an ambrosia beetle that together with its fungal symbiont causes dieback and mortality of a diverse range of woody plants. This beetle was detected through a sentinel project at Kwa-Zulu Natal Botanical Gardens in 2017 (Paap et al. 2018), however, it soon became clear that it was already widely established across much of the country (Paap et al. 2020). Although the 2017 detection raised national awareness, by 2021 the PSHB was present in eight out of South Africa's nine provinces, affecting many native, ornamental and commercially important tree species (van Rooyen et al. 2021; de Wit et al. 2022; Townsend et al. 2025). Retrospective analysis revealed that the PSHB had been present in South Africa as early as 2012, when it was sampled as part of a barcoding project (van Rooyen et al. 2021). Whether effective action could have been taken in 2012 is not known, but it is concerning that there were five years between the first detection and the subsequent detection that was acted upon.

Globally, DNA barcoding has emerged as a powerful tool for the detection of alien taxa, providing molecular evidence of their presence in a given region while supporting biosecurity efforts (Armstrong and Ball 2005; Hamelin 2012; Hanner et al. 2009; Madden et al. 2019). The method involves sequencing a standardised region of an organism's genome and comparing it against reference databases. The resulting matches can identify species with high confidence (see Baena-Bejarano et al. 2023), often more reliably than traditional morphological identification methods. Moreover, DNA barcoding proves particularly useful for identifying cryptic species or when morphological identification is challenging due to specimen damage or developmental stages. Integrative approaches combining molecular and morphological data have been shown to enhance identification accuracy, highlighting the limitations of relying solely on morphology (Yang et al. 2022).

The Barcode of Life Data System (BOLD; www.barcodinglife.org) is an open-access DNA barcode database that provides a centralised platform for barcode data (Ratnasingham et al. 2024). Each record is assigned a Barcode Index Number (BIN), a system for clustering DNA barcode sequences into operational taxonomic units that typically correspond to species (Ratnasingham and Hebert 2013). As of May 2025, BOLD contained 16.5 million public records from around the world, representing 1.3 million species across various kingdoms (e.g., Animalia, Bacteria, Chromista, Fungi and Plantae). BOLD's global coverage allows researchers to use this tool to detect and identify species across a broad range of geographic regions and taxonomic groups, supporting biodiversity monitoring and conservation efforts, and increasingly contributing to biosecurity applications worldwide (Hernández-Triana et al. 2019; Comia and Morris 2024).

Despite the potential for reducing the time between detection and response, the integration of molecular barcode data into national and regional biosecurity frameworks remains limited (Armstrong and Ball 2005). Many countries still rely heavily on morphological identification, which is often labour-intensive, error-prone, and limited by the availability of taxonomic expertise (Darling and Blum 2007). This gap in molecular-based surveillance systems leaves many regions vulnerable to undetected alien taxa introductions, particularly in biodiversity hotspots like South Africa, where rapid and accurate identification is crucial (Sethusa et al. 2014; van Wilgen et al. 2020).

This study aims to address this gap by developing a simple, generalisable workflow that cross-references barcode data from BOLD (from within RSA) with three datasets: (1) a national watch list of alien species of concern, (2) a list of high-risk pests intercepted at borders but not yet detected within the country, and (3) the Botanical Database of Southern Africa (BODATSA), which includes alien plants outside of cultivation. By identifying taxa that have been detected during molecular studies but are not yet recognised or managed as alien species in South Africa (i.e., included in the aforementioned lists), we aim to flag potential biosecurity risks to inform action and improve existing species lists for future surveillance. This workflow represents a step towards integrating molecular data into practical biosecurity applications, with the potential to be adapted for use in other regions facing similar challenges.

## Methods

To compare the species lists and identify potential undocumented alien species, we developed a workflow comprising three main steps: (1) acquisition, initial inspection, and preparation of the original datasets; (2) taxonomic standardisation of species names across datasets (following Faulkner 2025); and (3) pairwise comparison of standardised lists (Fig. 1, S1). The workflow was implemented in R (v4.4.2; R Core Team 2024) and relies on several R packages: tidyr (Wickham, Vaughan and Girlich 2024), dplyr (Wickham et al. 2023), stringr (Wickham 2023), rgbif (Chamberlain et al. 2025), stringdist (van der Loo 2014), rWCVP (Brown et al. 2023), purrr (Wickham and Henry 2023), lubridate (Grolemund and Wickham 2011), remotes (Csárdi et al. 2024), and rWCVPdata (Govaerts et al. 2021; Govaerts 2024).

### *1. Species lists*

The primary dataset was derived from the Barcode of Life Data System (BOLD v4. [www.barcodinglife.org/](www.barcodinglife.org/)), providing molecular evidence of species presence in South Africa. This list was compared against three reference datasets to flag species of concern. Below, we describe each dataset and its limitations.

### *1.1 Barcode of Life Database (BOLD)*

The BOLD database (Ratnasingham and Hebert 2007; Ratnasingham et al. 2024) is a global, open-access DNA barcoding platform that includes animals, plants, fungi, bacteria, and

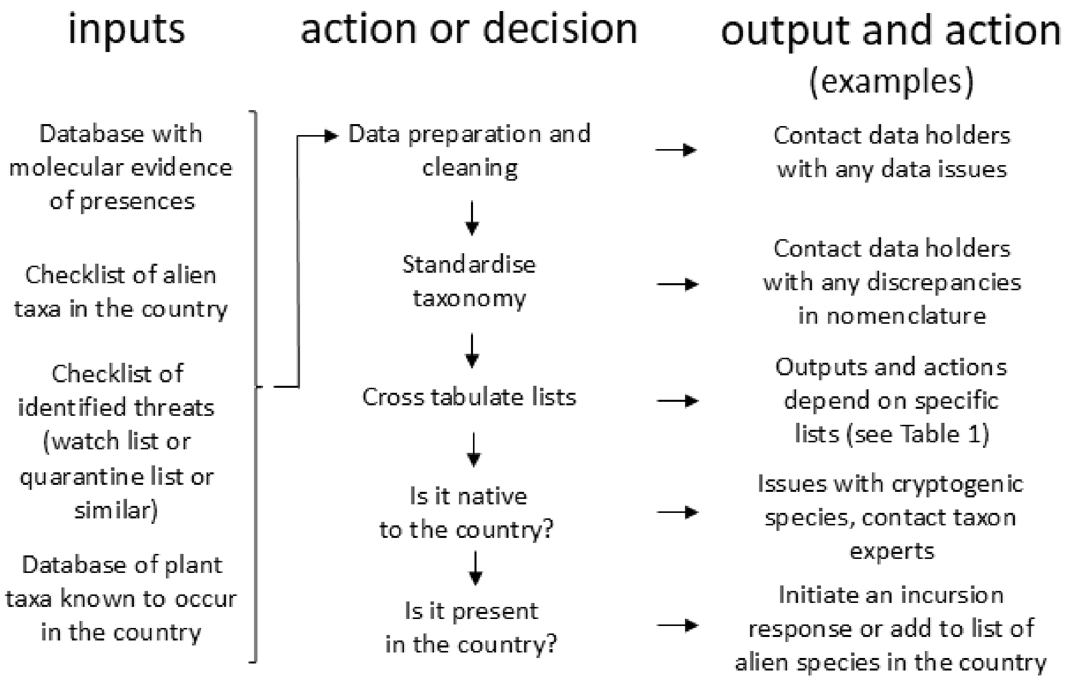

**Figure 1.** An overview of the workflow for comparing country specific species lists with molecular databases.

protists. Available metadata may include collection date and location, collector identity, specimen storage information, and associated media. Critically, the database does not indicate whether a taxon is native to the region where the specimen was collected, posing challenges for detecting alien species (discussed below). Not all records include DNA sequence data, so presence in BOLD does not necessarily confirm presence in South Africa. Data were downloaded on 22 November 2023.

