

# Unmasking hidden butterfly diversity with DNA barcodes: Fifteen putative species of Nymphalidae

John J. Wilson<sup>1</sup>, Mehrdad Hajibabaei<sup>1</sup>, Daniel H. Janzen<sup>2</sup>, Winnie Hallwachs<sup>2</sup> & Paul D. N. Hebert<sup>1</sup>

<sup>1</sup> Canadian Centre for DNA Barcoding, Biodiversity Institute of Ontario and Department of Integrative Biology, University of Guelph  
<sup>2</sup> Department of Biology, University of Pennsylvania



## INTRODUCTION

Butterflies are among the most conspicuous invertebrates in terrestrial ecosystems and represent a benchmark group for the assembly of an 'encyclopedia of life'. The Nymphalidae make an abundant and attractive component of this diversity, comprising the majority of the Costa Rican butterfly fauna (figure 1).

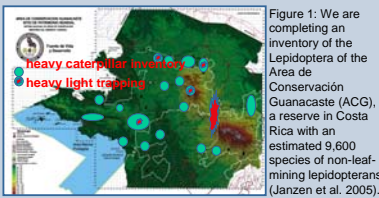


Figure 1: We are completing an inventory of the Lepidoptera of the Area de Conservación Guanacaste (ACG), a reserve in Costa Rica with an estimated 9,600 species of non-leaf-mining lepidopterans (Janzen et al. 2005).

DNA barcode sequence (*COI* mtDNA) patterns have flagged numerous records of potential species unrecognized by science (figure 2). Distinctive barcode clusters, often supported by ecological or morphological characters, can be used as the basis of 'primary' species hypotheses (figure 5). We are currently testing fifteen of these hypotheses through genealogical concordance with novel molecular sequences (figure 6).

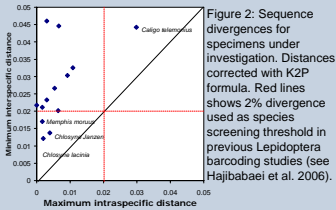


Figure 2: Sequence divergences for specimens under investigation. Distances corrected with K2P formula. Red lines shows 2% divergence used as species screening threshold in previous Lepidoptera barcoding studies (see Hajibabaei et al. 2006).

If these clusters represent phylogenetic species and not karyotypic structure, then this should be corroborated by diagnostic characters in sequences from the nuclear genome.

## DATA COLLECTION

Specimens reared from wild-collected caterpillars by DHJ, WH and a parataxonomist team during 27 years of biodiversity inventory in the ACG are sent to CCDB for molecular analysis (figure 3).



Figure 3: Production of a DNA barcode following Hajibabaei et al. (2006). We intend to recover nuclear sequences from the same extracts.

In initial testing in Saturniidae the *wingless* gene showed very similar patterns of divergence as *COI* barcodes (figure 4).

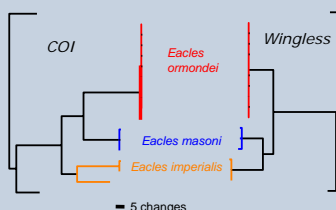


Figure 4: Mitochondrial and nuclear concordance in *Eacles*.

## DATA ANALYSIS

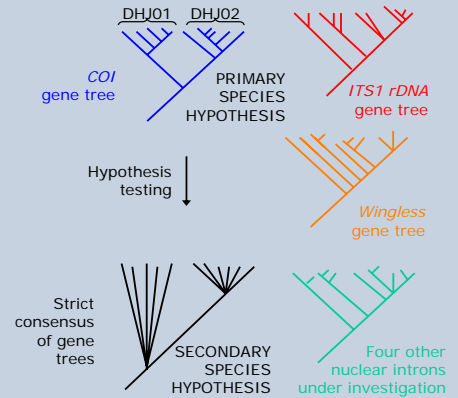


Figure 6: Testing species hypotheses using a genealogical concordance approach. Monophyletic clades in the consensus tree represent corroborated species. Adapted from Sites & Marshall (2004).

## DISCUSSION

This study has important implications:

- 1) It could reveal hidden diversity within a relatively well-known group. This has large implications for our estimates of species diversity within the Lepidoptera.
- 2) Alternatively this study could demonstrate cases where the barcode sequences do not conform to the pattern we normally see, raising new questions about *COI* and mtDNA evolution.
- 3) It provides empirical measures of the correlation between the barcode sequence and various nuclear gene regions, providing insights into how species limits are captured by a range of different markers.
- 4) It aims to provide an explicit and logical scientific framework for testing species hypotheses generated through DNA barcoding. This could subsequently be applied to other groups within the Lepidoptera and extended to other insects, the most diverse animal group on earth.

