DNA Barcodes for 1/1000 of the Animal Kingdom

Paul D.N. Hebert [1*]

Jeremy R. deWaard [2,3]

Jean-François Landry [4]

1 Biodiversity Institute of Ontario, University of Guelph, Guelph, ON N1G 2W1, Canada

2 Department of Forestry Science, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

3 Entomology, Royal British Columbia Museum, Victoria, BC V8W 9W2, Canada

4 Research Centre, Agriculture and Agri-Food Canada, Ottawa, ON K1A 0C6, Canada

* Corresponding author: phebert@uoguelph.ca

Running Title: DNA Barcodes for Lepidoptera

Summary: This study reports DNA barcodes for more than 1300 Lepidoptera species from the eastern half of North America, establishing that 99.5% of these species possess diagnostic barcode sequences. Intraspecific divergences averaged just 0.43% among this assemblage, but most values were lower. The mean was elevated by deep barcode divergences (>2%) in nearly 5% of the species, often involving the sympatric occurrence of two barcode clusters. A few of these cases have been analyzed in detail, revealing species overlooked by the current taxonomic system. This study also provided a large-scale test of the extent of regional divergence in barcode sequences, indicating that geographic differentiation is small, even when comparisons involve populations as much as 2800 km apart. The present results affirm that a highly effective system for the identification of Lepidoptera can be built with few records per species because of the limited intra-specific variation. As most terrestrial and marine taxa possess patterns of population structure similar to those in Lepidoptera, an effective DNA-based identification system can be developed with modest effort.

Keywords: DNA barcoding, cytochrome oxidase 1, species identification, cryptic species, Lepidoptera