### *1.2 Watch list*

This list, compiled by Faulkner et al. (2014), includes 400 alien species considered to pose a high risk of invasion but not known to be present in South Africa as of 2013. It comprises animals, plants (including algae and ferns), and other taxa such as fungi, micro-organisms, and oomycetes. As the list is static, some species on it (e.g. *Austropuccinia psidii*; Roux et al. 2013) have been detected since the analysis was completed in early 2013.

If a species on the watch list is found to have a DNA barcode collected from samples in South Africa, a response is warranted. The confidence of the identity of the DNA barcode and where it was collected should be confirmed, checks should be made to see if the species was already known in the country (and the watch list updated accordingly, particularly given the list was developed in 2013), and/or field confirmation should potentially trigger an incursion response.

### *1.3 High risk pest list ('inspections list')*

This list, from Saccaggi et al. (2021; available on figshare: Terblanche et al. 2021), contains 218 alien taxa intercepted on plant imports between 1994 and 2019 but not known to occur in South Africa. Taxa on this list includes mainly invertebrates (Phylum Arthropoda and Nematoda), but also viruses, fungi and bacteria. If a species on this list is found to have a DNA barcode collected from samples in South Africa, then the response should be the same as described for the watch list. Like the watch list, this list is not updated regularly, so overlapping species require verification to confirm they have not been reported since 2021.

### *1.4 Plants of Southern Africa (BODATSA)*

BODATSA (also known as NewPOSA; http://posa.sanbi.org/) is a living database curated by the South African National Biodiversity Institute (SANBI), documenting the presence and status (e.g. indigenous, naturalised, invasive) of plant species in South Africa. It excludes cultivated taxa unless they are recorded outside of cultivation. Thus, horticultural species present in BOLD may not be in BODATSA. Besides vascular plants, it includes non-vascular plant groups such as Bryophyta (mosses) and Marchantiophyta (liverworts) but excludes algae. The version used here was downloaded on 6th May 2025 and contained 68,153 rows (SANBI 2025). A recent version should be downloaded when performing this analysis, with the date of download recorded (as specified in R script 2).

If an alien plant is present on BOLD with a DNA barcode but is not listed in BODATSA, the first step is to confirm the validity of the BOLD record by checking whether it includes collection details such as GPS coordinates or a locality name. If the species is only present in South Africa under cultivation, it will not appear in BODATSA—even if it is known to exist in the country. In such cases, monitoring is recommended to ensure cultivated individuals do not spread beyond containment. However, if the BOLD record corresponds to wild individuals and the species has a known history of invasion or negative impacts in other countries, it should be flagged for inclusion in BODATSA and considered for an incursion response. Where appropriate, physical specimens should be collected and curated as evidence of the alien taxon’s presence in the country.

## *2 Workflow*

The workflow is reproducible, transparent, and based on existing templates (Reyserhove et al. 2020; Seebens et al. 2020), with R scripts provided in the Supplementary Material.

### *2.1 Data preparation (R script 1)*

The living database (BOLD) and static datasets (watch list and pest list) were cleaned and stored for reuse. Details are provided below. Note that BODATSA is not subjected to taxonomic standardisation because it serves as the taxonomic backbone for plants in South Africa. As such, it is not included in the first step of data preparation. Instead, it is processed later during the second step (and corresponding script), where it is prepared both for use as the taxonomic reference and for list comparisons in step 3.

#### *2.1.1 BOLD*

To ensure there is molecular evidence of a species’ presence, we retained only BOLD records with DNA sequences. This was assessed using the ‘nucleotides’ column in the “Combined” file export (which includes both Specimen and Sequence data). Records with blank, dashed, or missing values in this column were excluded. Note that presence of a BIN (Barcode Index

Number) indicates sequencing, but some sequenced records lack BINs, so the 'nucleotides' field is more inclusive. Unnecessary columns and duplicated species names were removed, as well as records with no data for species name. The ‘species_name’ column was renamed to ‘verbatimScientificName’ following Darwin Core standards. The list was saved as a .csv file.

The version used initially contained 498,735 rows. After removing 81,761 blank entries, 334 dashed lines, and 9 records from Israel (mislabelled under South Africa), plus the duplicated species, 10,084 entries remained.

### *2.1.2 Watchlist*

From the original dataset (884 taxa), we selected the 400 species designated as "Watch list" in the ‘Final.designation’ column. The ‘Species name’ field was renamed to ‘verbatimScientificName’, unnecessary columns were removed, and the list was saved as a .csv file.

### *2.1.3 Inspection list*

The dataset “Metadata of contaminants on SA plant imports 1994–2019.csv” (Terblanche et al. 2021) was downloaded from Figshare (878 taxa). We selected species recorded as “absent” in South Africa (column ‘SA.occurenceStatus.current’), removed unnecessary columns, and renamed the ‘species’ column to ‘verbatimScientificName’. The list was saved as a .csv file.

## *2.2 Taxonomic standardisation (R script 2)*

To ensure consistent taxonomy, we followed the standardisation workflow developed by Faulkner (2025), which was developed for South African alien species lists (Zengeya et al. 2025). The watch list, inspection list and list from BOLD were taxonomically standardised using this workflow (https://github.com/KatelynFaulkner/rsa-ans-workflow). As per Faulkner (2025), BODATSA was used for plant names and the GBIF backbone was used for the names of other taxa. If a plant taxon was not found in BODATSA, the taxonomic backbone of the World Checklist of Vascular Plants (via POWO) was used. The automated workflow was adapted so that the taxonomic standardisation of the lists was iterative. As described in Faulkner (2025) manual inspection of the outputs of the automated workflow was required to resolve ambiguous cases (e.g., taxonomic homonyms) and ensure clean outputs (Murray et al. 2017, see details in Supplementary material). Duplicates were removed to retain one row per species. The outputs were saved as .csv files, with the column ‘scientificName’ containing the taxonomically standardised taxon names.

As previously mentioned, the BODATSA list is not subjected to taxonomic standardisation because it serves as the taxonomic backbone for plants in South Africa. However, it still requires preparation to be used effectively in this role. We did not remove indigenous or endemic species at this stage to avoid misclassifying native taxa in BOLD as aliens during list comparison. The formatting in some columns was standardised (e.g., double spaces removed). The original dataset splits the scientific name across multiple columns (‘Genus’, ‘Sp1’, ‘Rank1’, ‘Sp2’), which were merged to a single ‘scientificName’ field which contained the scientific name with authorship and date information. Duplicated names in ‘scientificName’ were removed as well as unnecessary columns. Taxa labelled as “misapplied” in the 'TaxStat' column, indicating taxonomic misidentifications or records not considered valid for South Africa, were also excluded from the dataset. The list was saved as a .csv file.

### *2.3 List comparison (R script 3)*

Lists (i.e., scientificName columns) were compared in pairs, with BOLD always serving as the reference. While the comparison between BOLD and the watch list, and BOLD and the inspection list flags species present in both lists, the comparison of BOLD and BODATSA flags species present only in BOLD (i.e., missing from BODATSA).

#### *2.3.1 BOLD vs Watch & Inspection lists*

The complete BOLD list (10,084 records) was compared to the 400 watch list taxa and 185 pest species. Any matches were flagged as potential new detections. Because the watch and pest lists are static, all matches required literature checks to ensure the species had not already been reported in South Africa after 2013 (watch list) or 2021 (pest list).

#### *2.3.2 BOLD vs BODATSA*

For this comparison, only plant taxa from BOLD were used. The aim was to identify species in BOLD not recorded in BODATSA. Such mismatches may represent unrecognised or undocumented plant introductions. Subsequent checks were required to determine if these species are indeed new alien records for South Africa.

## **Results**

### *1 Number of species per species list*

The BOLD dataset downloaded on 22 November 2023 contained 498,735 records for South Africa, of which 11.5% were identified to species level (56,357 records), representing 10,084 species. After taxonomic standardisation, 801 synonyms were identified (verbatim name was a synonym of an accepted name that is included in the scientific name column), 313 duplicates for scientific name, and no duplicates for verbatim name.

