

RESEARCH

Open Access



Analysis of *Thaumatotibia leucotreta* (Lepidoptera: Tortricidae: Olethreutinae) mitochondrial genomes in the context of a recent host range expansion

Bart T.L.H. van de Vossenbergh^{1†}, Tom H. van Noort^{2*†}, Sanne H.Z. Hooiveld-Knoppers^{1,2}, Lucas P. van der Gouw¹, Jan E.J. Mertens² and Antoon J.M. Loomans²

Abstract

Background The false codling moth (FCM), *Thaumatotibia leucotreta* (Meyrick, 1913), is a significant pest of various important economic crops and is a EU quarantine pest. In the last decade the pest has been reported on *Rosa* spp. In this study we determined whether this shift occurred within specific FCM populations across seven eastern sub-Saharan countries or whether the species opportunistically switches to this novel host as it presents itself. To achieve this, we assessed the genetic diversity of complete mitogenomes of *T. leucotreta* specimens intercepted at import and analysed potential linkages with the geographical origin and host species.

Results Genomic, geographical and host information were integrated into a *T. leucotreta* Nextstrain build which contains 95 complete mitogenomes generated from material intercepted at import between January 2013 and December 2018. Samples represented seven sub-Saharan countries and mitogenomic sequences grouped in six main clades.

Discussion If host strains of FCM would exist, specialization from a single haplotype towards the novel host is expected. Instead, we find specimens intercepted on *Rosa* spp. in all six clades. The absence of linkage between genotype and host suggests opportunistic expansion to the new host plant. This underlines risks of introducing new plant species to an area as the effect of pests already present on the new plant might be unpredictable with current knowledge.

Keywords False codling moth, *Rosa*, Polyphagy, Nextstrain

[†]Bart T.L.H. van de Vossenbergh and Tom H. van Noort these authors contributed equally.

*Correspondence:

Tom H. van Noort
t.h.vannoort@nvwa.nl

¹Molecular Biology group, Netherlands Institute for Vectors, Invasive plants and Plant Health, NVWA, Geertjesweg 15, 6706 EA, Wageningen, the Netherlands

²Entomology group, Netherlands Institute for Vectors, Invasive plants and Plant Health, NVWA, Geertjesweg 15, 6706 EA, Wageningen, the Netherlands



Background

The false codling moth (FCM), *Thaumatotibia leucotreta* (Meyrick, 1913) (Lepidoptera: Tortricidae: Olethreutinae), described from Transvaal, Pretoria, South Africa, is a significant pest of various fruit crops and cut flowers. FCM is considered a species of high phytosanitary risk because of its polyphagous nature, moderate dispersal capabilities and wide climate tolerance. FCM is native to the Afrotropical region [1, 2] and occurs across most of sub-Saharan Africa and surrounding islands [3] along a wide range of climate zones [4]. Its climatic tolerance and dispersal capabilities are illustrated in the South-African Cape region where it was first recorded as a pest species in the Eastern Cape province in 1957 [5] after which it is thought to have extended its range into the Mediterranean climate of the Western Cape region by the 1970s [6]. In 1984, FCM was observed in Israel for the first time [7] where it has since spread across the Israeli coastal plain [8]. Many countries have listed *T. leucotreta* as quarantine pest (EPPO, 2020) and took measures to prevent the introduction and spread of the species via infested fruits, vegetables, and cut flowers or branches into their territory (EFSA, 2020). FCM was declared a quarantine pest for the European Union in 2018 [9].

The dispersal capacity of FCM in agricultural environments is usually limited to the surrounding host plants [10]. In habitats with a lower density of suitable hosts, adults have been reported to disperse over distances of several hundred meters [11]. The moth can have four to ten generations per year depending on the climatic conditions [12]. Human-assisted dispersal may currently be the greatest driver of FCM's distribution and potential introduction as an early infestation is easily overlooked. The small, translucent, and flattened eggs are deposited directly on the surface of (ripening) fruits, petals, and sepals, and are often concolorous with their substrate [13]. Newly hatched larvae immediately bore into the fruits or (as only recently observed in rose) flower buds, leaving only a small penetration hole. Clear indications of an infestation, such as frass or discolouration, tend to become apparent during the later larval stages [14]. The final instar larva exits the host to pupate on or in the soil near the host plant, well-camouflaged inside a cocoon made up of soil particles and plant debris [15].

FCM is known to feed on a wide range of host plants native to southern Africa, representing more than 40 plant families [16–18]. It was first recorded as a pest species on crops in the early 1920s. It attacked mostly oranges and, to a lesser extent walnuts, olives, persimmons and acorns [19]. As it appears, the polyphagous nature of the moth allowed it to gradually include various cultivated, often non-native, crops such as cotton [20], bell pepper [21], stone fruits [22, 23], avocado [24], and of more recent date (2007) also cut flowers (rose) [25,

26]. The latter is noteworthy as it is not the fruit but the flower bud that is targeted, representing a difference host stage selection. Nowadays, FCM is known to have more than 130 host plant taxa in 100 genera of plants covering 51 different plant families of both cultivated and wild species [25]. As a polyphagous species that occurs throughout large parts of the African continent [27], FCM could have developed regional host preferences, potentially having evolved into different strains or subspecies. South-African FCM populations were found to exhibit geographical population structuring; at a regional level most of the molecular variation was explained by differences between individuals and not geographical distance or host preference [28]. Indeed, results on mating preference show some premise of assortative mating among geographically distinct population [29]. However, based on mating success, FCM shows no signs of reproductive isolation in South-Africa [30].

The question remains whether populations from other African countries have developed into distinct strains and whether these are linked with changes in host plant preference (distinct host strains). With the more recent host range expansion towards *Rosa* plants, the research presented here aims to answer whether this shift occurred within specific FCM populations across the Afrotropical region or whether the species opportunistically switches to this novel host when and where it occurs. We assessed the genetic diversity of complete mitogenomes of *T. leucotreta* specimens intercepted at import to identify potential linkages with geographical origin and host species. The mitochondrial genome is frequently used in phylogenetic and population-level studies [e.g. [31–33].] as it typically represents a single maternally inherited molecule with low levels of recombination while being present in high copy numbers [34, 35]. We used Nextstrain [36], a collection of open-source bio-informatic tools, to create a webtool displaying the mitogenomic diversity in context of geographic origin and host species, and end-users can interactively interrogate the data provided in the webtool. The *T. leucotreta* Nextstrain build can be accessed via <https://nextstrain.nrcnvw.nl/Tleuco/20200708>.

