Submitted articles:

Mass Amputation
at the Royal BC Museum:
Sampling B.C.'s Geometrid Moths for DNA Barcoding

Written by Jeremy R. deWaard,
University of British Columbia, Department of Forest Sciences, Vancouver, BC V6T 1Z4
Royal British Columbia Museum, Entomology, 675 Belleville Street, Victoria, BC V8W 9W2

DNA barcoding has emerged in recent years, promising rapid, objective, and accurate species-level identifications using a standardized DNA fragment (Hebert *et al.* 2003a). Proof of principle studies since its inception have demonstrated the efficacy of sequence variation in a 648 base pair segment of the cytochrome oxidase 1 (COI) gene to differentiate species in a wide range of animals (e.g. Hebert *et al.* 2003b) and in distant sectors of life (e.g. Saunders 2005, Seifert *et al.* 2007). While its effectiveness in Lepidoptera has been explored previously (Hebert *et al.* 2003a, Hebert et al., *in prep*), the entire fauna of a region, particularly in a hyper-diverse family, has not been investigated. The Geometridae of BC provide an intriguing test group to this end, with approximately 345 species in BC alone. Moreover, it presents a group that would benefit greatly from a species-identification tool, due to its taxonomically challenging species groups, many of which can only be disentangled by labour-intensive genitalic dissections. In addition, a large proportion of the species are important defoliators, including several native pests such as the fall cankerworm (*Alsophila pometaria*) and invasive pests such as the winter moth (*Operophtera brumata*), making their reliable diagnosis imperative.

To establish and evaluate a DNA barcode library for BC's geometrid moths, I utilized the extensive geometrid collection of the Royal British Columbia Museum (RBCM). In several visits to the collection, 753 specimens of geometrid moths were selected and analyzed. Between 1 and 8 specimens were sampled for each species, with an effort made to cover the geographic range of each species. All specimens selected were databased, digitally photographed, and sampled for DNA analysis by removing a single leg. The tissue samples were analyzed at the Biodiversity Institute of Ontario in Guelph, ON following standard protocols (deWaard et al., in press). The data, images and resultant sequences were uploaded to the Barcode of Life Database (BOLD; www.barcodinglife.org). By comparing the new barcode sequences with the BOLD reference database, the identifications of the RBCM specimens could be confirmed, determined, or flagged as potentially incorrect. The latter two cases were checked using the dichotomous keys of McGuffin (e.g. 1987) and Bolte (1990) and corrected as necessary. Species groups that require genitalic

dissections for species delimitation have been left as originally identified; dissections are underway but were not completed for this report.

Of the 753 specimens analyzed, a DNA barcode was successfully generated from 707 (93.9%). This represents 250 species, providing a mean sampling intensity of 2.8 sequences per species. Almost all specimens analyzed were over 10 years old, and barcode sequences were recovered from specimens up to 29 years old. The 46 specimens that were not successfully barcoded contained only four species not represented by another sequence (*Eumacaria latiferrugata, Eupithecia behrensata*, *Eupithecia olivacea,* and *Leucobrephos brephoides,*). In general, specimens within a species form monophyletic groups that are divergent from other species clusters, as for instance, in the subfamily Geometrinae (Figure 1A). The COI divergence is typically low (mean = 0.94, range = 0-14.94, SE = 0.08) while the divergence between species of the same genus is relatively high (mean = 8.07, range = 0-15.62, SE = 0.069). However, this is not always the case – 36 of the 250 species could not be discriminated by DNA barcodes i.e. they either shared identical barcodes or had inter-mixed sequences with closely related species. An example of this is in the genus *Caripeta. C. divisata* and *angustiorata* can be distinguished from *aequaliaria*, but the former two can not be differentiated from one another (Figure 1B). The morphological identifications of these 2 species are unequivocal; given that the two share similar ranges and the adults fly at the same time (McGuffin 1987), this may suggest that they hybridize or continue to interbreed to some degree. Due to these 36 'misbehaving' species, the overall barcode success rate of 85.6% is nearly ten percent lower than previous estimates of the efficacy of barcodes in Lepidoptera (Hebert et al. 2003a, Hebert et al., *in prep*). This is may be due to several reasons, including some biological (e.g. increased taxonomic sampling should lead to a higher proportion of closely allied taxa, where a higher incidence of incomplete lineage sorting and introgression is expected), but it is probable that the success rate will increase when the pending genitalic dissections and final confirmations/corrections are complete.