The watch list from Faulkner et al. (2014) originally included 400 species. The taxonomic standardisation detected no duplicates for scientific name or verbatim name but encountered 44 synonyms. The high-risk pest list from Saccaggi et al. (2021) comprised 218 high-risk pest species, which was reduced to 185 following taxonomic checks and removal of duplicates, with 5 synonyms, 4 scientific name duplicates and no duplicates for verbatim name.

Finally, the BODATSA list, downloaded on 6th May 2025, originally contained 68,153 records, which were reduced to 67,105 species after data preparation and cleaning (e.g., 1,048 misapplied names were removed).

### *2 Taxonomic standardisations*

The inspection list after going through the steps of data preparation and taxonomic standardisation (scripts 1 and 2) resulted in 37 records (out of 218) that needed to be manually checked before moving to the next step of list comparison. These were ‘possible error flags' related to either uncertain (six), doubtful (one) or unmatched taxon names (30) (see Supplementary material and Faulkner 2025 for more details).

The watchlist processed file after the taxonomic standardisation contained 10 possible error flags to manually check out of 400 records (four uncertain, five doubtful and one unmatched). Finally, the BOLD list with 10,084 records had 1,589 possible error flags to be manually checked (16% of the total records: 193 uncertain, 61 doubtful and 1366 unmatched). Depending on the datasets under study, this step can be highly time-consuming, particularly when numerous flags require manual verification, however, the manual work is greatly reduced my implementing the automated workflow. Unmatched records were predominantly taxa identified only to the genus level, which could not be resolved to species level during taxonomic standardisation. These records were retained at the genus level, as the entire genus could potentially be absent from the country—even if species-level identification was not available. Certain genera are known to include pests or invasive species, so flagging the presence of a new genus in the country was considered important. After all flags were manually verified, we proceeded with the list comparison.

## *3 Lists comparison*

### *3.1 BOLD vs Watch List*

Seven species were shared between BOLD (10,084 species) and the Faulkner et al. (2014) watch list (400 species; Table 1; S2a). These were the fungal pathogen *Austropuccinia psidii*, four invertebrates (*Aedes aegypti*, *Cinara cupressi*, *Culex quinquefasciatus* and *Octolasion tyrtaeum*), the grey wolf *Canis lupus* and a cultivated plant (*Terminalia catappa*).

*Austropuccinia psidii* is a highly invasive fungus affecting plant species within the Myrtaceae family. Known as the causal agent of myrtle rust, it was first detected in South Africa in 2013 (Roux et al. 2013), after Faulkner et al. (2014) had analysed the data for their study and therefore does not represent a new introduction to the country.  (note that in the watchlist, Roux et al. 2013 and in BOLD the species appears with the synonym *Puccinia psidii*). Likewise, *Aedes aegypti* and *Cinara cupressi* were found to be in South Africa after Faulkner et al. (2014) was published (see Guarido et al. 2021 and Wondafrash et al. 2024 respectively). *Canis lupus* is also not flagged because, although there are 54 records in BOLD for this species, they pertain to domestic dogs (*Canis lupus familiaris* or *Canis familiaris*) and not grey wolves, which are the species of concern on the watch list. Nonetheless this highlights the value of the workflow for updating watch lists and similar.

*Culex quinquefasciatus* (Southern house mosquito) is a mosquito species native to Southeast Asia that has dramatically expanded its range over recent centuries, now established throughout much of the tropics and subtropics, particularly in Central America and Oceania (Bhattacharya and Basu 2016; Harvey-Samuel et al. 2021; GBIF 2023a). It is well known for its role as a vector of various pathogens, transmitting diseases such as lymphatic filariasis (*Wuchereria bancrofti*), West Nile virus, St. Louis encephalitis virus, Zika virus, and avian malaria (Bhattacharya and Basu 2016; Harvey-Samuel et al. 2021). Its adaptability allows it to thrive in diverse environments, including urban, peri-urban, and rural areas (Bhattacharya and Basu 2016). Notably, *C. quinquefasciatus* poses a significant threat to island avifauna by facilitating the spread of avian diseases, contributing to population

**Table 1.** Proposed management actions for when molecular evidence of presence is found for alien taxa listed under different check-list types. The numbers of species are based on the South African examples.

| Check-list type | Molecular evidence of presence | Listed on check-list | Proposed action(s) | Number of species in South African example | Example |
|---|---|---|---|---|---|
| National alien species checklist (BODATSA) | TRUE | TRUE | No action needed; species are recorded as present in both databases | 3035 | *Pinus halepensis* |
| | TRUE | FALSE | Confirm validity of BOLD record, incorporate species into checklist, and as necessary collect and curate physical specimen as evidence of the alien taxon's present in the country | 205 from which 131 were found to be alien, 9 flagged | *Ficus microcarpa*, *Jacobaea maritima*, and *Terminalia catappa* |
| | FALSE | TRUE | Arrange for a barcode (or similar) to be collected so there is molecular evidence of the alien taxon's presence in the country | 1,763 | *Colocasia esculenta* |
| | FALSE | FALSE | No action needed; species are not recorded as present in both databases | NA | NA |
| Watch list (Faulkner et al. 2014) | TRUE | TRUE | IMMEDIATE ACTION REQUIRED<br>Species are identified as threats if they were not previously recorded as present in the country. Confirm validity of BOLD record, confirm presence in the country, update watch list, initiate an incursion response | 7 initially but after manual check, 3 flagged | *Culex quinquefasciatus*, *Octolasion tyrtaeum* and *Terminalia catappa* |
| | TRUE | FALSE | No action needed; species are present in the country, but are not identified as significant threats (many are likely native) | 10,077 | *Achroia grisella* |
| | FALSE | TRUE | No action needed; species are identified as threats, but there is no evidence of presence in the country | 393 | *Cronartium ribicola* |
| | FALSE | FALSE | No action needed; species are neither recorded as present nor a threat. | NA | NA |
| Quarantine pest list (Saccaggi et al. 2021) | TRUE | TRUE | IMMEDIATE ACTION REQUIRED<br>Species are identified as threats and were not previously recorded as present in the country. Confirm validity of BOLD record, confirm presence in the country, update quarantine list, initiate an emergency pest response plan | 2 initially but after manual check, 1 flagged | *Cartodere constricta* |
| | TRUE | FALSE | No action needed; species are present in the country, but are not identified as quarantine pests (many are likely native) | 10,082 | *Culex annulioris* |
| | FALSE | TRUE | No action needed; species are identified as quarantine pests, but there is no evidence of presence in the country | 183 | *Aceria tulipae* |
| | FALSE | FALSE | No action needed; species are neither recorded as present nor a threat. | NA | NA |

declines and extinctions of endemic bird species, such as those observed in Hawaii (Harvey-Samuel et al. 2021).

*Octolasion tyrtaeum* is an earthworm species native to Europe that has been introduced to North America, where it is considered invasive in certain ecosystems, particularly in northern forests that historically lacked native earthworm populations (Ransom 2017). Its introduction has led to significant alterations in soil chemistry and structure, impacting nutrient cycling and affecting native plant and animal communities. Its activity influences soil microflora, microarthropod populations, and vegetation dynamics (Eisenhauer et al. 2007). In northern hardwood forests, *O. tyrtaeum* has been associated with shifts in the organic layer of soils, which may indirectly affect native fauna, such as salamander populations in the southern Appalachian Mountains (Ransom 2017). Its presence has also been shown to influence seed bank emergence, benefiting herbaceous species while negatively affecting legumes (Eisenhauer et al. 2009).