Results

Assembly and annotation of *T. leucotreta* mitogenomes *De Novo* assembly of Illumina data generated for the 126 *T. leucotreta* specimens resulted in assemblies containing 146 to 1,724 scaffolds representing 0.25 to 3.24 Mbp (median_{scaffolds} = 1,248; median_{assembly size} = 1.25 Mbp). Assemblies of the *T. leucotreta* specimens collectively resulted in 151,920 scaffolds of which 289 produced blastn hits with the *Helicoverpa armigera* (Hübner) mitogenome. These putative mitogenomic scaffolds ranged from 1,007 to 18,331 bp (median = 5,381),

of which 33 > 10 kbp scaffolds were used to build a chimeric consensus sequence that served as reference to obtain specimen-specific mitogenomic sequences (Fig. S1). Ninety-five of the 126 specimens produced reliable contiguous mapping of Illumina reads with > 10x mean read coverage. The obtained mitogenomes ranged from 15,041 to 15,397 bp in length with 21x to 1,443x mean read coverage.

The largest mitogenomic sequence (15,397 bp) was obtained for specimen 33,289,322 which was intercepted at import in a consignment of roses from Zambia (NCBI accession MT635674). This mitogenome encoding 37 genes, which are typically found on insect mitogenomes

(13 protein coding genes, 22 tRNAs, and 2 rRNAs), was used as reference for the species (Fig. 1).

Organization and orientation of the 37 mitogenomic genes were found conserved relative to publicly available complete Olethreutinae mitogenomes. In a single species (i.e. *Retinia pseudotsugaicola* Liu & Wu; NC_022865), one tRNA (tRNA-Leu) was predicted to be reverse-orientated relative to the other Olethreutinae species (Fig. S2). The “ATAGA+poly T” motif, which is conserved in lepidopteran mitogenomes and is believed to represent the origin of replication [37, 38] is located in the AT-rich control region between the tRNA-Met and 12 S rRNA genes.

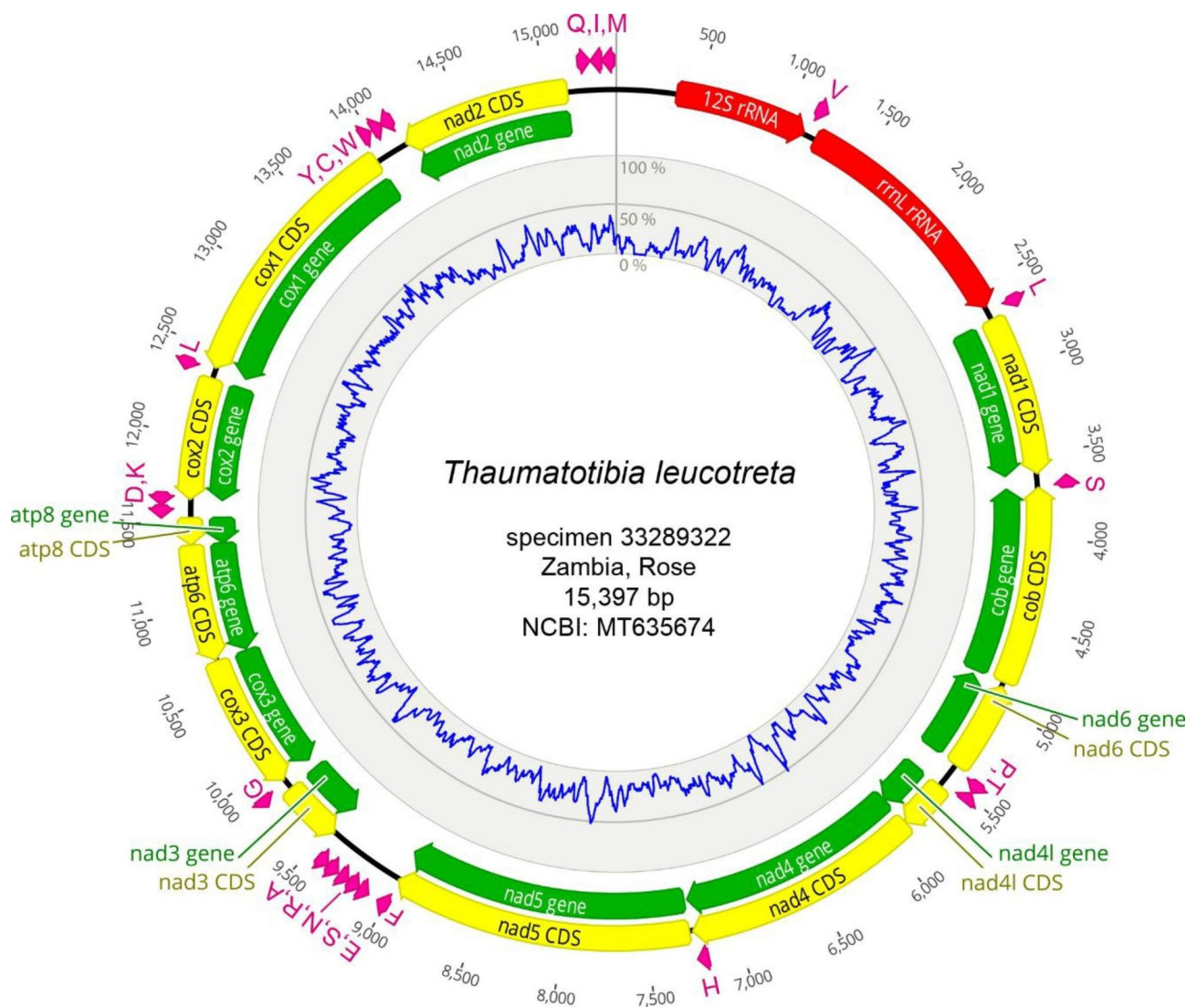


Fig. 1 The *Thaumatotibia leucotreta* mitogenome encoding thirteen protein coding genes (annotated in green, coding sequences are annotated in yellow), twenty-two transfer RNAs (purple), and two ribosomal RNAs (red). The direction of gene transcription is indicated by the arrows and annotation lengths are proportional to their nucleotide length. Transfer RNA genes are represented by the single letter IUPAC-IUB abbreviations for their corresponding amino acid. Abbreviations for protein coding genes: atp6, atp8, ATP synthase subunits 6 and 8 genes; cob, cytochrome oxidase b gene; cox1–cox3, cytochrome oxidase c subunit 1–3 genes; nad1–6, nad4L, NADH dehydrogenase subunits 1–6 and 4 L genes

Nextstrain build Genomic, geographical and host information were integrated in a *T. leucotreta* Nextstrain build which contains 95 complete mitogenomes generated from material sampled between January 2013 and December 2018. Information in the build is presented in three main panels: clustering of genomic diversity, geographical origin of the samples, and diversity relative to the mitogenome of specimen 33289322 (NCBI accession: MT635674). The associated metadata included in the build allows users to colour the external nodes in the tree according to the country of origin, host plant, material sampled, and the mitochondrial clade to which the specimen belongs. Internal node colours indicate the predicted ancestral state of a given trait, and the confidence of that state is conveyed by saturation of the colour of the internal node. Cladograms can be shown in different styles such as rectangular, radial and unrooted. The branch lengths of the tree can be shown based on divergence or in function of time. Based on the information provided in the build, Nextstrain determines the most likely transmission events, which can be animated from the webpage. The genotypes represented in the tree are plotted on a map, and users can set different levels of geographical resolution, i.e. continent and country. This allows simultaneous interrogation of phylogenetic and geographic relationships, with additional relevant metadata. Multiple filters can be applied to view (for instance) specific origin-host-mtDNA clade combinations. Users can download metadata and tree-files from the webpage, and can create screenshots of their views.

