



Phylogenetic analysis

In 1 collection

Joachim Nwezeobi¹, Onyeyirichi Onyegbule², Chukwuemeka Nkere², Joseph Onyeka², Sharon van Brunschot³, Susan Seal³, John Colvin³

¹National Institute of Agricultural Botany, Cambridge, United Kingdom,

²National Root Crops Research Institute, Umudike, Umuahia, Abia State, Nigeria,

³Natural Resources Institute, University of Greenwich, Central Avenue, Chatham Maritime, Kent, United Kingdom

1 Works for me dx.doi.org/10.17504/protocols.io.bd44i8yw

Joachim Nwezeobi

- 1 The partial mtCO1 sequences obtained from sequencing two randomly selected whiteflies from each of the 119 locations were first screened using quality checks in Geneious (v10.2.6) and those failing to meet the following criteria were discarded
- 2 First, the sequence chromatograms were manually checked and translated using the invertebrate mitochondrial genetic code. Poor quality sequences and sequences containing stop codons were excluded and the sequencing process repeated for these samples.
- 3 Two hundred and thirty-eight contig sequences were assembled from sequence data that passed the quality control and were trimmed to 657 bp and aligned with 691 *B. tabaci* mtCO1 reference sequences downloaded directly from GenBank.
- 4 The resulting alignment was used to plot an initial phylogenetic tree that was used to identify the whitefly genetic groups, using the FastTree program in Geneious (v10.2.6).
- 5 Twenty-nine unique haplotypes were identified in the 238 new sequences, which were extracted with Geneious 10.2.6 and used to generate the main phylogenetic tree. These unique sequences were combined with 18 selected reference sequences and aligned. A mtCO1 sequence of the spiralling whitefly *Aleurodicus dispersus* Russell (GenBank code: AJ748380) was used as an outgroup for the analysis.
- 6 The best model of nucleotide sequence evolution was evaluated using jModelTest (v2.1.10)
- 7 MrBayes (v3.2.6) was run using the nexus-formatted aligned sequences and the Generalised time-reversible (GTR) model [48,49] and Gamma distribution with invariant sites (G+I) as well as with four chains for 50 million generations and the trees were sampled every 1,000 generations.
- 8 The consensus tree (.nex.con.tre) file was used to view the phylogenetic tree and the posterior probabilities of the branches using FigTree (v1.4.3) [50]. The tree was further annotated using Adobe Illustrator (v22.1). The 29 unique haplotype sequences were submitted to the GenBank sequence database and were each assigned unique accession numbers.