*Terminalia catappa*, commonly known as tropical almond, is a tree species native to coastal areas of Southeast Asia and the Pacific (Thomson and Evans 2006). It has been introduced to various tropical and subtropical regions worldwide for various purposes, including ornamental planting, shade, medicinal applications, small-scale timber production, and agroforestry (Ramanan et al. 2024). In some areas, such as parts of the United States, Cuba and Brazil, *T. catappa* is considered invasive due to its capacity to form dense stands that outcompete native vegetation, alter plant community composition, facilitate the development of shade-tolerant species, and disrupt ecological interactions between native species). The two BOLD records for *T. catappa* are from cultivated plants collected in KwaZulu-Natal (Richard's Bay). The species is thus known to be present in the country under cultivation, likely explaining its absence from BODATSA. However, monitoring (and ideally management measures) is needed to prevent its spread into native landscapes.

In summary, three species were flagged through the BOLD–Watch List comparison: (1) *Culex quinquefasciatus* – invertebrate, (2) *Octolasion tyrtaeum* – invertebrate, and (3) *Terminalia catappa* – cultivated plant. The first two species require immediate attention, as they are recognised threats and had not been previously recorded as present in South Africa. For these taxa, the validity of the BOLD record must be verified, their presence in the country confirmed, the national watch list updated accordingly, and an incursion response initiated. For *T. catappa*, which current records indicate is present only in cultivation, ongoing monitoring is recommended to ensure it does not escape and establish in the wild. If monitoring confirms that the species has spread beyond cultivation and formed a naturalised or wild population, it should then be flagged for inclusion in BODATSA and assessed for an incursion response. In such cases, physical specimens should be collected and curated as supporting evidence.

### *3.2 BOLD vs High-Risk Pest List*

Two taxa were shared between the BOLD dataset (10,084 species) and the high-risk pest list (185 species) (Table 1; S2b): the plaster beetle *Cartodere constricta* and the fruit fly *Drosophila immigrans*. After checking the species manually, only one species was flagged: *C. constricta*. According to GBIF (https://www.gbif.org/species/1044381, accessed 11 May 2025), this beetle has been introduced into six countries. However, no evidence of environmental or economic impact has been recorded to date. *Drosophila immigrans* is already known to be present in South Africa (De Araujo et al. 2019), and therefore it is not flagged.

### *3.3 BOLD vs BODATSA*

Comparison of the plant species in BOLD (3,384 species) with the BODATSA list (67,105 species) revealed 349 species unique to BOLD (see Table 1; S2c). Of these, 144 were algae, which are not included in BODATSA, leaving 205 taxa for further assessment: 186 vascular plants (Phylum Tracheophyta), and 19 non-vascular plants [10 mosses (Bryophyta) and 9 liverworts (Marchantiophyta)]. Twenty-six of these 205 taxa are already present in BODATSA and therefore they are not flagged; at least six are native to South Africa (see Suppl. Material S2c). Of the remaining taxa, 47 were determined to be native (according to POWO) but missing from BODATSA, while one was considered data deficient. This left 131 taxa confirmed as alien to South Africa. From this, nine species were flagged: *Combretum indicum*, *Conocarpus erectus*, *Ficus microcarpa*, *Jacobaea maritima*, *Olea europaea* subsp. *europaea*, *Phormium tenax*, *Terminalia catappa* (see above), *Voacanga africana*, and *Yucca gloriosa*.

*Combretum indicum* is a fast-growing, thorny climber native to tropical Asia. It can thrive in diverse conditions, escape cultivation, and naturalise in non-native environments. Although not widely listed as invasive, its potential impacts require further evaluation (CABI 2015). Its BOLD record was collected at Durban Botanical Gardens, suggesting it is currently in cultivation; however, monitoring is needed to ensure it does not ‘jump the fence’ and invade native communities. *Conocarpus erectus*, or button mangrove, is highly invasive in Hawaii, where it colonises mudflats and coastal forests, displacing native vegetation and causing drainage problems (Keighery and Long 2022). Like *Combretum*, its BOLD record was collected at Durban Botanical Gardens, indicating it is cultivated but not yet naturalised in South Africa.

*Ficus microcarpa*, commonly known as Chinese banyan, has an aggressive root system and is considered invasive in urban areas like Sydney (Australia), where it damages infrastructure and alters local ecosystems (Ossola et al. 2023). Its BOLD record was from Robben Island rather than mainland South Africa, and biosecurity measures should be taken to prevent dispersal to the mainland by tourists or staff. *Yucca gloriosa*, native to the southeastern United States, has been introduced to over 20 countries, with evidence of impacts in Japan and New Zealand (GBIF 2023b). It has been rated as high-risk in weed risk assessments in the U.S. (PlantPono 2017). Its BOLD record also came from Robben Island, so similar biosecurity precautions apply. *Phormium tenax*, or New Zealand flax, is problematic in Australia, where it invades freshwater wetlands (Carr et al. 1992), and has been rated as high-risk in the Pacific Islands (PIER 2005). The two BOLD records for this species also came from Robben Island, highlighting the need for strict biosecurity to prevent its spread to the mainland.

*Jacobaea maritima* (dusty miller) is a Mediterranean ornamental species that has naturalised in several regions, including Colombia and Portugal (GBIF 2023c; POWO 2025). In the Crimean Peninsula, it disrupts native ecosystems and is considered a transformer species (Bagrikova and Skurlatova 2021). The BOLD record for this species was mined from GenBank and does not include collection locality coordinates, making it difficult to assess current presence or spread. *Olea europaea* subsp. *europaea*, the European olive, is highly invasive in parts of Australia, where it displaces native plants and animals by forming dense, long-lived stands (Weeds Australia 2024). While likely present in South Africa as a cultivated species—the two BOLD records were collected at Addo Elephant National Park—its status outside cultivation should be verified and formally documented. Note that this species appears on BOLD as *Olea racemosa*, which is considered a synonym of *Olea europaea* subsp. *europaea* by POWO.

*Voacanga africana*, although native to tropical Africa, has been introduced to China and flagged for potential impacts (GBIF 2023d); however, available data are limited and further research on its invasive traits and impacts is needed. Its BOLD record was collected at Lowveld Botanical Gardens, suggesting it is also cultivated. Given the lack of evidence for environmental spread or impact, this species warrants further investigation.

In summary, while several flagged species are confirmed invaders elsewhere, their current status in South Africa varies. Some are present under cultivation (*C. indicum*, *C. erectus*, *O. europaea* subsp. *europaea*, *T. catappa* and *V. africana*) others are recorded only from offshore islands (*F. microcarpa*, *P. tenax* and *Y. gloriosa*), and one lacks locality data (*J. maritima*). All may already be spreading and therefore warrant further scrutiny. For those four not under cultivation, once the validity of the BOLD records is confirmed, they should be incorporated into BODATSA, and physical specimens should be collected and curated as evidence of the alien taxa's presence in the country.

## Discussion

Comparison between the species lists flagged four taxa that warrant further investigation. These include three species from the watch list and one from the list of high-risk quarantine pests. Given their prior designation as species of concern, risk analyses should be prioritised and an immediate response considered. Additional plant taxa were also flagged from the comparison with BODATSA. These records require field verification to determine whether they represent new introductions or are already established but undocumented; in either case these taxa should be evaluated for potential inclusion in BODATSA.

Importantly, not all flagged species necessarily represent novel incursions. Taxonomic inconsistencies, synonymy, or unresolved native versus alien status can lead to apparent mismatches across datasets (Pyšek et al. 2013; Essl et al. 2018). This highlights the importance of treating each flagged species as a candidate for expert scrutiny, rather than assuming it to be a new alien species. Misidentification or misclassification can generate false positives, and careful interpretation is necessary before initiating management actions (Cristescu and Hebert 2018; Hayes 2021). In this context, we will notify BODATSA of all species that appear as native to South Africa in POWO but are not listed in BODATSA, to support efforts toward achieving a taxonomic consensus. Discrepancies between authoritative sources such as POWO and BODATSA present a significant hurdle for end users, complicating the interpretation of presence and nativity. Resolving such inconsistencies is essential to improve the reliability of species checklists and support evidence-based biosecurity responses.