Diversity, and host and origin associations The mitogenomes assembled and annotated in this study are highly similar with 97.5–99.7% pairwise similarity relative to the reference sequence MT635674 (Fig. 2B). Nucleotide sequences of the thirteen protein coding genes were used to create a phylogenetic cladogram. Concatenation of the aligned gene sequences resulted in a 11,252 bp alignment with up to 116 single nucleotide polymorphisms (SNPs) among the *T. leucotreta* specimens, which equals 0–1.0% intra-species variation. Bayesian clustering of the sequences, using the leafroller species *Adoxophyes honmai* Yasuda as outgroup, results in grouping the *T. leucotreta* protein coding gene sequences in six main clades. Support for these clusters is high with 85–100% posterior probability scores for the major internal nodes (Fig. 2A). Most specimens were found to belong to mtDNA clades 1 and 2 with 44 and 27 specimens respectively. The RAxML clustering generated with entire mitogenomic sequences of the *T. leucotreta* specimens with Nextstrain is, apart from some branch swapping in the major mtDNA clades, identical to the thirteen gene Bayesian clustering.

Included in this study are *T. leucotreta* specimens from seven sub-Saharan African countries (Fig. 3A & B).

Mitogenomic diversity observed in the overall sample set is highest in South Africa with specimens from that country being represented in all six mtDNA clades. Specimens from other sub-Saharan countries were found representing two or more mtDNA clades. The three specimens obtained at an outbreak location in the Netherlands are found in mtDNA clade 2 and are 100% identical on a nucleotide level for the thirteen protein coding genes. The specimen from Israel is found in mtDNA clade 1 and matches closest to a specimen intercepted from a *Citrus sinensis* L. x Osbeck consignment from South Africa. There are no linkages observed between mitochondrial genotype and origin of the specimens (Fig. 3B). Similarly, no linkages were observed between mitochondrial genotype and host. Specimens found on citrus fruits and rose are present in all six haplogroups (Fig. 4A), whereas specimens on *Capsicum* are found in three of the four biggest haplogroups (Fig. 4B).

Discussion

We generated 95 complete mitogenomes of *T. leucotreta* specimens intercepted at import or sampled at outbreak locations. The import specimens represented seven sub-Saharan countries and 12 host plant species across eight genera. These sequences were used to determine the phylogenetic relationship between the specimens and to analyse potential linkages between host plant or country of origin. Based on these potential linkages we aimed to find an explanation for the occurrence of *T. leucotreta* interceptions on novel hosts, such as rose.

Mitogenome data generation

With genome sizes in Lepidoptera species ranging from 199 Mbp to 1,253 Mbp [39], our Illumina whole genome shotgun data resulted in assemblies with median size of 1.25 Mbp representing only a small proportion of the entire *T. leucotreta* genome. However, as the mitochondrial genome is present in multiple copies per cell, the entire mitogenome could easily be assembled even when few nuclear DNA sequences were assembled. Furthermore, even when using small amounts of biological material (e.g. a single egg) as input for the sequencing, library preparation and Illumina sequencing, the complete assembled and annotated mitogenome could be obtained. Insect mitochondrial genomes are highly conserved in Lepidoptera species [40], and the *T. leucotreta* mitogenome is no exception. The 37 genes typically found in arthropods are present in the mitogenome, i.e. 13 protein coding genes, 22 transfer RNAs and 2 ribosomal RNAs, and the organization and orientation of the mitogenomic components were found conserved relative to seven publicly available complete Olethreutinae mitogenomes.

