

# Identifying the specific status and geographical origin of European incursions of an invasive and cryptic pest, *Bactrocera dorsalis* (Diptera: Tephritidae).

Emeline CHARBONNEL<sup>1,2,3</sup>, David OUVARD<sup>3</sup>, Laure BENOIT<sup>1,2</sup>, Marie-Pierre CHAPUIS<sup>1,2</sup>

<sup>1</sup> CBGP, CIRAD, INRAE, Institut Agro, IRD, Univ Montpellier, Montpellier, France

<sup>2</sup> CIRAD, CBGP, F-34398 Montpellier, France

<sup>3</sup> ANSES, Laboratoire de la Santé des Végétaux, Unité entomologie et botanique, Montferrier-sur-Lez, France

Biological invasions are increasing exponentially and pose significant threat to biodiversity, ecosystems, and agriculture. Advances in genome sequencing create new opportunities to understand invasion routes, invasion success and colonization dynamics, which helps in detection and management of invasive species. The Oriental fruit fly *Bactrocera dorsalis*, native to Asia, has invaded the vast majority of Africa and the Indian Ocean islands, and is threatening the United States of America, Australia and the European Union, where it is classified as a priority quarantine pest. Gaps in the invasion history reconstruction and in the identification of population sources of incursions call for population genomics studies. However, their potential is limited by possible taxonomic confusions since *B. dorsalis* belongs to a complex in which species delimitation is unclear for a several taxa, hybridization was reported and attraction to methyl-eugenol, the male chemical lure used in trapping, is shared by several species of the complex.

In this context, we first proposed a protocol for ensuring the correct identification of *B. dorsalis* specimens, combining morphological expertise to both mt and nu available DNA barcodes (COI and EIF3L). This approach is especially required for specimens from Southeast Asian islands since half of them displayed characters ambiguous with *B. carambolae* and/or *B. occipitalis*. We also optimized a laboratory protocol of restriction-site associated DNA (RAD) sequencing accurate (i.e. genotyping error < 0,05%) and robust for low quality DNA samples (i.e. molecular weight of 5-10kb). We then applied these protocols to (i) 66 genomes belonging to 13 species of the *B. dorsalis* complex attracted to methyl-eugenol, including all species with morphological similarity and/or reported to hybridize with *B. dorsalis*, and (ii) 84 *B. dorsalis* genomes representative of the whole distribution range.

Using population genetics approaches, we (i) observed a clear genomic delineation supporting all species, without evidence for nuclear introgression, and (ii) resolved part of the demographic history in both native and invasion areas. In the native Asia, continental countries were differentiated from Malaysia, the southernmost continental point, with evidence of gene flow, and from Southeast islands, which were also differentiated from each other. Two distinct genetic groups in Africa, one from the East and South, the other from the West, with asymmetrical gene flow from the former to the latter, suggested two independent



invasion events, both from continental Asia. We identified potential sources of invaded islands: the Papua New Guinean sample grouped with Indonesian samples; Mayotte and Madagascan samples grouped together with Mozambique; La Réunion island and French Polynesian samples were related to continental Asia. From these genomics resources, we designed probes to capture 19,678 loci specific to *B. dorsalis* and genotyped 1,950 individuals belonging to 70 worldwide populations (in addition to the European incursion samples) in order to refine the invasion history and inform on the geographic origin of the European incursions.