The decision to remove misapplied names from BODATSA (>1000 records) was deliberate to prevent confusion during the comparison process. Misapplied names represent instances where species were incorrectly identified in the past, leading to false assumptions about their presence in the country. By eliminating these records, the workflow ensures that any flagged species appearing in BOLD for South Africa can be confidently considered as potential incursions or established populations. For example, the name *Acacia retinoides* was historically misapplied to naturalised populations of *A. provincialis* in South Africa, a discrepancy that was only resolved relatively recently (Magona et al. 2018). Another decision that needed to be made was how to handle taxa with ambiguous identifications—such as those labelled with *aff.* (affinis), *cf.* (confer), or *nr* (near). These were manually reviewed during the standardisation process. Such qualifiers typically indicate uncertainty at the species level, suggesting that the specimen is similar to, but not confidently assignable as, a particular species. To reflect this uncertainty and avoid incorrectly flagging an entire genus, these records were standardised to the format *Genus sp.* This approach ensures that any potential flagging or follow-up actions remain appropriately focused on a single, unidentified species within the genus, rather than generalising to the entire taxonomic group.

For taxa detected exclusively through eDNA, it is crucial to establish evidence of physical presence before assuming incursion. Burgess et al. (2021), in their study on *Phytophthora* in Australia, emphasised that while eDNA and metabarcoding are valuable for species detection, physical verification remains necessary. They recommend that eDNA detections be followed up with traditional isolation methods to confirm the presence of live organisms. In alignment with ISPM8 guidelines (Anonymous 2017), taxa detected solely by molecular methods should not be considered present without accompanying biological confirmation. This precaution guards against false positives that could mislead biosecurity efforts (Darling and Mahon 2011; Rees et al. 2014). Extending this principle beyond *Phytophthora* to other taxa is essential for ensuring robust detection and response strategies.

The proposed workflow offers a practical tool for bridging the gap between molecular detection and management response. By integrating different types of species lists and standardising taxonomic information, it fosters more consistent communication between researchers, biosecurity agencies, and decision-makers (Gatto et al. 2013; Pagad et al. 2018; Sandall et al. 2023). This standardised approach enables more efficient triage of flagged taxa and encourages routine surveillance using molecular records (van Rees et al. 2022; Pocock et al. 2024).

We recommend applying this workflow at least annually, given that the BOLD database is regularly updated with new records from diverse molecular studies (Ratnasingham et al. 2024). Although many of these studies are not focused specifically on alien species, they often generate data with incidental biosecurity value (Comtet et al. 2015). Regular application of the workflow ensures that newly deposited records are not overlooked, and it strengthens the early warning capacity of national biosecurity systems.

An important consideration when using BOLD as a detection tool is its reliance on complete metadata for optimal effectiveness. For BOLD records to be truly informative, it is crucial that users provide detailed metadata, including the collection location, date, and any available images. These supplementary data fields significantly enhance field verification and incursion response efforts. Pitcher et al. (2024) emphasise that BOLD records are particularly valuable when accompanied by comprehensive metadata, underscoring the need for better curation practices and the encouragement of metadata completeness.

Challenges remain, particularly regarding invertebrate detection. Invertebrates are often underrepresented in barcoding databases, not only in South Africa but globally (Pitcher et al. 2024). Boykin et al. (2012) discuss how the lack of comprehensive barcoding for invasive insects poses risks to biosecurity. Addressing these gaps requires ongoing projects like the QBOL project in Europe, which focuses on gathering DNA barcoding data of quarantine organisms (see Bonants et al. 2010), and multiple efforts from South Africa. For example, in 2011, the International Barcode of Life Project (iBOL) node was established in South Africa, followed by the launch of the Urban DNA Barcode Project in KwaZulu-Natal. Additionally, funding mechanisms like the National Research Foundation's Foundational Biodiversity Information Programme have supported these developments (da Silva and Willows-Munro 2016; Pitcher et al. 2024). These projects have made significant progress in expanding barcoding efforts for invasive and native species alike. However, continuous funding and support for DNA barcoding remain critical to sustain these initiatives and enhance detection capabilities.

Even when clear evidence of invasive species presence is available—such as in the case of the PSHB in South Africa—mobilising a rapid and coordinated response can be difficult. Decision-makers and the public often require visible or large-scale impacts before taking action, by which point management options may be limited (Simberloff et al. 2013). A similar delay was observed in Australia during the early spread of myrtle rust (*Austropuccinia psidii*), despite warnings from scientists (Carnegie and Pegg 2018). This tendency to delay action until impacts are widely visible illustrates the socio-political and economic barriers to effective incursion response (Hester and Bland 2024).

Finally, it is important to acknowledge the limitations of this approach. Cryptic species may go undetected, and the method is constrained to taxa with DNA barcode representation (Beng and Corlett 2020). It is not a substitute for preventative measures, but it complements them by enhancing incursion response (Wilson et al. 2017; Clarke et al. 2023). Once a species has entered the country, rapid identification and assessment are critical to reduce the risk of spread, and workflows like the one we propose can support those efforts. For the workflow to be effective in practice it would need to be institutionalised and included as part of standard operating procedures. Surveillance operations need to be in place, the data gathered through such exercises need to be picked up and applied to this workflow, and if risks are flagged there is a clear reporting chain to someone who can initiate an incursion response. In this sense we believe the workflow is a tool to assist with integrated governance (Roy et al. 2024).

## Acknowledgements

Francois Roets and Davina Saccaggi for useful discussion on the first steps of the paper when looking for lists of high-risk pests.

## Funding

This work was supported by the Centre for Invasion Biology (CIB) at Stellenbosch University. The South African Department of Forestry, Fisheries and the Environment (DFFE) are thanked for funding, noting that this publication does not necessarily represent the views or opinions of DFFE or its employees. KF and JW also received funding from the B-cubed project (Biodiversity Building Blocks for policy) which is funded by the European Union’s Horizon Europe Research and Innovation Programme (ID No 101059592). Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them. TP was funded by a University of Pretoria Postdoctoral Research Fellowship.

## Author contributions

T.P. and J.R.U.W. conceived the idea; L.F.W. obtained the species lists, performed preliminary analysis and led the writing of the manuscript; K.T.F. developed the R scripts. All coauthors provided insightful advice, valuable feedback on the manuscript and agreed to submit the final version of the manuscript.

## Competing interests

The authors have declared that no competing interests exist.

## Author ORCIDs

Laura Fernández Winzer https://orcid.org/0000-0003-0675-9117

Katelyn T. Faulkner https://orcid.org/0000-0002-3955-353X

Trudy Paap https://orcid.org/0000-0003-1364-4350

John R. U. Wilson https://orcid.org/0000-0003-0174-3239

## Data availability

All the data that support the findings of this study and the R scripts underpinning the analysis reported in this paper are available either in the Supplementary Information or upon request from the corresponding author.

## References

Armstrong KF, Ball SL (2005) DNA barcodes for biosecurity: invasive species identification. Philosophical Transactions of the Royal Society B: Biological Sciences 360: 1813–1823. https://doi.org/10.1098/rstb.2005.1713

Baena-Bejarano N, Reina C, Martínez Revelo DE, Medina CA, Tovar E, Uribe-Soto S, et al. (2023) Taxonomic identification accuracy from BOLD and GenBank databases using over a thousand insect DNA barcodes from Colombia. PLoS ONE 18(4): e0277379. https://doi.org/10.1371/journal.pone.0277379

Bagrikova NA, Skurlatova MV (2021) The Materials to the "Black Book" of the Flora of the Crimean Peninsula. Russian Journal of Biological Invasions 12: 244–257. https://doi.org/10.1134/S2075111721030036

Beng KC, Corlett RT (2020) Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. Biodiversity and Conservation 29: 2089–2121. https://doi.org/10.1007/s10531-020-01980-0

Bhattacharya S, Basu P (2016) The southern house mosquito, *Culex quinquefasciatus*: profile of a smart vector. Journal of Entomology and Zoology Studies 4(2): 73-81.