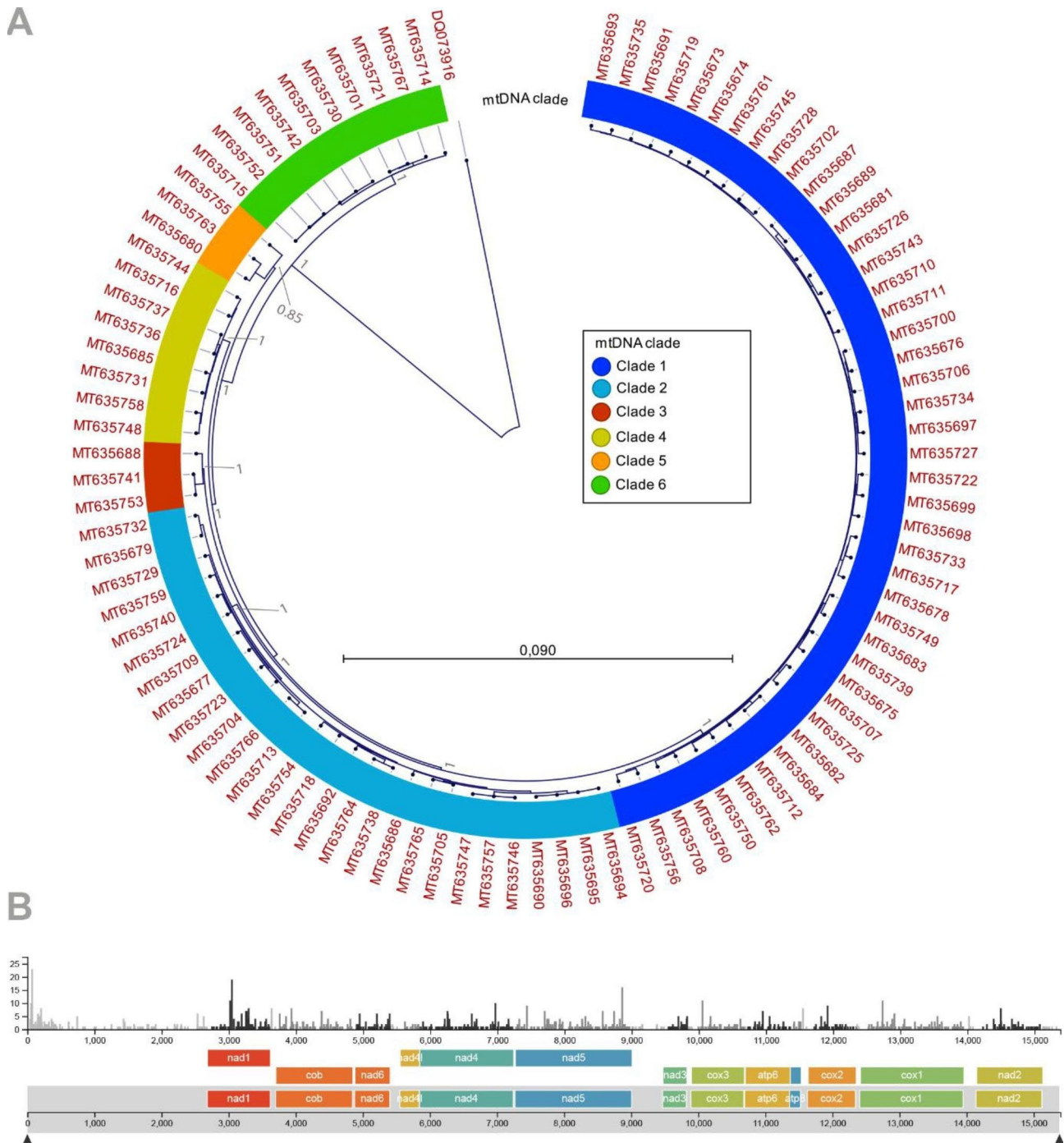


Fig. 2 *Thaumatotibia leucotreta* intra-specific variation **(A)** MrBayes tree based on nucleotide sequences of the thirteen mitochondrial protein coding genes of 95 *T. leucotreta* specimens using *Adoxophyes honmai* as outgroup. The external codes represent the NCBI accession of the complete mitogenomic sequences from which the protein coding genes were extracted. The coloured ring indicates the major clusters (mtDNA clades) observed in the circular phylogram. Posterior probabilities are shown for the major nodes, and the scale bar represents the number of substitutions per site. **(B)** Number of single nucleotide polymorphisms per mitogenomic position relative to the 15,397 bp reference sequence (MT635674) represented in the overall *T. leucotreta* dataset. The mitogenomic diversity plot is taken from the public *T. leucotreta* Nextstrain build described in this paper

Genetic structuring between countries

We found no proof of genetic structuring across the seven sub-Saharan countries when comparing the mitogenomes of FCM; all countries were represented by

at least two different clades of FCM, with South-African interceptions spanning all six clades. The active dispersal capabilities of FCM could, in theory, be sufficient to have regular mixing of genetic material among neighbouring

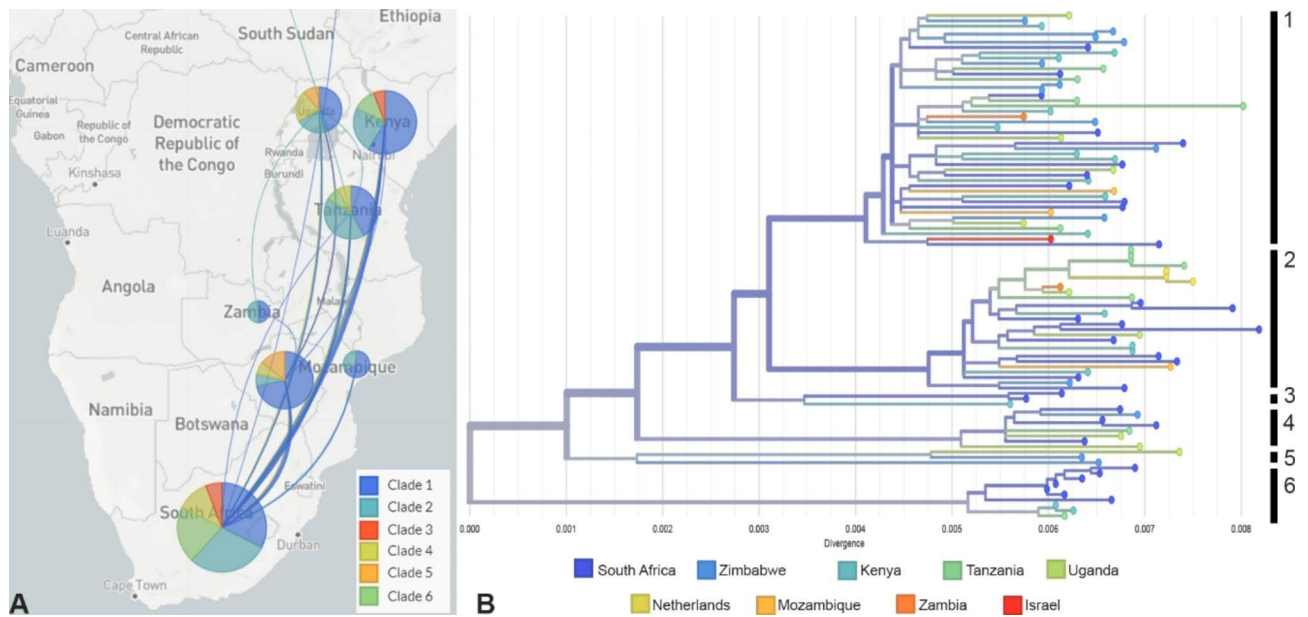


Fig. 3 **A.** Origins of sub-Saharan *T. leucotreta* specimens in this study and mtDNA clades represented from these specimens. Bubble size indicates the number of specimens included from each country of origin. The map was generated with Mapbox, which is included in the Nextstrain tool **B.** RAXML cluster based on entire mitogenomes with external nodes coloured based on the origin of the *T. leucotreta* specimen sampled. The vertical line represents the major mtDNA clades

