Integrative approach to discovering species diversity within the Mediterranean group of the *Bemisia tabaci* complex

SOŇA VYSKOČILOVÁ

A thesis submitted in partial fulfilment of the requirements of the University of Greenwich for the Degree of Doctor of Philosophy

April 2019

DECLARATION

I certify that the work contained in this thesis, or any part of it, has not been accepted in substance for any previous degree awarded to me, and is not concurrently being submitted for any degree other than that of Doctor of Philosophy being studied at the University of Greenwich. I also declare that this work is the result of my own investigations, except where otherwise identified by references and that the contents are not the outcome of any form of research misconduct.

_____	_____
Soňa Vyskočilová	Date
Student	
_____	_____
Prof John Colvin	Date
First supervisor	
_____	_____
Prof Susan Seal	Date
Second supervisor	

ACKNOWLEDGEMENTS

My profound gratitude belongs to my supervisors, Prof John Colvin and Prof Susan Seal, who nurtured my professional growth and made me feel welcome and valued since day one. I thank the University of Greenwich and the African Cassava Whitefly Project funded by The Bill and Melinda Gates Foundation for their financial support throughout the four years of my doctoral studies.

I would like to acknowledge Dr Sharon van Brunschot and Dr Hua-Ling Wang for their help in generating the high-throughput sequencing data, as well as for numerous other occasions they advised me during my laboratory, insectary and/or computational work. I am grateful to Dr Paul Visendi who trained me in quality control of sequencing reads and performed the read mapping to endosymbiont genomes. I acknowledge Mitulkumar Patel and Nikunj Naik for extracting candidate genes from MED transcriptomes and Stephen Young for help with statistical analyses. My thanks also belong to Dr Habibu Mugerwa, Dr Moritz Bömer, Dr Gonçalo Silva, Dr Sophie Bouvaine, Rebecca Grimsley, Natalie Morley and Dr Simon Springate for their help and support during my laboratory and/or insectary work at NRI. I am indebted to Dr Wee Tek Tay for his expertise and encouragements that helped me publish, as well as his help with the sliding window analysis and facilitating my visit to Australia. I also acknowledge CSIRO Health and Biosecurity for hosting my visit in Canberra and its staff for fruitful discussions. Lastly, I appreciate the opportunity to mentor the intern Quentin Bédrune, who, besides broadening my music horizons, helped me to handle the insectary workload.

I am very thankful for all my postgraduate peers inside and outside the Tower. I am proud of the NRI Postgraduate Society, the Medway Postgraduate Café and our regular events, which brought all students and staff closer together and made this journey much more enjoyable for all of us. I deeply wish that these efforts will keep their momentum and continue to be shaped by new “generations” of postgraduate students at NRI and the Medway campus.

I am also grateful to Grzegorz Caban, without whom I would not have embarked on this PhD project; to my good friends inside and outside the campus always ready to listen and cheer me up; to London salsa scene occasionally unsticking me from the office chair; and to my family loving and supporting me all the way.

ABSTRACT

Bemisia tabaci is a complex of cryptic whitefly species, which includes some of the world's most damaging agricultural pests. The Mediterranean (MED) putative species is globally invasive and its populations are often resistant to insecticides. The intra-species genetic variability identified in partial sequences of the mitochondrial cytochrome *c* oxidase 1 (mtCOI) gene has led to the recognition of four MED groups: Q1, Q2, Q3 and African silver-leafing (ASL). A lack of hybridisation between Q1 and ASL populations has been reported, but the taxonomic and biological significance of these groups remained unclear.

The aim of this study, therefore, was to evaluate the species status of MED groups using an integrative approach, combining (i) molecular analyses of high-throughput sequencing-derived mitogenomes, (ii) reciprocal crossing experiments to investigate reproductive compatibility among Q1, Q2 and ASL populations and developing a molecular marker for hybrid verification, (iii) detection of bacterial endosymbionts and (iv) bioassays to compare their host-plant ranges and performances on 13 plants. Our mitogenome phylogeny showed close relationships among Q1 and Q2 populations, while ASL was placed outside the Q1/Q2 cluster with 100% bootstrap support. Using the mitogenomes as a reference enabled the identification of sequence errors in 155 of 289 published MED mtCOI haplotypes. Crossing experiments revealed that only Q1 from Spain and Q2 from Israel were compatible, confirming that they belong to the same species. In contrast, ASL from Uganda and Q1 from Sudan both failed to interbreed with any other population. Parental origin of the Q1xQ2 hybrids was verified by the novel nuclear marker GC1 and their fertility by backcrossing. The F_2 offspring showed asymmetry in numbers and sex ratio. Hypotheses were formulated about the involvement of endosymbionts *Rickettsia* and/or *Hamiltonella* in the F_2 asymmetry, and *Wolbachia* and/or *Cardinium* in the incompatibility between Spain Q1 and Sudan Q1 populations. Lastly, significant differences ($P<0.05$) in host use occurred amongst all four populations. Spain Q1 had the widest host range and Israel Q2 the narrowest, but the most distinct pattern of host use was observed for Uganda ASL. In addition, despite its name, we found no evidence of the capacity of ASL to induce silver-leafing symptoms in squash.

From this combined evidence, we conclude that the ASL group belongs to a distinct biological species and we highlight the inconsistency between the biological species concept and the currently used species delimitation based on the partial mtCOI sequence and 3.5% nucleotide distance threshold. Accurate knowledge of the number and biology of cryptic species will allow more targeted and efficient design of pest control strategies.