By comparing the new barcode sequences with the BOLD reference database, the identification of specimens were confirmed (653), corrected (41) or determined (13). While the validity in extrapolating this to the entire RBCM Geometridae collection is debatable (since only specimens <30 years old were analyzed), this might suggest that roughly 6% of the specimens are misidentified. Again, this number is likely to change when all the necessary dissections are complete, allowing further confirmation and correction to the assigned identifications. Of the determinations provided by barcoding, 2 resulted in new species for the RBCM collection (*Eupithecia lachrymosa* and *E. intricata*). Similarly, one of the corrections provided a new species for the collection—and for Canada (*Lampropteryx suffumata*; deWaard et al., *in prep*).

These examples nicely highlight the utility of DNA barcoding for assisting faunal inventories, as well as its potential for invasive alien species detection.

In summary, this small study on the RBCM Geometridae collection demonstrates the reciprocal contributions that natural history collections and the DNA barcoding programme can provide one another. Collections receive 'value- added' to each specimen in the form of online images, DNA sequence data, archived DNA extractions and the onfirmation/correction/determination of species identification, all with negligible damage to specimens. The barcode programme on the other hand, receives access to endless authoritatively identified specimens for generating the barcode reference libraries, as well as access to the invaluable expertise and resources of the 'collectionsphere'. In tandem, the two institutions are certain to make significant contributions, both in taxonomic discoveries and with the countless applications of the DNA barcoding tool.