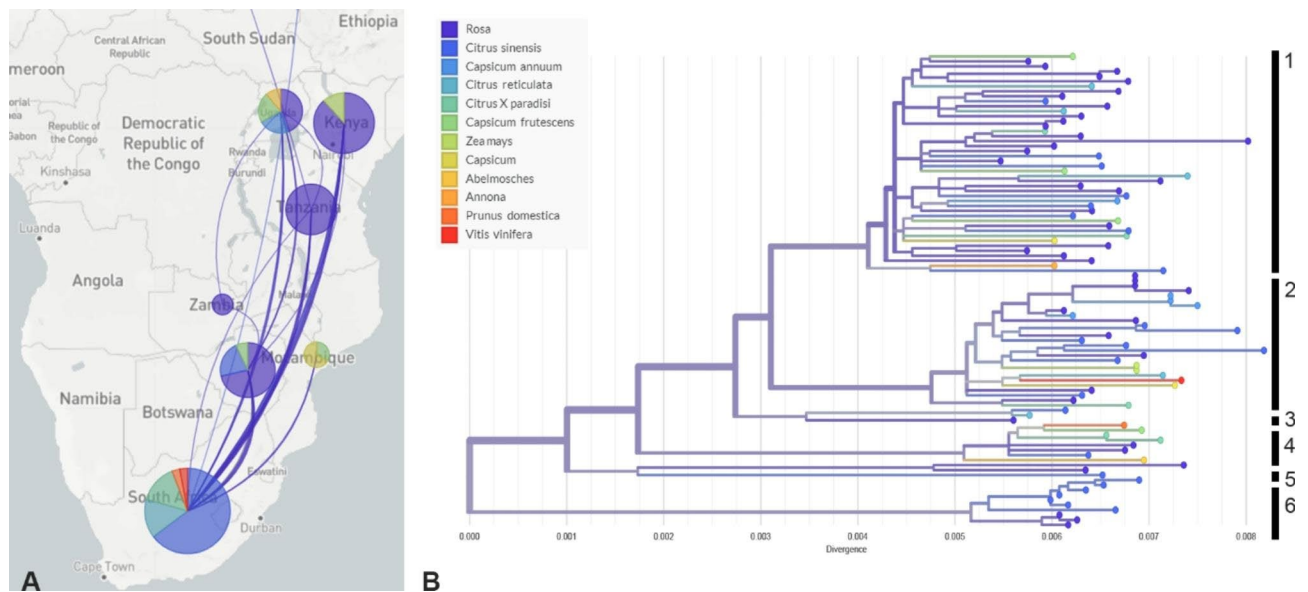


Fig. 4 **A.** Host plants from which the *T. leucotreta* specimens were obtained from the different sub-Saharan origins included in this study. Bubble size indicates the number of specimens included from each country of origin. The map was generated with the open source Mapbox, which is included in the Nextstrain tool **B.** RAXML cluster based on entire mitogenomes with external nodes coloured based on the host plant from which the *T. leucotreta* specimen was sampled. The vertical line represents the major mtDNA clades

populations or even countries [41]. However, FCM adults do not usually disperse far when host plants are abundant [10, 42]. This is further supported by regional genetical structuring of South African FCM populations [28] and genetic variation at different altitudes in Kenya and Tanzania [43]. A study on the migratory beetle armyworm (*Spodoptera exigua* (Hübner)) found little genetic

structuring among the populations that migrate due to the monsoon season, whereas the clade that undergoes no or little migration had a much higher genetic differentiation and phylogeographic structure [44].

FCM is not known to show migratory behaviour. Therefore, under natural circumstances, we would expect to find a degree of genetic structuring among the countries

included in this study. Instead, we find a higher genetic diversity between import consignments from within the same country of origin (e.g. all six clades in South-Africa) than those between the countries. This is in accordance with Mkiga et al. (2021) who found no genetic variability between populations in Tanzania and Kenya [43]. This absence of linkage between genotype and origin suggests that infected fruits and/or plants with soil have been transported between African countries. Citrus, as an example, has been introduced to sub-Saharan Africa on multiple occasions since late 15th century [45–47] and it is very likely to have crossed country borders from there onward, especially because of FCM's hidden nature as larvae inside the host plant and camouflaged pupae in the soil [14, 15, 48].

Strains linked to different host plants

Our results find no different FCM strains related to any of the 12 host plant species included in this study and *Rosa* was present as a host plant in all six clades. Moreover, specimens collected from the same host species are often less related to each other from a mitogenomic perspective than specimens collected on other hosts. The lack of linkage between genotype and plant host suggests FCM's capability to opportunistically switch host taxa, including non-native species such as citrus in the past and *Rosa* more recently. Many factors influence whether a newly introduced plant can be a suitable host: the phenology of both parties, presence of and resistance against plant defences, landscape composition (e.g. the presence and diversity of native host plants), etc. [49–51]. Generalist and more widespread species are more likely to prefer newly introduced hosts [52] and, as a study on *Vanessa cardui* (L.) suggests, could possess adaptations that allow them to include new plant host species regardless of their phylogenetic relation to already accepted hosts [53]. FCM has many traits to its advantage: a short generation time, the possibility to disperse moderate distances to find suitable hosts, and a broad range of potential host plants in several plant families including a plethora of natural host plants. FCM has attacked new hosts on a regular basis in the past [19–25] and although its fitness varies between hosts [54], it is clearly able to sustain viable populations on cultivated hosts, possibly shifting to native hosts when needed, without having evolved specific strains related to different host plants.

Almost all samples used in this study were obtained from interceptions at Dutch import locations. Consequently, there is a sampling bias towards countries that trade large volumes of commodities on which FCM is present with the Netherlands [55]; Eastern sub-Saharan countries are well-represented, whereas Western sub-Saharan countries were absent in our samples. Even though the current study does not represent FCM's

complete distribution range, the diversity observed among the south-eastern Sub-Saharan samples is large enough to justify the determination of linkages between mitogenomic diversity and geographic origin or host on which the specimen was found.

This study showcases a unique situation where we are able to document a pest's early inclusion of a novel host. Large-scale cultivation of cut flowers slowly shifted to tropical regions, including Africa during the late 1990s [56, 57]. Our oldest data point dates back to 2013, the first notifications of FCM on *Rosa* were made in 2007 from several African countries. If host strains of FCM would exist one would expect to find only one haplotype to be associated with this novel host. Instead, we find *Rosa* as a host in all six clades. Although our data does not focus on *Citrus* species, another important host taxon of FCM, it is likely that the association with *Citrus* and other host species is also opportunistic. It has to be noted that *Prunus* spp., and thus the Rosaceae family, have been a known host of *T. leucotreta* well before the first observations on *Rosa* sp. were made [16, 58]. Taxonomically, it is not unexpected that polyphagous species can feed on other genera within an already included plant family. However, on *Prunus* spp., the eggs are laid on young fruits and the larvae will bore into the fruits to feed and develop. In *Rosa* sp., no fruit develops and the larvae feed on the flower base and petals. This illustrates the ability of the larvae to feed on plant parts or developmental stages other than the soft tissue inside fruits, which is exceptional and was previously not known for *T. leucotreta*. This stipulates the risks of introducing new crop species to an area; the effect an already present polyphagous pest on the new crop might be unpredicted with current knowledge.

