

**Linking genomes to Museum vouchers:  
an open-access database for African “true” fruit flies (Diptera, Tephritidae)**

Esselens L.<sup>1,2</sup>, Ody J.<sup>1</sup>, Sonet G.<sup>2</sup>, Deschepper P.<sup>1</sup>, Vanbergen S.<sup>1</sup>, Ferrari G.<sup>3</sup>, Hollingsworth P.<sup>3</sup>,  
De Meyer M.<sup>1</sup>, Vangestel C.<sup>2</sup>, Virgilio M.<sup>1</sup>

<sup>1</sup>Royal Museum for Central Africa (RMCA), Biology Department, Tervuren, Belgium. Email: [lore.esselens@africamuseum.be](mailto:lore.esselens@africamuseum.be), <sup>2</sup>Royal Belgian Institute of Natural Sciences (RBINS), Brussels, Belgium, <sup>3</sup>Royal Botanic Garden Edinburgh, UK

*Background:* Museum specimens are of particular importance for research on taxonomy, systematics and biogeography and provide tools to tackle a wide range of scientific questions in disciplines such as ecology, evolution and conservation. The rapid advances of -omic technologies have led to a substantial reduction in costs, so that the routine whole genome sequencing (WGS) of museum specimens represents a cutting-edge intervention to valorise the collections’ genetic resources. In this context, the Joint Experimental Molecular Unit (JEMU) of RMCA-RBINS is defining the best practices for the genotyping of suboptimal insects. This approach relies on the archiving of *genomic vouchers* as a complement to the archiving of morphological and digital museum vouchers.

*Methods:* The performances of (a) different DNA extraction procedures and kits and (b) different genomic library preparation approaches were quantitatively compared on approximately 300 tephritid vouchers of different ages sampled from the collections of RMCA. After DNA extraction and quality check, samples were subjected to WGS at 10x coverage on an Illumina NovaSeq platform (150 PE reads, 6Gb raw data output / sample). Relationships between quantity/quality of extracted insect DNA (as estimated by fluorometric methods and by levels of DNA fragmentation) and quantity/quality of WGS output (including total amount of filtered reads, proportion of high quality reads, etc.) were investigated using a correlative approach.

*Results and Conclusions:* Correlation was observed between sample age, DNA fragmentation and WGS performances. Accordingly, a cost- and time-effective approach to DNA extraction and genomic library preparation has been defined for insect vouchers collected up to 30 years ago. A decision map on wet- and dry-lab procedures for the WGS of museum insects is proposed to optimize results based on initial quality checks. These results show that *a priori* and large-scale genotyping of relatively fresh collection specimens represent a complementary and cost -effective approach to the targeted genotyping of degraded museum specimens. The WGS data and metadata obtained for more than 300 fruit fly vouchers have been archived on the RMCA servers. The setup of an open-access database for voucher archiving (including voucher metadata, genomic data, digital imaging and linked to the Global Genome Biodiversity Network) has been initiated and will be finalized by 2023 in the framework of the projects InsectMOoD and SYNTHESYS+ (respectively funded by the Belgian Science Policy and by the European Union).

**Keywords:** Museomics, genomics, genomic voucher, DNA archiving

**Session 3.** Genetics and Biotechnology