‘Aliens in Europe’

ID 420

Sophie Gombeer[1], Kenny Meganck[2], Yoo Ree Van Bourgonie[1], Nathalie Smitz[2], Marc De Meyer[2], Thierry Backeljau[1,3]

[1] Royal Belgian Institute of Natural Sciences, Barcoding of Organisms and Tissues of Policy Concern (BopCo), Vautierstraat 29, B-1000 Brussels, Belgium.
[2] Royal Museum for Central Africa, Barcoding of Organisms and Tissues of Policy Concern (BopCo), Leuvensesteenweg 13, B-3080 Tervuren, Belgium.
[3] Evolutionary Ecology Group, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium.
* Corresponding author e-mail: sophie.gombeer@naturalsciences.be

ABSTRACT

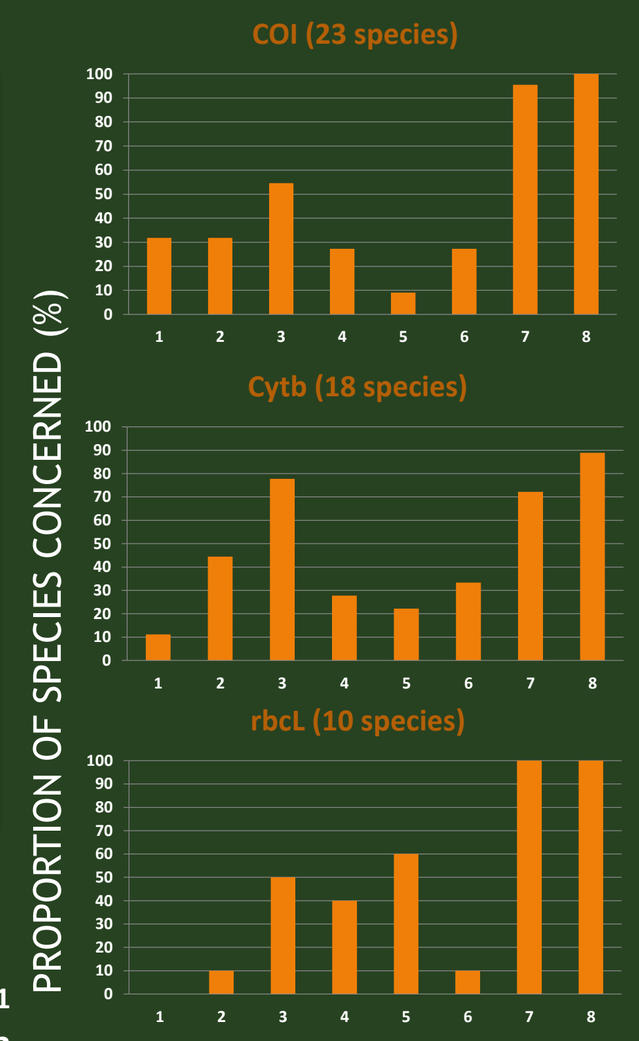
Background: An increasing number of species are being introduced in Europe, whether by accident or deliberately. Some are able to establish viable populations and may outcompete other species or disrupt ecosystem functioning: these species are called ‘invasive alien species’ (IAS). In order to (i) protect native biodiversity and ecosystem services, and (ii) mitigate potential impacts on human health and socio-economical activities, the European Commission issued Regulation 1143/2014, reporting on 37 IAS. The Regulation foresees in three types of interventions: (i) prevention, (ii) early detection and rapid eradication, and (iii) management of established populations. Aside from compiling this list and gathering information on presence, distribution, ecology, impacts and management, accurate methods for rapid identification are required when suspicious biological material is being encountered. In cases where a morphological identification is problematic (ex. cryptic species, trace material), DNA-based identifications may represent an alternative method. The purpose of the present work is, therefore, to investigate and evaluate the available molecular identification techniques for each IAS ‘*in silico*’. **Results:** We investigated the usefulness and accurateness of BOLD (COI for animals; *rbcL*, ITS and *matK* for plants) and encountered some limits when using BOLD barcodes as only tool for species identification. Knowledge gaps regarding (i) the sequence coverage of the IAS and their sister species, and (ii) the metadata on the vouchers (ex. subspecies, locality) have been detected, which can hamper reliable identification. We therefore investigated the feasibility to complement the BOLD database, as well as the usefulness of complementary markers and methods (e.g. RFLP) in providing reliable and rapid identifications. **Significance:** The present project aims to provide an up-to-date status on the molecular tools and methods available for rapid and accurate identification of IAS, as well as to optimize and complement them whenever necessary.