Methods

Insect specimens

A total of 126 *T. leucotreta* samples obtained from regular import inspections performed by the Dutch National Plant Protection Organization (NPPO-NL) from January 2013 to December 2018 were included in this study. Larvae and adults were identified as *T. leucotreta* using an identification key described by Gilligan and Epstein [12] also as implemented in EPPO standard PM7/137 (1) [26]. For selected specimens Sanger Sequencing of the partial mitochondrial *cox1* gene using primers LCO1490/HCO2198 according to the EPPO DNA Barcoding standard PM7/129(1) [59] was performed in support of species identification. Alternatively, eggs were identified as cf. *T. leucotreta* based on host plant and origin association. Specimens mostly originated from interceptions from sub-Saharan Africa, i.e. South Africa: 34; Kenya: 17; Zimbabwe: 14; Tanzania: 12; Uganda: 9; Mozambique: 3; and Zambia: 2. In addition, one specimen from Israel

and 3 specimens from an outbreak location in the Netherlands were included. Host plants from which the specimens were obtained belonged to several plant genera, i.e. *Rosa*: 43; *Citrus*: 35; *Capsicum*: 11; *Zea*: 2; *Abelmoschus*: 1; *Annona*: 1; *Prunus*: 1; and *Vitis*: 1. An overview of specimen information and associated molecular data is provided in Table S1.

DNA extraction and Illumina sequencing

Single eggs, larvae and parts of adult specimens were used as input for the DNA extraction. Insect tissue was ground in a lysis buffer with a micro-pestle prior to DNA extraction with the QuickPick SML genomic DNA kit (Bio-Nobile, Finland) on a KingFisher 96 flex (ThermoFisher, MA, USA). Genomic DNA was eluted in 100 μ L elution buffer, and aliquots were sent to GenomeScan (Leiden, the Netherlands) for 150PE Illumina sequencing with the NextSeq 500 V2 platform. In short, after fragmentation of the genomic DNA with the Bioruptor Pico (Diagenode, Belgium), sequencing libraries were prepared with the NEBNext Ultra DNA Library Prep kit for Illumina following manufacturer's instructions. Each sample was tagged to create a unique index-flowcell combination, and libraries were split over four sequencing lanes resulting in at least 2 Gb output per sample.

Mitogenome assembly and annotation

Illumina sequence reads were uploaded to CLC genomics workbench v11 and subjected to a pipeline that combined *de novo* assembly and blast-based identification of putative mitochondrial contigs. The pipeline consisted of the following analysis steps: trim reads (Quality trim: 0.05; Ambiguous limit: 2), *de novo* assembly (mapping mode: map reads back to contigs; update contigs: ON; length fraction: 0.8; similarity fraction: 0.9; minimum length: 1 kbp), extract consensus (threshold: 10; low coverage handling: split contigs), a size selector to discard consensus sequences shorter than 500 bp, and a blastn analysis of consensus sequences to the complete *Helicoverpa armigera* mitogenome (GU188273). Consensus sequences with an E-value < 1e-50 were regarded putative mitogenomic sequences. Putative mtDNA contigs > 10 kbp were used to create a single (chimeric) circular consensus mtDNA sequence in Geneious v11 (Biomatters, New Zealand). This was achieved by iteratively mapping putative mtDNA contigs to the largest scaffold, which allowed the identification and removal of non-*T. leucotreta* sequences. Sequences that wrapped around the 5' and 3' end of the largest scaffold were used to identify circularity of the mitogenome. Mitochondrial sequences for the individual *T. leucotreta* specimens were obtained by mapping reads of these specimens to the circular consensus mtDNA sequence (CLC Genomics workbench, length fraction: 0.8, similarity fraction:

0.9). Specimens with ≤ 10 x mean read coverage were excluded from the analysis, resulting in 95 specimens which were used in the phylogenetic and genotype association analyses. The 95 circular mitogenomes were annotated using the online MITOS v2 tool using the invertebrate mitochondrial code [60], and protein coding sequence annotations were manually checked to determine the reliability of the automated annotation. The annotated *T. leucotreta* mitogenomes were submitted in NCBI under accessions MT635673 - MT635767. The organization and orientation of *T. leucotreta* mtDNA genes was compared with that of other publicly available Olethreutinae mitogenomes using the MAUVE (version: 2015-02-26; algorithm: progressive) genome alignment tool [61] incorporated in Geneious v11, i.e.: *Celypha flavipalpana* H.-F. (NC_046051); *Cydia pomonella* (L.) (NC_020003); *Grapholita dimorpha* Komai (NC_024582); *Grapholita molesta* (Busck) (NC_014806); *Retinia pseudotsugaicola* (NC_022865); *Rhyacionia leptotubula* Liu & Bai (NC_019619); *Spilionota lechriaspis* Meyrick (NC_014294).

Nextstrain build

To identify potential links between genotypes and geographical origin or host plants, a Nextstrain build was created based on the complete mitogenomic sequences of the 95 *T. leucotreta* specimens. Augur (github.com/Nextstrain/augur), a bioinformatics toolkit for phylogenetic analysis used in the Nextstrain pipeline, was used to align the mitogenomic sequences with MAFFT [62] and to perform a RAxML clustering [63]. The initial tree was refined with metadata and a time tree was built using TreeTime [36] with optimization of scalar coalescent time, and using a fixed clock rate of 2.3×10^{-8} following Brower [64]. Internal nodes were assigned to their marginally most likely dates, and confidence intervals for node dates were estimated. Amino acid changes were determined based on the longest annotated *T. leucotreta* mitogenome, i.e. MT635674: 15,397 bp, specimen 33289322. Next, Augur output was exported and visualized in auspice. The *T. leucotreta* Nextstrain build is deposited on GitHub (<https://github.com/NPPO-NL/nextstrain-Tleuco>).

Clustering analysis

Sequences of the 13 protein coding genes were extracted from the mitogenome of each of the 95 specimens. Per gene a MAFFT v7.388 [62] alignment (algorithm: FFT-NS-1) was created using gene sequences from the *Adoxophyes honmai* (Lepidoptera: Tortricidae: Tortricinae) mitogenome (DQ073916=NC_008141) [65] as outgroup. The resulting thirteen alignments were concatenated and a Bayesian Inference of phylogeny was constructed with MrBayes v 3.2.6 [66] in Geneious v11 using the optimal

substitution model as defined in the model testing module of CLC genomics workbench (i.e. JC69 : nst=1; rate variation=equal). A total of 1×10^6 replications were run, and 10% was discarded as burn-in. Trees were sampled every 1,000 replications, and the consensus tree was exported to CLC genomics workbench to colour nodes according to the mtDNA clade to which they belonged. Posterior probability values were manually added in Adobe Illustrator.

Abbreviations

EFSa	European Food Safety Authority
EPPO	European and Mediterranean Plant Protection Organization
FCM	False Codling Moth
NVWA	Netherlands Food and Consumer Product Safety Authority
USDA-APHIS-PPQ	The United States Department of Agriculture - Animal and Plant Health Inspection Service - Plant Protection and Quarantine

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12862-023-02139-5>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We would like to thank Marja van der Straten (NVWA, the Netherlands) and Todd Gilligan (USDA-APHIS-PPQ, United States) for their contribution in species identification. We also thank Tslila Ben-David (Hebrew University of Jerusalem, Israel) for providing biological material from Israel. Finally, the laboratory support by Tim Warbroek and Naomi te Braak, and bioinformatic support by Michael Visser (NIPINVWA, the Netherlands) is greatly acknowledged.

