

DNA metabarcoding of insects and allies: an evaluation of primers and pipelines

G.-J. Brandon-Mong^{1,2}, H.-M. Gan^{3,4}, K.-W. Sing^{1,2},
 P.-S. Lee^{1,2}, P.-E. Lim⁵ and J.-J. Wilson^{1,2*}

¹Museum of Zoology, Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia: ²Ecology and Biodiversity Program, Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia: ³School of Science, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500 Petaling Jaya, Selangor, Malaysia: ⁴Monash University Malaysia Genomics Facility, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway 47500 Petaling Jaya, Selangor, Malaysia: ⁵Institute of Ocean and Earth Sciences (IOES), University of Malaya, 50603 Kuala Lumpur, Malaysia

Abstract

Metabarcoding, the coupling of DNA-based species identification and high-throughput sequencing, offers enormous promise for arthropod biodiversity studies but factors such as cost, speed and ease-of-use of bioinformatic pipelines, crucial for making the leapt from demonstration studies to a real-world application, have not yet been adequately addressed. Here, four published and one newly designed primer sets were tested across a diverse set of 80 arthropod species, representing 11 orders, to establish optimal protocols for Illumina-based metabarcoding of tropical Malaise trap samples. Two primer sets which showed the highest amplification success with individual specimen polymerase chain reaction (PCR, 98%) were used for bulk PCR and Illumina MiSeq sequencing. The sequencing outputs were subjected to both manual and simple metagenomics quality control and filtering pipelines. We obtained acceptable detection rates after bulk PCR and high-throughput sequencing (80–90% of input species) but analyses were complicated by putative heteroplasmic sequences and contamination. The manual pipeline produced similar or better outputs to the simple metagenomics pipeline (1.4 compared with 0.5 expected:unexpected Operational Taxonomic Units). Our study suggests that metabarcoding is slowly becoming as cheap, fast and easy as conventional DNA barcoding, and that Malaise trap metabarcoding may soon fulfill its potential, providing a thermometer for biodiversity.

Keywords: Arthropoda, biodiversity, COI, high-throughput sequencing, Illumina MiSeq, Malaise trap

(Accepted 19 July 2015; First published online 7 September 2015)

Introduction

Much of our knowledge of biodiversity patterns and changes comes from the data based on mammals, birds and

vascular plants (e.g., Gillison *et al.*, 2013). Yet these taxa represent only a fraction of biodiversity; the major component of terrestrial biodiversity comprises insects (Mora *et al.*, 2011). A recent meta-analysis of biodiversity studies revealed the dearth of information about most of the world's tropical biota (Gillison *et al.*, 2013), highlighting the fact that in order to decipher biodiversity patterns and change the major component can no longer be ignored. The absence of data on insects in biodiversity surveys, with the exception of small

*Author for correspondence
 Fax: +603-7967-4178
 E-mail: johnwilson@um.edu.my