DNA barcoding and forest biosecurity

L.M. Humble, J.R. deWaard, R. Hanner and P.D.N. Hebert

Introduction

The ability to distinguish non-indigenous species (NIS) from native species is critical to the success of any surveillance program. Unfortunately there are numerous problems inherent with detection of NIS including:

- large samples needed to detect NIS when they are present at low levels
- immature life stages often intercepted but usually cannot be fully identified
- often inadequate knowledge of native fauna

We provide examples from forest bio-surveillance in Canada showing how these problems can be circumvented by the application of DNA barcoding (DNAB) for species identification.

What is DNA barcoding?

- DNAB uses the sequence variation in a short, standardized DNA fragment to identify organisms
- DNA fragment used for animals is a 658 base pair segment of the mitochondrial gene cytochrome *c* oxidase subunit 1 (COI) (Hebert *et al.* 2003)
- It compares unknown sequences against a reference library of DNA sequences
- It meets or exceeds minimum standards required for diagnostic protocols under ISPM No. 27 (Floyd *et al.* 2010)

Figure 1. Light trapping for nocturnal moths.

Biosurveillance for NIS

Methods

- moths sampled with UV lights (Fig. 1)
- a single leg was removed and barcoded using standardized procedures (Fig. 2) **(**deWaard *et al*. 2009**)**

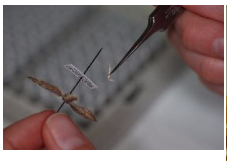

Figure 2. Sampling tissue from moths for barcoding.

Biosurveillance (cont'd)

Results

The identification engine in the Barcode of Life database (BOLD-ID) was used to obtain initial identifications for the 925 specimens (Fig. 3).

- ~190 species clusters with a 3% sequence divergence cut-off (Hebert *et al.* 2003)
- 124 clusters assigned to species, 61 to genus using BOLD-ID (all species assignments were also confirmed morphologically)
- 66 remaining clusters were identified morphologically (the only step in the identification process that required taxonomic specialists)

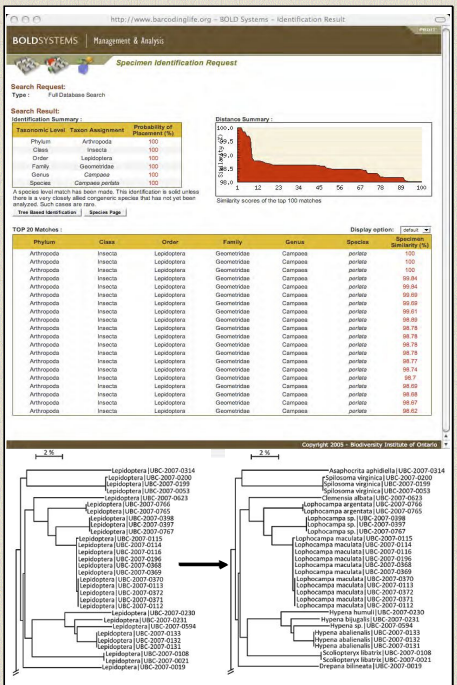

Figure 3. Species identification report and neighbour-joining trees from BOLD-ID.

- 31 species and 16% of all moths captured were NIS.
- two NIS, *Argyresthia pruniella* and *Dichelonia histrionana* (Fig. 4), were new introduction records for North America
- two NIS, *Paraswammerdamia lutarea* and *Prays fraxinella* recorded for the first time from western Canada

DNA barcoding provides an efficient and rapid means of assessing large samples. It enhances both species recognition and the detection of new NIS by:

- minimizing valuable specialist time;
- detecting species at low density

Figure 4. *Argyresthia pruniella* (left) and *Dichelonia histrionana* (right)

Identification of immature life stages

Eggs, larvae or pupae are the most frequently intercepted life stages of many quarantine organisms. They usually cannot be fully identified to species'.