Author contributions

BV and AL designed the study. TN collected and identified (where relevant) biological material. BV and SH organized sequence data collection. BV, SH and LG performed bioinformatic analyses. BV, TN, AL, JM drafted the manuscript. All authors contributed to the final version of the manuscript and gave final approval for publication.

Funding

This work was supported by the Netherlands Food and Consumer Product Safety Authority grant OS2018380c to Bart T.L.H. van de Vossenbergh. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The annotated *T. leucotreta* mitogenomes were submitted in NCBI under accessions MT635673 - MT635767. Sequence read archives 31054601 to 31054524 are submitted to NCBI under BioProject PRJNA938910.

Declarations

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Received: 9 February 2023 / Accepted: 23 June 2023

Published online: 03 July 2023

References

- De Jager ZM. Biology and ecology of the false codling moth, *Thaumatotibia leucotreta* (Meyrick). Stellenbosch: Stellenbosch University; 2013.
- Hill DS. Agricultural insect pests of the tropics and their control. CUP Archive; 1983.
- EPPO. : *Thaumatotibia leucotreta*. EPPO datasheets on pests recommended for regulation 2022.
- Rubel F, Brugger K, Haslinger K, Auer I. The climate of the European Alps: Shift of very high resolution Köppen-Geiger climate zones 1800–2100. *Meteorol Z*. 2017;26(2):115–25.
- Taylor JS. Notes on Lepidoptera in the Eastern Cape Province (Part IV). *J Entomol Soc South Afr*. 1957;20(2):315–32.
- Honiball SJ. Opnames van gasheerplante vir valskodlingmot (*Cryptophlebia leucotreta*) in die citrusdal-omgewing. CRI Group Annual Research Report. 2004;2004:140–2.
- Wysoki M. New records of lepidopterous pests of macadamia in Israel. *Phytoparasitica* 1986, 14(2).
- Kaspi R, Steinitz H, Nemny-Lavi E, Lebedev G, Melamed E, Gazit Y. Nontarget host Risk Assessment of the Egg Parasitoid *Trichogrammatoidea cryptophlebiae* (Hymenoptera: Trichogrammatidae) for classical Biological Control of the false Codling Moth (Lepidoptera: Tortricidae) in Israel. *J Econ Entomol*. 2019;113(2):1023–7.
- EU: Commission Implementing Regulation (EU) 2019/2072. In., vol. II. Online. ; 2019: 279.
- Timm AE. Morphological and molecular studies of Tortricid moths of economic importance to the south african fruit industry. Stellenbosch: University of Stellenbosch; 2005.
- Stotter RL. Spatial and temporal distribution of false codling moth across landscapes in the Citrusdal area (western Cape Province, South Africa). Stellenbosch: University of Stellenbosch; 2009.
- Gilligan T, Epstein M. Discovery of false Codling Moth, *Thaumatotibia leucotreta* (Meyrick), in California (Lepidoptera: Tortricidae). *Proc Entomol Soc Wash*. 2011;113:426–35.
- Stibick J. New Pest Response Guidelines: False Codling Moth *Thaumatotibia leucotreta*. In. Edited by Programs UAPEAd. Riverdale, Maryland; 2006.
- Moore SD. Moths and butterflies: False codling moth. In: *Integrated production guidelines for export citrus: integrated pest and disease management* Edited by Grout TG, vol. III. Nelspruit, South Africa: Citrus Research International (Pty) Ltd; 2019: 10.
- Daiber C. A study of the biology of the false codling moth [*Cryptophlebia leucotreta* (Meyr.)]: the cocoon. *Phytophylactica*. 1979;11(4):151–7.
- Brown J, Copeland R, Aarvik L, Miller S, Rosati M, Luke Q. Host records for fruit-feeding Afrotropical Tortricidae (Lepidoptera). *Afr Entomol*. 2014;22:343–76.
- Datasheet *Thaumatotibia leucotreta*. (false codling moth (FCM)) [<https://www.cabi.org/isc/datasheet/6904#top-page>].
- De Prins J, De Prins W. Afromoths, online database of afrotropical moth species (Lepidoptera). In <https://www.africanmoths.net>; 2019.
- Gunn D. The False Codling-Moth (*Argyroplote leucotreta*, Meyr.). *Union S Afric bept Agric* 1921, 21.
- Reed W. The false codling moth, *Cryptophlebia leucotreta* Meyr. (Lepidoptera: Olethreutidae) as a pest of cotton in Uganda. *Cotton Grow Rev*. 1974;51(3):213–25.
- Nyirira Z. Pest status of thrips and lepidopterous species on vegetables in Uganda. *East Afr Agricultural Forestry J*. 1973;39(2):131–5.
- Daiber C. A survey of the false codling moth (*Cryptophlebia leucotreta* Meyr.) in peach orchards. *Phytophylactica*. 1976;8(4):96–102.
- Blomefield T. Economic importance of false codling moth, *Cryptophlebia leucotreta*, and codling moth, *Cydia pomonella*, on peaches, nectarines and plums. *Phytophylactica*. 1989;21(4):435–6.
- Grové T, Steyn W, De Beer M. The false codling moth, *Cryptophlebia leucotreta* (Meyrick) (Lepidoptera: Tortricidae) on avocado: a literature review. *South Afr Avocado Growers' Association Yearbook*. 1999;22:31–3.
- Loomans EFSa, van Noort A, Schenk T, Delbianco M, Vos A. Pest survey card on *Thaumatotibia leucotreta*. EFSa Supporting Publications. 2020;17(8):1916E.