Bonants P, Groenewald E, Rasplus JY, Maes M, De Vos P, Frey J, Boonham N, Nicolaisen M, Bertacini A, Robert V, Barker I (2010) QBOL: a new EU project focusing on DNA barcoding of quarantine organisms. Eppo Bulletin 40(1): 30-33. https://doi.org/10.1111/j.1365-2338.2009.02350.x

Boykin LM, Armstrong KF, Kubatko L, De Barro P (2012) Species delimitation and global biosecurity. Evolutionary Bioinformatics 8: 1–37. https://doi.org/10.4137/EBO.S8532

Burgess TI, Edwards J, Drenth A, Massenbauer T, Cunnington J, Mostowfizadeh-Ghalamfarsa R, Dinh Q, Liew EC, White D, Scott P, Barber PA (2021) Current status of *Phytophthora* in Australia. Persoonia - Molecular Phylogeny and Evolution of Fungi. 47: 151–177. https://doi.org/10.3767/persoonia.2021.47.05.

Brown MJM, Walker BE, Black N, Govaerts RHA, Ondo I, Turner R, Nic Lughadha E (2023) rWCVP: a companion R package for the World Checklist of Vascular Plants, *New Phytologist* 240(4): 1355–1365.

CABI (2015) *Combretum indicum* (Rangoon creeper), CABI Digital Library, Accessed 09 June 2025, https://doi.org/10.1079/cabicompendium.46654

Carnegie AJ, Pegg GS (2018) Lessons from the incursion of myrtle rust in Australia. Annual review of Phytopathology 56: 457-478. https://doi.org/10.1146/annurev-phyto-080516-035256

Carr GW, Yugovic JV, Robinson KE (1992) Environmental weed invasions in Victoria: conservation and management implications. Dept. of Conservation and Environment and Ecological Horticulture Pty Ltd. In: Agriculture Victoria (2020) New Zealand flax (*Phormium tenax*), Invasive Plants, Victorian Resources Online, Accessed 09 June 2025, https://vro.agriculture.vic.gov.au/dpi/vro/vrosite.nsf/pages/weeds_new-zealand-flax

Chamberlain S, Barve V, Mcglinn D, Oldoni D, Desmet P, Geffert L, Ram K (2025) *rgbif: interface to the Global Biodiversity Information Facility API, R package version 3.8.0.*

Clarke LJ, Shaw JD, Suter L, Atalah J, Bergstrom DM, Biersma E,Convey P, Greve M, Holland O, Houghton MJ, Hughes KA, Johnston EL, King CK, McCarthy AH, McGaughran A,Pertierra LR, Robinson SA, Sherman CDH, Stark JS, Stevens MI, Strugnell JM, von Ammon U, Wilson NG, Zaiko A, MacDonald AJ (2023) An expert-driven framework for applying eDNA tools to improve biosecurity in the Antarctic. Management of Biological Invasions 14(3): 379–402, https://doi.org/10.3391/mbi.2023.14.3.01

Comia CF, Morris MR (2024) Beyond BOLD: using DNA barcoding to identify Prussian carp (*Carassius gibelio* Bloch, 1782) in southern Alberta, Canada. BioInvasions Record. 13(2): 497–513. https://doi.org/10.3391/bir.2024.13.2.15

Comtet T, Sandionigi A, Viard F, Casiraghi M (2015) DNA (meta) barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. Biological Invasions 17: 905-22. https://doi.org/10.1007/s10530-015-0854-y

Cristescu ME, Hebert PD (2018) Uses and misuses of environmental DNA in biodiversity science and conservation. Annual Review of Ecology, Evolution, and Systematics. 49(1): 209-230. https://doi.org/10.1146/annurev-ecolsys-110617-062306

Csárdi, G., Hester, J., Wickham, H., Chang, W., Morgan, M. & Tenenbaum, D., 2024, *remotes: R package installation from remote repositories, including 'GitHub', R package version 2.5.0.*

da Silva JM, Willows-Munro S (2016) A review of over a decade of DNA barcoding in South Africa: a faunal perspective. African Zoology 51: 1-12. https://doi.org/10.1080/15627020.2016.1151377

Darling JA, Blum MJ (2007) DNA-based methods for monitoring invasive species: a review and prospectus. Biological Invasions 9: 751–765. https://doi.org/10.1007/s10530-006-9079-4

Darling JA, Mahon AR (2011) From molecules to management: adopting DNA-based methods for monitoring biological invasions in aquatic environments. Environmental research. 111: 978-988. https://doi.org/10.1016/j.envres.2011.02.001

De Araujo LI, Karsten M, Terblanche JS (2019) Three new Drosophilidae species records for South Africa. Bothalia-African Biodiversity & Conservation 49(1): 1-5, https://doi.org/10.4102/abc. v49i1.2429

de Wit MP, Crookes DJ, Blignaut JN, de Beer ZW, Paap T, Roets F, van der Merwe C, van Wilgen BW, Richardson DM (2022) An assessment of the potential economic impacts of the invasive Polyphagous Shot Hole Borer (Coleoptera: Curculionidae) in South Africa. Journal of Economic Entomology 115: 1076-1086. https://doi.org/10.1093/jee/toac061

Eisenhauer N, Partsch S, Parkinson D, Scheu S (2007) Invasion of a deciduous forest by earthworms: changes in soil chemistry, microflora, microarthropods and vegetation.

Soil Biology and Biochemistry 39(5): 1099-110. https://doi.org/10.1016/j.soilbio.2006.12.019

Eisenhauer N, Straube D, Johnson EA, Parkinson D, Scheu S (2009) Exotic ecosystem engineers change the emergence of plants from the seed bank of a deciduous forest. Ecosystems 12: 1008-16. https://doi.org/10.1007/s10021-009-9275-z

Essl F, Bacher S, Genovesi P, Hulme PE, Jeschke JM, Katsanevakis S, ... Richardson DM (2018) Which taxa are alien? Criteria, applications, and uncertainties. BioScience 68: 496–509.

Faulkner KT (2025) An automated workflow to standardise taxon names for South African alien species lists (under review)

Faulkner KT, Robertson MP, Rouget M, Wilson JR (2014) A simple, rapid methodology for developing invasive species watch lists. Biological Conservation 179: 25–32.

Gatto F, Katsanevakis S, Vandekerkhove J, Zenetos A, Cardoso AC (2013) Evaluation of online information sources on alien species in Europe: the need of harmonization and integration. Environmental Management 51: 1137-1146. https://doi.org/10.1007/s00267-013-0042-8

GBIF, GBIF Secretariat. GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei Accessed via GBIF.org on 2025-05-06.

GBIF Secretariat (2023a) *Culex quinquefasciatus* Say, 1823. GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei accessed via GBIF.org on 2025-05-19.

GBIF Secretariat (2023b) *Ficus microcarpa* L.fil. GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei accessed via GBIF.org on 2025-06-11.

GBIF Secretariat (2023c) *Jacobaea maritima* (L.) Pelser & Meijden. GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei accessed via GBIF.org on 2025-05-19.

GBIF Secretariat (2023d) *Voacanga africana* Stapf ex Scott Elliot. GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei accessed via GBIF.org on 2025-06-09.

Govaerts R (2024) *WCVP: World Checklist of Vascular Plants. Facilitated by the Royal Botanic Gardens, Kew, version 13*.

Govaerts R, Nic Lughadha E, Black N, Turner R, Paton A (2021) ‘The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity’, *Scientific Data* 8(1): 215.

Grolemund G, Wickham H (2011) ‘Dates and times made easy with lubridate’, *Journal of Statistical Software* 40(3): 1–25.