- DNA barcodes are invariant during a species development; any life stage can be identified from its DNA sequence

European poplar shoot borer, *Gypsonoma aceriana*, was first reported from North America in 2001

- barcoding of adults in museum collections confirmed its long-term presence in western Canada (Fig. 5)
- barcoding of larvae from delimitation surveys used to define its range in British Columbia (Humble *et al.* 2009)

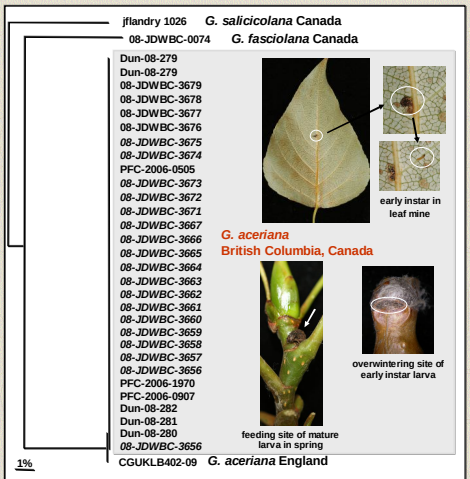

Figure 5. Neighbour-joining tree of COI sequences and geographic origin of samples for *Gypsonoma* species. Larval samples are denoted with italics.

Building DNA reference libraries

DNA barcoding identifies unknown species by comparing their COI sequences to reference sequences derived from reliably identified species sampled from museum collections. Development of the sequence libraries is done in collaboration with taxonomic specialists. Two examples of reference library development follow.

DNA barcoding and forest biosecurity (cont'd)

Building DNA reference libraries

1. *Lymantria* tussock moths

Lymantria includes serious quarantine & forest pests [e.g. gypsy moth (*L. dispar*), pink gypsy moth (*L. mathura*) and nun moth (*L. monacha*)]

- Species are often transported globally as dormant egg masses on vessels and cargo
- DNA reference library constructed for the identification of 36 *Lymantria* spp. (deWaard *et al.* 2010a)

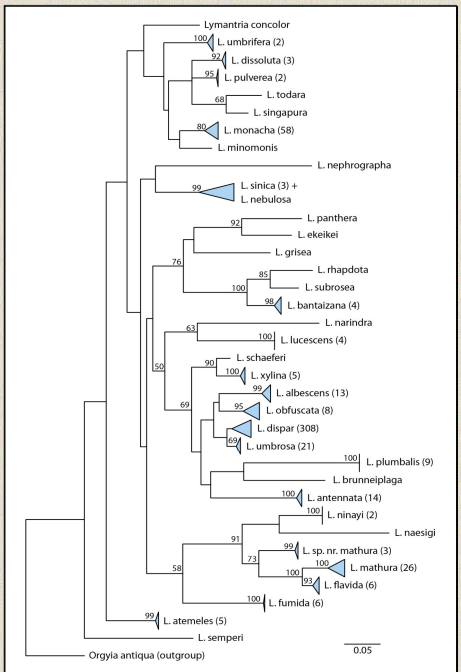

Figure 6. Maximum likelihood tree for 36 species of *Lymantria* constructed with the barcode region of the COI gene. The number of specimens sampled per species is noted in parentheses (after deWaard 2010).

Results

- 518 adult *Lymantria* from 35 countries barcoded (Fig. 6)
- barcode data led to morphological and taxonomic re-evaluation of specimens in two clusters (deWaard *et al.* 2010a)
  - *L.* sp. nr. *mathura* (Fig. 6) now considered to be *L. subpallida*
  - *L. nebulosa* is a valid species distinct from *L. sinica*

Lymantria spp. results (cont'd)

- 142 COI haplotypes identified across all 36 species
- no haplotypes shared between species
- 91 COI haplotypes within *L. dispar*
- haplotypes of Asian subspecies *L. dispar asiatica* and *L. dispar japonica* cluster separately from the European and North American subspecies *L. dispar dispar*
  - allows rapid identification of "Asian gypsy moth" recovered from monitoring programs

2. Geometridae

DNA reference library was developed for the 349 spp. of Geometridae in British Columbia (deWaard 2010).

- specimens from 8 museum collections sampled, databased and imaged; DNA extracted and COI sequenced
- all data is publicly available from the Barcode of Life Data Systems (BOLD) (Ratnasingham and Hebert 2007) in project 'GOBCL – Geometridae of BC Library'

Results

- 2392 COI sequences generated from 400 species in 125 genera
- 374 (93.2%) of the species could reliably be distinguished with barcodes
- only 27 species (6.8%) had undifferentiated or overlapping barcodes
- both a new NIS (Fig. 7) and a new native species for Canada were detected by barcoding museum collections (deWaard *et al.* 2008, 2010b)

