DNA barcoding as identification tool in fruit fly interception and surveying activities

Floris Breman[1,2], Kurt Jordaens[1,2], Marc De Meyer[1,2] & Massimiliano Virgilio[2,3]

[1] Joint Experimental Molecular Unit, Brussels & Tervuren, Belgium
[2] Royal Museum for Central Africa, Entomology Section, Tervuren, Belgium
[3] Royal Belgian Institute of Natural Sciences, Invertebrates Section, Brussels, Belgium

Introduction

DNA barcoding uses the sequences of a particular part of the mitochondrial COI gene as a unique identifier for all zoological species. DNA barcodes can provide a valuable alternative to morphological identification as far as libraries of DNA barcodes are representative of the taxonomic diversity of particular groups. Complete taxon coverage in tephritids is currently not feasible because of the large numbers of tephritid species and the unavailability of specimens for the rarer species. However, for quarantine interception and monitoring activities with para-pheromone traps, identification of fruit flies through DNA barcodes could rely on regional libraries that contain the main economically important (EI) species.

female of *Bactrocera cucurbitae* (©R.S. Copeland, 1995)

Methods

This study focuses on three fruit fly genera of economic importance (EI): *Ceratitis*, *Dacus* and *Bactrocera*. A DNA barcode library, including 589 sequences from 128 indigenous African and alien invasive tephritid taxa, has been established for those taxa which are regularly encountered in para-pheromone trapping activities in different parts of the African continent. It includes all major EI species found on the continent. The reliability of this library for tephritid identification was tested on a set of 83 tephritid specimens intercepted in European NPPO's or collected during recent monitoring surveys in Africa. Specimens originated from 9 different countries (Burkina Faso, Cameroon, Egypt, Ivory Coast, Kenya, Mali, Mozambique, South Africa, Togo). We quantified the proportion of correctly identified queries using three distance-based criteria (Best Match: BM; Best Close Match: BCM, Best Match with a distance threshold of 3%= 3%T) and a tree-based identification method (tree-identification: NJT) as defined in Virgilio et al. (2010: BMC Bioinformatics 11:206).

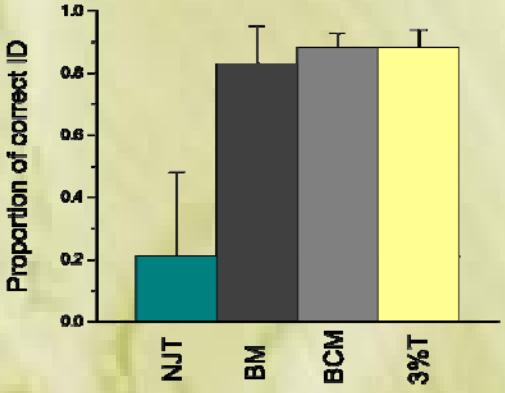

Figure 1: Proportion of correctly identified queries through Neighbor-Joining Tree (NJT), Best Match (BM), Best Close Match (BCM) and 3% threshold (3%T).

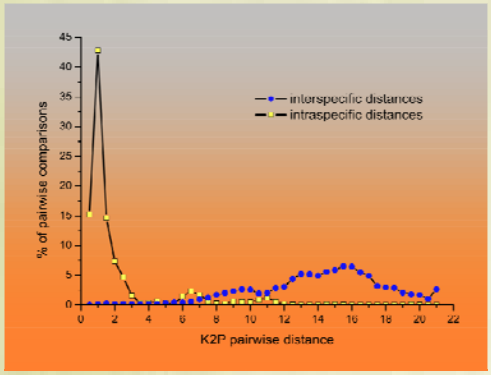

Figure 2: Distribution of interspecific (blue circles) and intraspecific (yellow squares) pairwise Kimura-2-Parameter distances.

Results and discussion

The percentage of correctly identified queries (± standard deviation) over the three genera, *Ceratitis*, *Dacus* and *Bactrocera* were remarkably low for the tree-based ID (NJT=19±24%) but considerably increased when distance-based criteria were used (BM=84±13%; BCM=89±7%; 3%T=89±8%) (Figure 1). Three main sources of error were identified:

1) lack of reference sequences in the library used
2) inadequate resolution of the molecular marker (particularly for species complexes)
3) human errors in sequencing and morphological identification of samples

As already observed in a number of insect orders (Virgilio et al. 2010), the distributions of intra- and interspecific-congeneric distances are largely overlapping (Figure 2). Hence, the threshold proposed by Hebert et al. (2004: PLoS Biology 2:e312) (i.e. the ratio mean congeneric interspecific divergence/mean intraspecific divergence ≥10) can not be used to separate the tephritid taxa considered in this study. Despite the lack of a clear barcoding gap, the distance-based criteria show a relatively high identification success, which could be improved when the regional library for African tephritids will be extended in order to include a larger number of reference taxa and when problematic cases of complexes are covered in an adjusted identification protocol.

Acknowledgements: We would like to thank Valérie Balmès (Laboratoire National de la Protection des Végétation, France), Sharon Reid (Food and Environment Research Agency, UK), Komina Amevoin (Université de Lomé, Togo) and Mervyn Mansell (USDA/APHIS Pretoria Branch, South Africa) for providing samples. The financial support of Belgian Science Policy (through JEMU and Action 1 of M. Virgilio) is acknowledged.