## REFERENCES

Hajibabaei M et al. (2006) DNA barcodes distinguish species of tropical Lepidoptera. *PNAS* 103: 968-971.  
 Janzen DH et al. (2005) Wedding a tropical inventory with DNA barcoding. *Phil. Trans. R. Soc. B* 360: 1835-1845.  
 Sites JW & Marshall JC (2004) Operational criteria for delimiting species. *Annu. Rev. Ecol. Syst.* 35: 199-227.

## ACKNOWLEDGEMENTS

Thanks to the many collaborators involved in the ACG 'BioLeP' project both in Costa Rica and the Canadian Centre for DNA Barcoding. Research is supported by grants to DHJ and WH from the Wege Foundation, NSF, Guanacaste Dry Forest Conservation Fund and ACG and by grants to PDNH from the Gordon and Betty Moore Foundation, NSERC, and Genome Canada through the Ontario Genomics Institute.

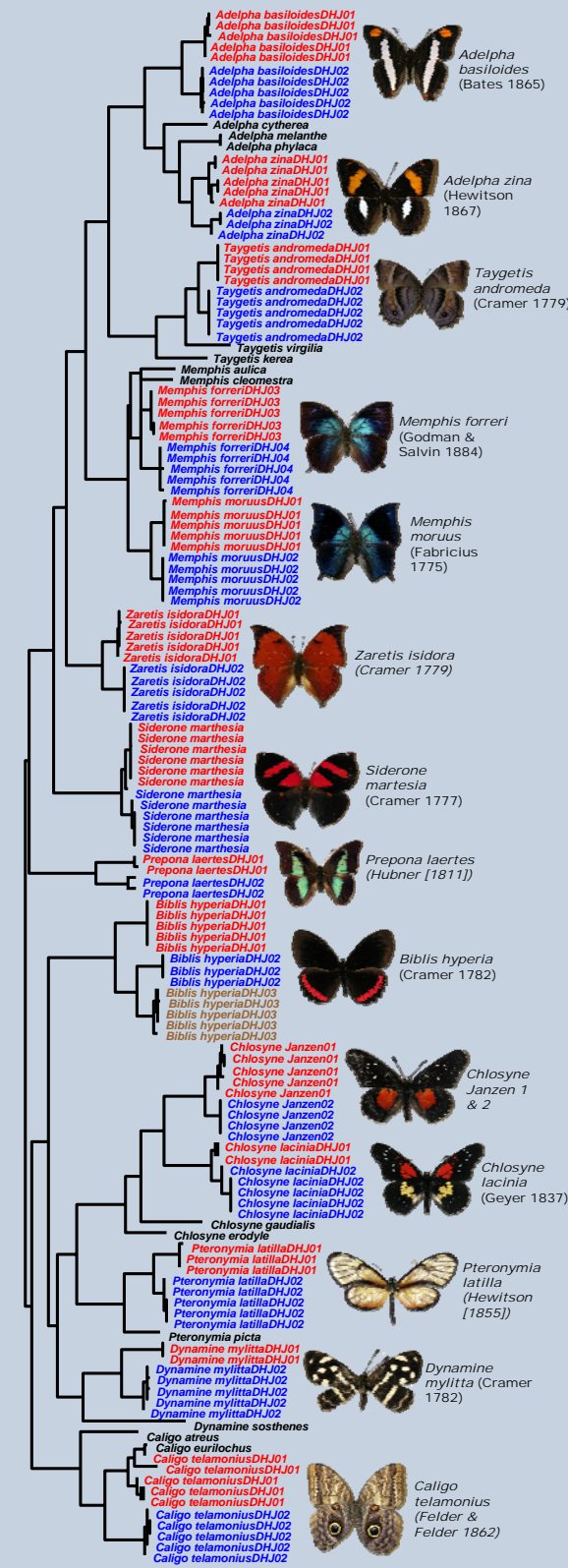


Figure 5: One of 200 most parsimonious cladograms inferred from barcode sequences showing 15 potential new Nymphalidae species. Converted to a phylogram to visualize number of nucleotide changes supporting the branches.

For more information please email: [jwilso04@uoguelph.ca](mailto:jwilso04@uoguelph.ca)