Guarido MM, Riddin MA, Johnson T, *et al.* (2021) *Aedes* species (Diptera: Culicidae) ecological and host feeding patterns in the north-eastern parts of South Africa, 2014–2018. *Parasites Vectors* 14: 339. https://doi.org/10.1186/s13071-021-04845-9

Hamelin RC (2012) Contributions of genomics to forest pathology. Canadian Journal of Plant Pathology 34: 20–28.

Hanner RH, Lima J, Floyd R (2009) DNA barcoding and its relevance to pests, plants and biological control. ISHS Acta Horticulturae 823: 41–48. https://doi.org/10.17660/ActaHortic.2009.823.3.

Harvey-Samuel T, Ant T, Sutton J, Niebuhr CN, Asigau S, Parker P, Sinkins S, Alphey L. (2021) *Culex quinquefasciatus*: status as a threat to island avifauna and options for genetic control. CABI Agriculture and Bioscience 2(9): 1-21. https://doi.org/10.1186/s43170-021-00030-1

Hayes KA (2021) Taxonomic shortcuts lead to long delays in species discovery, delineation, and identification. Biological Invasions 23: 1285–1292. https://doi.org/10.1007/s10530-020-02438-8

Hernández-Triana LM, Brugman VA, Nikolova NI, Ruiz-Arrondo I, Barrero E, Thorne L, de Marco MF, Krüger A, Lumley S, Johnson N, Fooks AR (2019) DNA barcoding of British mosquitoes (Diptera, Culicidae) to support species identification, discovery of cryptic genetic diversity and monitoring invasive species. ZooKeys 19: 832:57. https://doi.org/10.3897/zookeys.832.32257

Hester SM, Bland LM (2024) Prepare, Respond, and Recover: Selecting Immediate and Long-Term Strategies to Manage Invasions. In: Biosecurity. Taylor & Francis. https://doi.org/10.1201/9781003253204-10

IPBES (2023) Thematic Assessment Report on Invasive Alien Species and their Control of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Roy HE, Pauchard A, Stoett P, Renard Truong T (eds.). IPBES secretariat, Bonn, Germany. https://doi.org/10.5281/zenodo.7430682

Keighery GJ, Long V (2022) The Button Mangrove *Conocarpus erectus* (Combretaceae) is naturalised in Western Australia, Department of Biodiversity, Conservation and Attractions. Nuytsia, The journal of the Western Australian Herbarium 33: 29-33 https://florabase.dbca.wa.gov.au/science/nuytsia/1020.pdf

Liebhold AM, Berec L, Brockerhoff EG, Epanchin-Niell RS, Hastings A, Herms DA, Kean JM, McCullough DG, Suckling DM, Tobin PC, Yamanaka T (2016) Eradication of invading insect populations: from concepts to applications. Annual review of entomology 61(1): 335-52. https://doi.org/10.1146/annurev-ento-010715-023809

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: e0222291. https://doi.org/10.1371/journal.pone.0222291

Magona N, Richardson DM, Le Roux JJ, Kritzinger-Klopper S, Wilson JR (2018) Even well-studied groups of alien species might be poorly inventoried: Australian *Acacia* species in South Africa as a case study. NeoBiota 39: 1–29. https://doi.org/10.3897/neobiota.39.23135

Marciniak B, Machado LP, de Campos LLF, Hirota M, Dechoum MS (2022) Direct and indirect effects of an invasive non-native tree on coastal plant communities. Plant Ecology 223: 935–949. https://doi.org/10.1007/s11258-022-01246-5

Murray BR, Martin LJ, Phillips ML, Pyšek P (2017) Taxonomic perils and pitfalls of dataset assembly in ecology: a case study of the naturalized Asteraceae in Australia. NeoBiota 34: 1–20. https://doi.org/10.3897/neobiota.34.11139

Ossola A, Yu M, Le Roux J, Bustamante H, Uthayakumaran L, Leishman M (2023) Research note: Integrating big data to predict tree root blockages across sewer networks. Landscape and Urban Planning: 240: 104892. https://doi.org/10.1016/j.landurbplan.2023.104892

Paap T, De Beer ZW, Migliorini D, Nel WJ, Wingfield MJ (2018) The polyphagous shot hole borer (PSHB) and its fungal symbiont *Fusarium euwallaceae*: a new invasion in South Africa. Australasian Plant Pathology 47: 231–237.

Paap T, Wingfield MJ, De Beer ZW, Roets F (2020) Lessons from a major pest invasion: The polyphagous shot hole borer in South Africa. South African Journal of Science 116. https://doi.org/10.17159/sajs.2020/8757

Pagad S, Genovesi P, Carnevali L, Schigel D, McGeoch MA (2018) Introducing the global register of introduced and invasive species. Scientific data 23: 1-2. https://doi.org/10.1038/sdata.2017.202

PIER (2005) *Phormium tenax*, Risk assessment, Pacific Island Ecosystems at Risk, Accessed 09 June 2025, http://www.hear.org/pier/wra/pacific/phormium_tenax_htmlwra.htm

Pitcher T, Ndaba A, Jacobs A, Hamer M, Janion-Scheepers C (2024) DNA barcoding of alien invertebrates and biological control agents in South Africa: a review. African Entomology 32: e19290. https://doi.org/10.17159/2254-8854/2024/a19290

Plant Pono (2017) *Yucca gloriosa*, risk assessment, Accessed 09 June 2025, https://plantpono.org/hpwra/yucca-gloriosa/

R Core Team (2024) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Plant Pono (2017) *Yucca gloriosa*, Risk Assessment, accessed 09 June 2025, https://plantpono.org/hpwra/yucca-gloriosa/

Pocock MJ, Adriaens T, Bertolino S, Eschen R, Essl F, Hulme PE, Jeschke JM, Roy HE, Teixeira H, De Groot M (2024) Citizen science is a vital partnership for invasive alien species management and research. IScience. 19: 27(1). https://doi.org/10.1016/j.isci.2023.108623

POWO (2025) *Jacobaea maritima* (L.) Pelser & Meijden, Plants of the World Online, Kew Royal Botanical Gardens, https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:77075194-1 [accessed 19 May 2025]

Pyšek P, Hulme PE, Meyerson LA, Smith GF, Boatwright JS, Crouch NR, Figueiredo E, Foxcroft LC, Jarošík V, Richardson DM, Suda J, Wilson JRU (2013) Hitting the right target: Taxonomic challenges for, and of, plant invasions. AoB Plants 5(0): plt042. https://doi.org/10.1093/aobpla/plt042

R Core Team (2024) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ramanan SS, Arunachalam A, Singh R, Verdiya A (2024) Tropical almond (*Terminalia catappa*): A holistic review. Heliyon. 11: e41115 https://doi.org/10.1016/j.heliyon.2024.e41115

Ransom TS (2017) Local distribution of native and invasive earthworms and effects on a native salamander. *Population Ecology* 59: 189–204. https://doi.org/10.1007/s10144-017-0578-1

Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7: 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x.

Ratnasingham S, Hebert PDN (2013) A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. PLoS ONE 8(8): e66213. https://doi.org/10.1371/journal.pone.0066213

Ratnasingham S, Wei C, Chan D, Agda J, Agda J, Ballesteros-Mejia L, Ait Boutou H, El Bastami Z M, Ma E, Manjunath R, Rea D, Ho C, Telfer A, McKeowan J, Rahulan M, Steinke C, Dorsheimer J, Milton M, Hebert PDN (2024) BOLD v4: A Centralized Bioinformatics Platform for DNA-Based Biodiversity Data. In DNA Barcoding: Methods and Protocols, pp. 403-441. Chapter 26. New York, NY: Springer US.

Reaser JK, Burgiel SW, Kirkey J, Brantley KA., Veatch SD, Burgos-Rodríguez J (2020) The early detection of and rapid response (EDRR) to invasive species: a conceptual framework and federal capacities assessment. Biological Invasions 22: 1–19.