MATERIAL, METHODS AND RESULTS

DNA-database compilation and evaluation method

The present EU regulation includes: 3 bird, 9 mammal, 6 crustacean, 1 insect, 2 fish, 1 reptile, 1 amphibian and 14 plant species. All available DNA sequences for the 37 target IAS and their congeners were retrieved from the online repositories BOLD and GenBank. After preliminary filtering and alignment steps, NJ trees were reconstructed (500BT, Jukes-Cantor distance model) for each marker with sufficient material. To evaluate their capacity at providing a reliable species ID, we classified the different potential issues encountered into eight categories. Each marker of each species was then evaluated based on these criteria:

(1) Taxonomic issue of the target species;
(2) Less than 5 DNA sequences available for the target species;
(3) Poor geographical coverage (native or invasive range missing or scarce);
(4) Non-recovery or unsupported cluster (< 85 bootstrap value) of the target species;
(5) Low overall DNA sequence genetic variation (target and congeners);
(6) Potential species misidentification of a voucher specimen;
(7) Incomplete representation of congener species in the repositories;
(8) Less than 3 DNA sequences available per congeneric species.

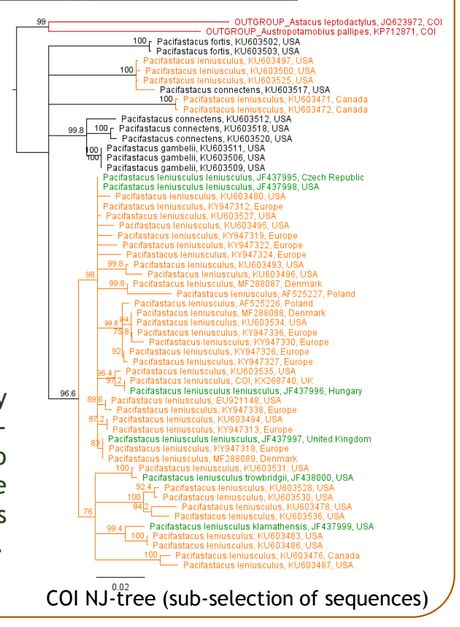


Some examples:

(Cat. 1,4,6,8) Non-recovery of the target species as supported cluster:

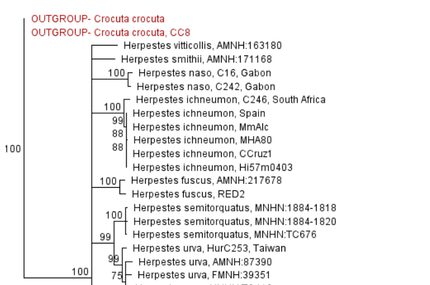
Pacifastacus leniusculus

Larsen *et al*. (2012, 2016) found substantial cryptic diversity within *P. leniusculus* and demonstrated that morphology-based assignment to subspecies does not match assignment to *Pacificastacus* clusters as defined by COI. The three subspecies (indicated in green) were originally described as distinct species, yet they all end up in one cluster in our tree.



COI NJ-tree (sub-selection of sequences)

(Cat. 1, 6) Taxonomic issue and potential misidentification of vouchers:



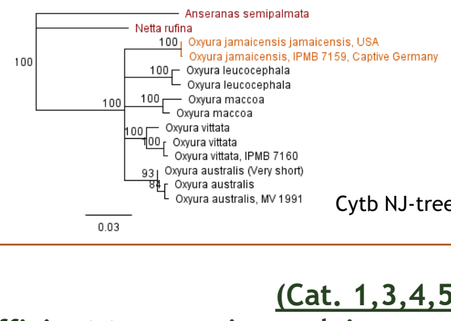
Cytb NJ-tree

Herpestes javanicus

H. auropunctatus and *H. javanicus* were considered as one species previously but are now split into two distinct species (5% genetic divergence- Veron *et al*. 2007). Moreover, the many introduced populations of *H. javanicus sensu lato* around the world are all believed to be *H. auropunctatus* (Veron *et al*. 2007).

