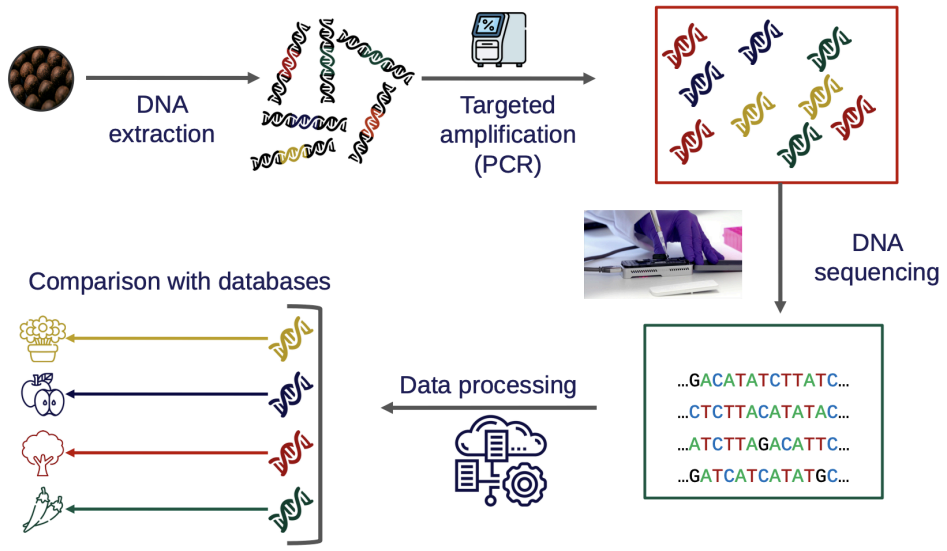
Clody Manon, Thébault Axelle and Wurch Inès

DNA Sequencing and species identification
by Lucas MARMIN

Nanopore sequencing, combined with DNA barcoding, offers a powerful method to identify species from complex biological samples. The technique relies on the amplification and sequencing of specific genes that must be sufficiently conserved to enable stable PCR amplification, but also sufficiently variable to distinguish different species.



DNA barcoding and sample preparation: To identify plant species, 4 universal genes are targeted for PCR amplification. To reduce analysis costs, the amplicons obtained are grouped together in a single sequencing run using a multiplexing process. This requires the addition of sample-specific “barcodes” to the PCR primers for identification purposes.

Nanopore sequencing: Nanopore technology uses semi-permeable membranes equipped with pores capable of recognizing adapters attached to DNA. As a DNA molecule passes through the pore, variations in ionic current enable the nucleotide sequence to be read. This process is particularly sensitive to DNA modifications.

Analysis and species identification: Raw sequencing data, generated as .pod5 files, are converted into usable files such as BAM or FASTQ via a “basecalling” process. Then, the sequences are grouped into clusters, enabling a consensus sequence to be generated for each group. This step is crucial to distinguish real variations from technical errors. The consensus sequences are then compared with reference databases via BLAST, enabling species identification and, in some cases, the construction of phylogenetic trees to explore genetic relationships.

Nanopore sequencing method offers great flexibility, low cost thanks to multiplexing, and the ability to analyze molecular changes directly. Although some challenges remain, notably related to sample purity and DNA fragmentation, it is a promising tool for accurate species identification in contexts such as food control.