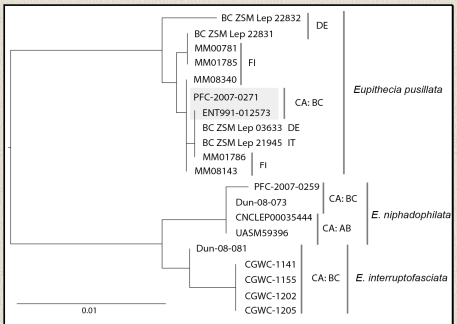

Figure 7. Neighbour-joining tree of the European species, *Eupithecia pusillata,* and two closely related native spp., *Eupithecia niphadophilata* and *E. interruptofasciata.* All feed on juniper (after deWaard *et al.* 2010b)

Abbreviations: DE – Germany, FI – Finland, IT – Italy, CA – Canada, BC – British Columbia, AB – Alberta

Literature cited

deWaard, J.R., Schmidt, B.C., Anweiler, G.G. and Humble, L.M. 2008. First Canadian records of *Lampropteryx suffumata* ([D. & Schiff.], 1775) (Geometridae: Larentiinae). J. Ent. Soc. Brit. Columbia 105: 19–25.

deWaard, J.R., Landry, J.-F., Schmidt, B.C., Derhousoff, J., McLean, J.A. and Humble, L.M. 2009. In the dark in a large urban park: DNA barcodes illuminate cryptic and introduced moth species. Biodiversity and Conservation 18: 3825–3839

deWaard, J.R. 2010. Forest biomonitoring, biosecurity and DNA barcoding. Ph.D.Dissertation. University of British Columbia, Forest Sciences Department. https://circle.ubc.ca/handle/2429/30496

deWaard, J.R., Mitchell, A., Keena, M.A., Gopurenko, D., Boykin, L.M., Armstrong, K.F., Pogue, M.G., Lima, J., Floyd, R., Hanner, R.H. and Humble, L.M. 2010a. Towards a global barcode library for *Lymantria* (Lepidoptera: Lymantriinae) tussock moths of biosecurity concern. PLoS ONE 5: e14280.

deWaard, J.R., Schmidt, B.C.S., and Humble, L.M. 2010b. DNA barcoding flags the first North American records of a Eurasian moth, *Eupithecia pusillata* (D. & Schiff., 1775) (Lepidoptera: Geometridae). J. Entomol. Soc. Brit. Columbia 107:1–7.

Floyd, R., Lima, J., deWaard, J., Humble, L. and Hanner, R. 2010. Common goals: policy implications of DNA barcoding as a protocol for identification of arthropod pests. Biological Invasions 12 (9): 2947-2954

Hebert, P.D.N., Cywinska, A., Ball, S.L. and deWaard, J.R. 2003. Biological identifications through DNA barcodes. Proc. of the Royal Soc. of London B 270: 313–321

Humble, L.M., deWaard, J.R. and Quinn, M. 2009. Delayed recognition of the European poplar shoot borer, *Gypsonoma aceriana* (Duponchel) (Lepidoptera: Tortricidae) in Canada. J. Ent. Soc. Brit. Columbia 106: 61–70.

Ratnasingham, S. and Hebert, P.D.N. 2007. BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7: 355–364.

Acknowledgments

The contributions of all of the co-authors of the publications listed above are gratefully acknowledged. We also acknowledge funding support from the Ontario Centres of Excellence research program, the Ontario Ministry of Agriculture, Food and Rural Affairs, the Canadian Food Inspection Agency and the Canadian Forest Service. Additional support was provided by a Forest Investment Account - Forest Science Program Student Grant, Canadian Forest Service - Pacific Graduate Student Supplements and an NSERC Graduate Scholarship (to JRdW). This research was also supported through funding to the Canadian Barcode of Life Network from Genome Canada (through the Ontario Genomics Institute), NSERC and other sponsors listed at http://www.BOLNET.ca.

For further information

Please contact the authors (see below) for publications. More information on these and related projects can be obtained from the Canadian Centre for DNA Barcoding - *http://www.ccdb.ca/* or the Barcode of Life Data Systems - *http://www.boldsystems.org/views/login.php*.

Leland Humble – Natural Resources Canada, Canadian Forest Service, Victoria, BC V8Z 1M5 & Forest Sciences Department, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (leland.humble@nrcan-rncan.gc.ca); **Jeremy deWaard –** Forest Sciences Department, University of British Columbia, Vancouver, BC V6T 1Z4, Canada & Entomology, Royal British Columbia Museum, Victoria, BC V8W 9W2, Canada (jeremy.dewaard@gmail.com); and **Robert Hanner** & **Paul Hebert –** Biodiversity Institute of Ontario & Department of Integrative Biology, University of Guelph, Guelph, ON N1G 2W1, Canada (rhanner@uoguelph.ca & phebert@uoguelph.ca).