(Cat. 2,3,8) Limited availability of DNA sequences and geographical coverage of both target and congeners:

Oxyura jamaicensis



Cytb NJ-tree

(Cat. 1,3,4,5,7) Insufficient taxonomic resolving power:

Pueraria montana* var. *lobata

trnE-trnT intergenic spacer NJ-tree

Lack of resolving power to distinguish closely related taxa that may hybridize.

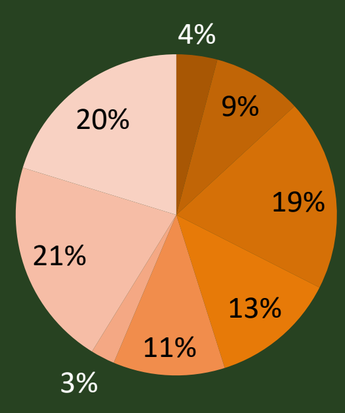


Fig. 1: Average proportion of the occurrence of each category (all DNA sequences, all species included)

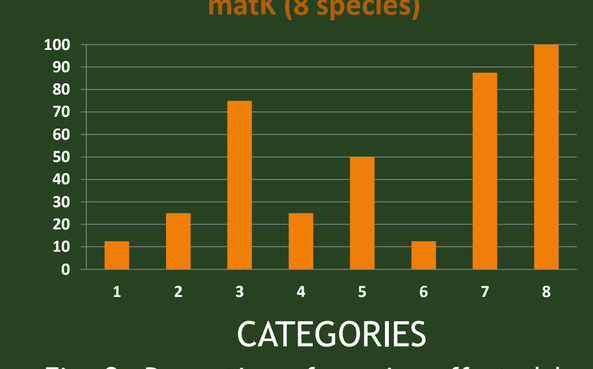


Fig. 2: Proportion of species affected by each potential issue, displayed for the four most represented markers

Outcomes

Overall, a total of 35 different markers were evaluated, including COI, Cytb, D-Loop, 12S, 16S, *rbcL*, ITS, *matK* and *psbA-trnH*. Per species, 1 to 10 different markers (mean of 4) were investigated. On average, the most common issues identified were (3), (7) & (8) (Fig. 1). The pattern recorded for the four most common markers are displayed in Fig. 2.

DNA-based identifications:

Despite the fact that the ideal situation was never met, we presently consider that 15 of the 37 IAS are reliably identifiable using DNA sequences. For 12 other species, DNA markers with a high potential were identified, yet they are discarded for the time being due to a lack of available material from the target and its congeners. The 10 last IAS are presently not considered as identifiable using DNA sequences.

Take-home message:

Although DNA barcoding is an identification method offering several advantages, the reliability of a species identification entirely depends on a comprehensive investigation of the available reference sequences of the target and its congeners. Shortcomings for more than half of the 37 IAS were highlighted in the present work.

DISCUSSION

REFERENCES

- Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39:783-791.
- Larson, E. R., Abbott, C. L., Usio, N., *et al*. (2012). The signal crayfish is not a single species: cryptic diversity and invasions in the Pacific Northwest range of *Pacifastacus leniusculus*. *Freshwater Biology*, *57*(9):1823-1838.
- Larson, E. R., Castelin, M., Williams, B. W., *et al*. (2016). Phylogenetic species delimitation for crayfishes of the genus *Pacifastacus*. *PeerJ*, *4*:e1915.
- Veron, G., Patou, M.-L., Pothet, G., *et al*. (2007). Systematic status and biogeography of the Javan and small Indian mongooses (Herpestidae, Carnivora). *Zoologica Scripta*, 36:1-10.

museum Africa belspo LifeWatch