Literature Cited

- Bolte, K.B. (1990) Guide to the Geometridae of Canada (Lepidoptera). VI, subfamily Larentiinae, 1. Revision of the genus *Eupithecia*. No. 150.
- deWaard, J.R., Ivanova, N.V., Hajibabaei, M. and Hebert, P.D.N. Assembling DNA Barcodes: Analytical Protocols. In Methods in Molecular Biology: Environmental Genetics (Ed: Cristofre Martin; Humana Press Inc., Totowa, USA). In press.
- deWaard, J.R., Schmidt, B.C., Anweiler, G.G., and Humble, L.M. First Canadian records of *Lampropteryx suffumat*a (Denis & Schiffermüller, 1775) (Geometridae: Larentiinae). In preparation.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., and deWaard, J.R. (2003a) Biological identifications through DNA barcodes. Proceedings of the Royal Society of London B 270:313-321.
- Hebert, P.D.N., Ratnasingham, S., and deWaard, J.R. (2003b) Barcoding animal life: COI divergences among closely related species. Proceedings of the Royal Society of London B 270:S96-S99.
- Hebert, P.D.N., deWaard, J.R., and Landry, J-F. DNA barcodes deliver: Species identifications and revelations for 1/1000 of the Animal Kingdom. In preparation.
- McGuffin, W.C. (1987) Guide to the Geometridae of Canada (Lepidoptera). II, subfamily Ennominae 4. No. 138.
- Saunders, G.W. (2005) Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. Philosophical Transactions of the Royal Society of London, B, *Biological Sciences* 360: 1879-1888.
- Seifert, K.A., Sampson, R.A., deWaard, J.R., Houbracken, J.A., et al. (2007) Prospects for fungus identification using CO1 DNA barcodes, with *Penicillium* as a test case. *Proceedings of the National Academy of Sciences* 104:3901-3906.
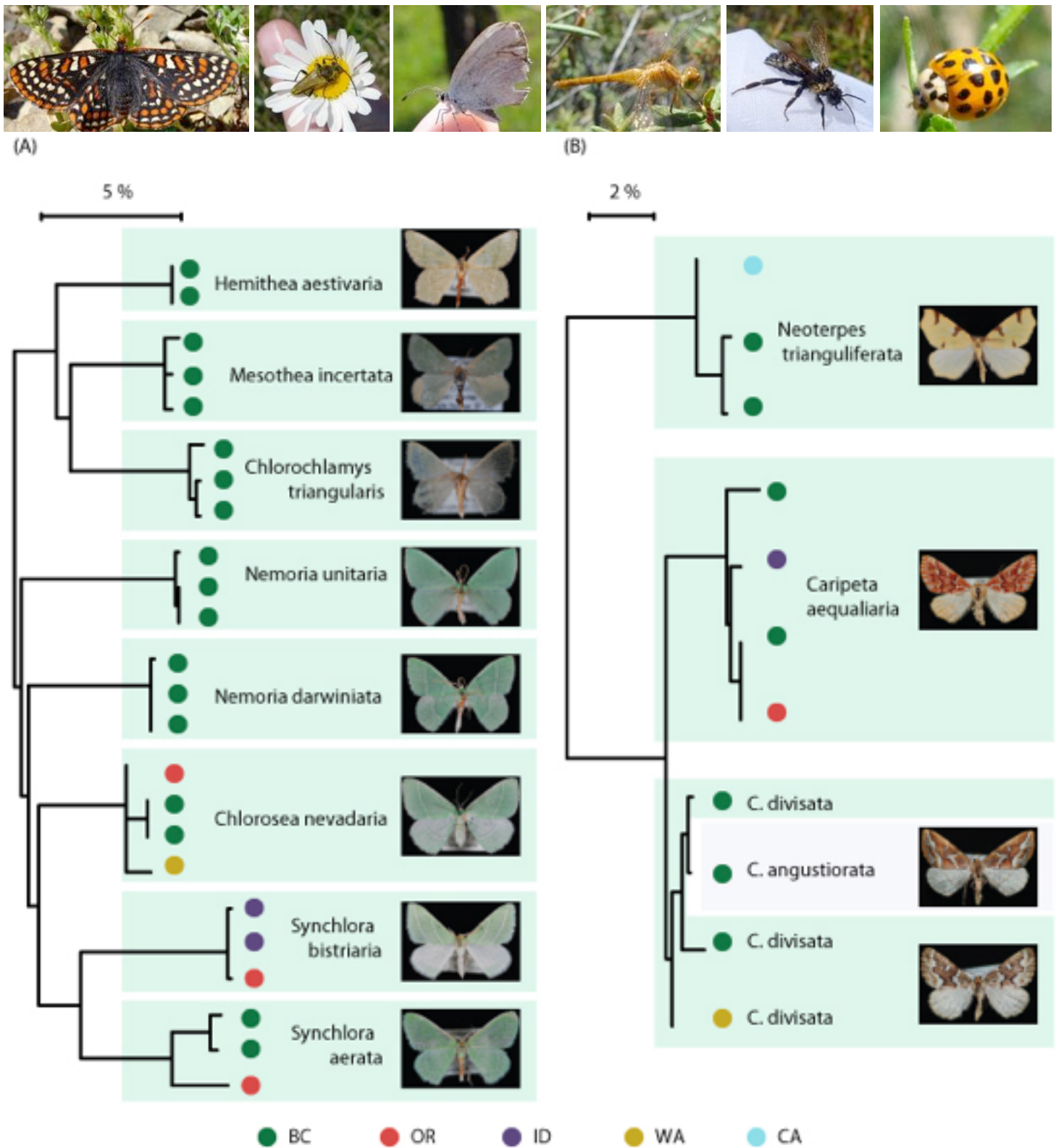
Figure 1. Neighbour-joining trees constructed with COI barcode sequence data for (A) the subfamily Geometrinae and (B) the genus *Caripeta* (with *Neoterpes* as outgroup).