26. EPPO/OEPP. PM 7/137 (1) *Thaumatotibia leucotreta*. EPPO Bull. 2019;49(2):248–58.
27. Razowski J, Trematerra P. Tortricidae (Lepidoptera) from Ethiopia. J Entomol Acarological Res 2010, 42.
28. Timm AE, Geertsema H, Warnich L. Population genetic structure of economically important Tortricidae (Lepidoptera) in South Africa: a comparative analysis. Bull Entomol Res. 2010;100(4):421–31.
29. Upfold JK. Sexual attraction and mating compatibility between *Thaumatotibia leucotreta* populations and implications for semiochemical dependent technologies. 2019.
30. Mgocheki N, Addison P. Investigating the validity of the species status of the false codling moth in south african deciduous fruit orchards using mating studies and mtDNA. Bull Entomol Res. 2016;106(5):598–605.
31. Blacket MJ, Malipatil MB, Semeraro L, Gillespie PS, Dominiak BC. Screening mitochondrial DNA sequence variation as an alternative method for tracking established and outbreak populations of Queensland fruit fly at the species southern range limit. Ecol Evol. 2017;7(8):2604–16.
32. Hua J, Li M, Dong P, Cui Y, Xie Q, Bu W. Phylogenetic analysis of the true water bugs (Insecta: Hemiptera: Heteroptera: Nepomorpha): evidence from mitochondrial genomes. BMC Evol Biol. 2009;9:134.
33. van de Vossenbergh BTLH, Warbroek T, Ingerson-Mahar J, Waalwijk C, van der Gouw LP, Eichinger B, Loomans AJM. Tracking outbreak populations of the pepper weevil *Anthonomus eugenii* (Coleoptera: Curculionidae) using complete mitochondrial genomes. PLoS ONE. 2019;14(8):e0221182.
34. Friedman JR, Nunnari J. Mitochondrial form and function. Nature. 2014;505(7483):335–43.
35. Cameron SL. Insect mitochondrial Genomics: implications for evolution and phylogeny. Ann Rev Entomol. 2014;59(1):95–117.
36. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics. 2018;34(23):4121–3.
37. Taanman J-W. The mitochondrial genome: structure, transcription, translation and replication. Biochim et Biophys Acta (BBA) - Bioenergetics. 1999;1410(2):103–23.
38. Xin Z-Z, Yu L, Zhu X-Y, Wang Y, Zhang H-B, Zhang D-Z, Zhou C-L, Tang B-P, Liu Q-N. Mitochondrial genomes of two Bombycoidea Insects and Implications for their phylogeny. Sci Rep. 2017;7(1):6544.
39. Jia Z, Gao X, Xu Y, Tao H, Yang B, Gui M, Liu Y, Li Y. Analysis of the defence responses to Tomato spotted Wilt Orthotospovirus (TSWV) and Hippeastrum Chlorotic Ringspot Orthotospovirus (HCRV) infection in *Capsicum Annuum* based on Transcriptome. Res Square 2020.
40. Liu Q-N, Xin Z-Z, Bian D-D, Chai X-Y, Zhou C-L, Tang B-P. The first complete mitochondrial genome for the subfamily Limacodidae and implications for the higher phylogeny of Lepidoptera. Sci Rep. 2016;6(1):35878.
41. Baker EFSA, Gilioli R, Behring G, Candiani C, Gogin D, Kaluski A, Kinkar T, Mosbach-Schulz M, Neri O. F: *Thaumatotibia leucotreta*. Pest Report to support ranking of EU candidate priority pests. 2019.
42. Stotter RL, Terblanche JS. Low-temperature tolerance of false codling moth *Thaumatotibia leucotreta* (Meyrick) (Lepidoptera: Tortricidae) in South Africa. J Therm Biol. 2009;34(6):320–5.
43. Mkiga AM, Mohamed SA, du Plessis H, Khamis FM, Ekesi S. Bio-ecology of false codling moth, *Thaumatotibia leucotreta* (Meyrick) (Lepidoptera: Tortricidae) within citrus orchards in Kenya and Tanzania. Agric For Entomol. 2021;23(1):13–22.
44. Wang X, Yang X, Zhou L, Wyckhuys KA, Jiang S, Van Liem N, Vi LX, Ali A, Wu K. Population genetics unveils large-scale migration dynamics and population turnover of *Spodoptera exigua*. Pest Manag Sci. 2022;78(2):612–25.
45. Alpern SB. The european introduction of crops into West Africa in Precolonial Times. History in Africa. 1992;19:13–43.
46. Ravenstein EG. A journal of the first voyage of Vasco da Gama. Hakluyt Society; 2017. pp. 1497–9.
47. Webber HJ. History and development of the citrus industry. citrus Ind 1967:1–37.
48. Mazza G, Strangi A, Marinelli L, Del Nista D, Roversi PF. *Thaumatotibia leucotreta* (Meyrick) (Lepidoptera Tortricidae) intercepted for the first time in Italy. Redia. 2014;97:147–9.
49. Pearse IS, Harris DJ, Karban R, Sih A. Predicting novel herbivore–plant interactions. Oikos. 2013;122(11):1554–64.
50. Castells E, Morante M, Goula M, Pérez N, Dantart J, Escolà A. Herbivores on native and exotic *Senecio* plants: is host switching related to plant novelty and insect diet breadth under field conditions? Insect Conserv Divers. 2014;7(5):420–31.
51. Yang L, Liu H, Pan Y, Li H, Lu Y. Landscape simplification increases the risk of infestation by the polyphagous pest *Helicoverpa armigera* for walnut, a novel marginal host. Landsc Ecol 2022.
52. Jahner JP, Bonilla MM, Badik KJ, Shapiro AM, Forister ML. Use of exotic hosts by Lepidoptera: widespread species colonize more novel hosts. Evolution. 2011;65(9):2719–24.
53. Celorio-Mancera MdP, Wheat CW, Huss M, Vezzi F, Neethiraj R, Reimegård J, Nylin S, Janz N. Evolutionary history of host use, rather than plant phylogeny, determines gene expression in a generalist butterfly. BMC Evol Biol. 2016;16(1):59.
54. Mkiga AM, Mohamed SA, du Plessis H, Khamis FM, Ekesi S. Field and laboratory performance of false codling moth, *Thaumatotibia leucotreta* (Lepidoptera: Tortricidae) on orange and selected vegetables. Insects. 2019;10(3):63.
55. Adom M, Fening K, Billah M, Wilson D, Hevi W, Clottey V, Ansa-Amprofi F, Bruce A. Pest status, bio-ecology and management of the false codling moth, *Thaumatotibia leucotreta* (Meyrick) (Lepidoptera: Tortricidae) and its implication for international trade. 2020.
56. Steen M. A world of flowers: dutch flower auctions and the market for cut flowers. J Appl Hortic. 2010;12(2):113–21.
57. UNCTAD-ICC. An investment guide to Ethiopia: opportunities and conditions. In. UN: ICC; 2004.
58. Prinsloo G, Uys V. Insects of cultivated plants and natural pastures in Southern Africa. Hatfield, South Africa: Entomological Society of Southern Africa; 2015.
59. EPPO/OEPP. PM 7/129 (1) DNA barcoding as an identification tool for a number of regulated pests. EPPO Bull. 2016;46(3):501–37.
60. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013;69(2):313–9.
61. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004;14(7):1394–403.
62. Katoh K, Misawa K, Kuma K-i, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059–66.
63. Stamatakis A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312–3.
64. Brower A. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences* 1994, 91(14):6491–6495.
65. Lee ES, Shin KS, Kim MS, Park H, Cho S, Kim CB. The mitochondrial genome of the smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidae). Gene. 2006;373:52–7.
66. Huelsenbeck JP, Ronquist F. MRBAYES: bayesian inference of phylogenetic trees. Bioinformatics. 2001;17(8):754–5.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.