Rees HC, Maddison BC, Middleditch DJ, Patmore JR, Gough KC (2014) The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. Journal of applied ecology. 51: 1450-9. https://doi.org/10.1111/1365-2664.12306

Reyserhove L, Desmet P, Oldoni D, Adriaens T, Strubbe D, Davis AJ, ... & Groom Q (2020) A checklist recipe: making species data open and FAIR. Database, https://doi.org/10.1093/database/baaa084

Roux J, Greyling I, Coutinho TA, Verleur M, Wingfeld MJ (2013) The Myrtle rust pathogen, *Puccinia psidii*, discovered in Africa. Ima Fungus 4: 155–159. https://doi.org/10.5598/imafungus.2013.04.01.14

Roy HE, Pauchard A, Stoett PJ, Renard Truong T, Meyerson LA, Bacher S, Galil BS, Hulme PE, Ikeda T, Kavileveettil S, McGeoch MA, Nuñez MA, Ordonez A, Rahlao SJ, Schwindt E, Seebens H, Sheppard AW, Vandvik V, Aleksanyan A, Ansong M, August T, Blanchard R, Brugnoli E, Bukombe JK, Bwalya B, Byun C, Camacho-Cervantes M, Cassey P, Castillo ML, Courchamp F, Dehnen-Schmutz K, Zenni RD, Egawa C, Essl F, Fayvush G, Fernandez RD, Fernandez M, Foxcroft LC, Genovesi P, Groom QJ, González AI, Helm A, Herrera I, Hiremath AJ, Howard PL, Hui C, Ikegami M, Keskin E, Koyama A, Ksenofontov S, Lenzner B, Lipinskaya T, Lockwood JL, Mangwa DC, Martinou AF, McDermott SM, Morales CL, Müllerová J, Mungi NA, Munishi LK, Ojaveer H, Pagad SN, Pallewatta NPKTS, Peacock LR, Per E, Pergl J, Preda C, Pyšek P, Rai RK, Ricciardi A, Richardson DM, Riley S, Rono BJ, Ryan-Colton E, Saeedi H, Shrestha BB, Simberloff D, Tawake A, Tricarico E, Vanderhoeven S, Vicente J, Vilà M, Wanzala W, Werenkraut V, Weyl OLF, Wilson JRU, Xavier RO, Ziller SR (2024) Curbing the major and growing threats from invasive alien species is urgent and achievable. Nature Ecology & Evolution 8: 1216–1223. doi:10.1038/s41559-024-02412-w

SANBI (2025) South African National Biodiversity Institute, Botanical Database of Southern Africa (BODATSA) [https://posa.sanbi.org/sanbi/Explore], downloaded 06 May 2025

Sandall EL, Maureaud AA, Guralnick R, McGeoch MA, Sica YV, Rogan MS, Booher DB, Edwards R, Franz N, Ingenloff K, Lucas M (2023) A globally integrated structure of taxonomy to support biodiversity science and conservation. Trends in Ecology & Evolution 38(12): 1143–53.

Saccaggi DL, Arendse M, Wilson JR, Terblanche JS (2021) Contaminant organisms recorded on plant product imports to South Africa 1994–2019. Scientific data, 8(1), 83. https://doi.org/10.1038/s41597-021-00869-z

Seebens H, Clarke DA, Groom Q, Wilson JR, García-Berthou E, Kühn I, ... & McGeoch M (2020) A workflow for standardising and integrating alien species distribution data. NeoBiota 59: 39–59. https://doi.org/10.3897/neobiota.59.53578

Sethusa MT, Yessoufou K, Van der Bank M, Van der Bank H, Millar IM, Jacobs A (2014) DNA barcode efficacy for the identification of economically important scale insects (Hemiptera: Coccoidea) in South Africa. African Entomology.22(2): 257-66. https://doi.org/10.4001/003.022.0218

Simberloff D, Martin JL, Genovesi P, Maris V, Wardle DA, Aronson J, Courchamp F, Galil B, García-Berthou E, Pascal M, Pyšek P (2013) Impacts of biological invasions: what's what and the way forward. Trends in ecology & evolution 28(1): 58-66.

Terblanche J, Saccaggi DL, Arendse M, Wilson JRU (2021) Contaminant organisms recorded on plant product imports to South Africa 1994-2019. figshare. Collection. https://doi.org/10.6084/m9.figshare.c.5180681.v1

Thomson LAJ, Evans B (2006) *Terminalia catappa* (tropical almond), ver. 2.2. In: Elevitch, CR (ed.). Species Profiles for Pacific Island Agroforestry. Permanent Agriculture Resources (PAR), Hōlualoa, Hawai‘i. <http://www.traditionaltree.org>. https://www.researchgate.net/publication/228374099_Terminalia_catappa_Tropical_Almond [accessed May 19 2025].

Townsend G, Hill M, Hurley BP, Roets F (2025) Escalating threat: increasing impact of the polyphagous shot hole borer beetle, *Euwallacea fornicatus*, in nearly all major South African forest types. Biological Invasions 27: 88. https://doi.org/10.1007/s10530-025-03551-2

Van der Loo MPJ (2014) The stringdist package for approximate string matching, *The R Journal* 6: 111–122.

van Rees CB, Hand BK, Carter SC, Bargeron C, Cline TJ, Daniel W, Ferrante JA, Gaddis K, Hunter ME, Jarnevich CS, McGeoch MA (2022) A framework to integrate innovations in invasion science for proactive management. Biological Reviews 97(4): 1712-1735.

van Rooyen E, Paap T, De Beer W, Townsend G, Fell S, Nel WJ, Morgan S, Hill M, Gonzalez A, Roets F (2021) The polyphagous shot hole borer beetle: current status of a perfect invader in South Africa. South African Journal of Science 10: 117(11-12). https://doi.org/10.17159/ sajs.2021/9736

Van Wilgen BW, Measey J, Richardson DM, Wilson JR, Zengeya TA (2020) Biological Invasions in South Africa. Springer Nature. 975p.

Weeds Australia (2024) European Olive, *Olea europaea subsp. europaea*, Centre for Invasive Species Solutions, Accessed 09 June 2025, https://weeds.org.au/profiles/european-olive/

Wickham H (2023) *stringr: simple, consistent wrappers for common string operations, R package version 1.5.1*.

Wickham H, Henry L (2023) *purrr: functional programming tools, R package version 1.0.2*.

Wickham H, François R, Henry L, Müller KDV (2023) *dplyr: a grammar of data manipulation, R package version 1.1.4*.

Wickham H, Vaughan D, Girlich M (2024) tidyr: tidy messy data, R package version 1.3.1.

Wilson JR, Panetta FD, Lindgren C (2017) Detecting and responding to alien plant incursions. Cambridge University Press.

Wondafrash M, Wingfield MJ, Hurley BP, Slippers B, Mutitu EK, Jenya H, Paap T (2024) DNA sequence data confirms the presence of two closely related cypress-feeding aphid species on African cypress (*Widdringtonia* spp.) in South Africa. *Southern Forests: A Journal of Forest Science* 86(4): 278–285. https://doi.org/10.2989/20702620.2024.2390863

Yang B, Zhang Z, Yang CQ, Wang Y, Orr MC, Wang H, Zhang AB (2022) Identification of species by combining molecular and morphological data using convolutional neural networks. Systematic Biology 71(3): 690-705. https://doi.org/10.1093/sysbio/syab076

Zengeya TA, Faulkner KT, Mtileni MM, Winzer LL, Kumschick S, McCulloch-Jones EJ, Miza-Tshangana SA, Robinson TB, Sifuba A, Engelbrecht W, van Wilgen BW (2025) A checklist of alien taxa for South Africa. bioRxiv, 2025-05. https://doi.org/10.1101/2025.05.22.655507

## Supplementary material

S1. Supplementary material – Notes containing additional methodological details and R scripts.

S2 a, b, c. R outputs - Three Excel spreadsheets containing the results of the list comparisons.