CONTENTS

ABBREVIATIONS

A	adenine
AA	amino acid
ASL	African silver-leafing
ATP	adenosine triphosphate
BIC	Bayesian Information Criterion
BLAST	Barcode of Life Data System
BOLD	Basic Local Alignment Search Tool
bp	base pair
C	cytosine
CABI	Centre for Agriculture and Bioscience International
CAPS	cleaved amplified polymorphic sequences
CarE	Carboxylesterase
CBSD	cassava brown streak disease
CLCuD	cotton leaf curl disease
CMD	cassava mosaic disease
Da	dalton
DNA	deoxyribonucleic acid
dsRNA	double stranded RNA
EPG	electrical penetration graph
F_1	the first filial generation
F_2	the second filial generation
FL	Florida
FU	fluorescence unit
G	guanine
GC1/2	Glutamate carrier 1/2
GST	Gluthatione-S-transferase
HTS	high-throughput sequencing
INDEL	insertion/deletion
IO	Indian Ocean
IPM	integrated pest management
ITS1	Internal transcribed spacer 1
IUPAC	International Union of Pure and Applied Chemistry
kb	kilobase
kDa	kilodalton
L:D	light:dark
MEAM1	Middle East-Asia Minor 1
MED	Mediterranean

mtCOI	Mitochondrial cytochrome *c* oxidase 1
mtDNA	mitochondrial DNA
NADH	nicotinamide adenine dinucleotide
NCBI	National Center for Biotechnology Information
ND-75	NADH-Q-oxidoreductase, 75kDa subunit gene
NEB	New England Biolabs
nt	nucleotide
NUMT	nuclear mitochondrial DNA segment
NW	New World
P450(s)	Cytochrome P450 monooxygenase(s)
PCR	polymerase chain reaction
prp8	Pre-mRNA processing factor 8 gene
qPCR	quantitative polymerase chain reaction
RAPD	randomly amplified polymorphic DNA
rDNA	ribosomal DNA
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNAi	RNA interference
RNApylI	RNA polymerase II gene
SC	species concept
SCAR	sequence characterised amplified regions
SdhD	Succinate dehydrogenase, subunit D gene
SG	subgroup
shaw	Shaker cognate w gene
SNP	single nucleotide polymorphism
SpQ1	Spain Q1
SSA	sub-Saharan Africa
SuQ1	Sudan Q1
T	thymine
TMV	*Tobacco mosaic virus*
tRNA	transfer RNA
TYLCD	tomato yellow leaf curl disease
TYLCV	*Tomato yellow leaf curl virus*
U	uracil
UGT	UDP-glucosyltransferases
UK	United Kingdom
US	United States
USD	US dollar

TABLES

FIGURES

1. Introduction

Agricultural sustainability and food security are threatened by multiple factors, such as climate change, drought, loss of pollinators and soil erosion, among others. Other important factors are the pests, weeds and pathogens that compete with humans for food. Herbivorous insects and plant pathogens have caused devastating losses and famine in the past, such as the 19th century Irish potato famine caused by a fungus (Turner, 2005), insect-borne viral diseases of cassava threatening food security in sub-Saharan Africa (Legg and Fauquet, 2004; Alicai *et al.*, 2007), or swarms of locusts causing widespread starvation throughout human history (Krall, 1995; Lenné, 2000).

One of the most damaging agricultural pests worldwide is the whitefly (Hemiptera: Aleyrodidae) species complex *Bemisia tabaci sensu lato,* which is listed among the 100 worst invasive organisms in the world (Lowe *et al.*, 2004). Members of the *B. tabaci* species complex cause plant damage by feeding on the phloem sap, inducing phytotoxic disorders, depositing honeydew on which sooty mould develops and by vectoring over 200 species of begomoviruses, criniviruses, carlaviruses, ipomoviruses and torradoviruses that infect plants (Jones, 2003; Seal *et al.*, 2006; Navas-Castillo *et al.*, 2011; Gilbertson *et al.*, 2015). Diseases caused by these viruses can lead to devastating losses of up to 100% of the yield (Pennisi, 2010; Cathrin and Ghanim, 2014).

B. tabaci was first described as an indigenous tobacco pest in Greece by Gennadius in 1889. Later the *B. tabaci* complex gained global economic importance due to the intercontinental spread of two putative species, currently called “Middle East-Asia Minor 1” (MEAM1) and “Mediterranean” (MED) (Dinsdale *et al*., 2010). These members of the complex are particularly threatening because of their invasiveness, ability to feed on a wide range of plants (polyphagy) and insecticide resistance.

Documented outbreaks of MEAM1 started in the 1980s and this species has since spread to at least 54 countries in six continents (De Barro *et al.*, 2011). Its invasiveness, high fecundity, polyphagy and resistance to some insecticides was unprecedented and led to a higher awareness and more frequent field screening for whiteflies. Since the late 1990s and early 2000s, the second invasive *B. tabaci* species, MED, was identified and has spread to five continents, apart from Australia (Dennehy *et al.*, 2005; Zhang *et al.*, 2005; Grille *et al.*, 2011; Gnankiné *et al.*, 2013). The major worry about MED is its very low susceptibility to insecticides, including the neonicotinoid group (Cahill *et al.*, 1996; Elbert and Nauen, 2000; Nauen *et al.*, 2002; Rauch and Nauen, 2003). MED also vectors economically important plant viruses such as *Tomato chlorosis virus*, *Tomato yellow leaf*

curl virus or *Papaya leaf curl China virus* (Navas-Castillo *et al.*, 2000; Shi *et al.*, 2014; Guo *et al.*, 2015).

Biological differences between pest species can have significant implications for their control. Traits such as the range of colonised crops, the diversity of transmitted plant viruses and the susceptibility to insecticide active ingredients and biological control agents can all vary among species. The main prerequisite to understanding these differences, however, is the delimitation of populations of these organisms that constitute different species. The definition of a species is a contentious topic and the preferred interpretation varies among fields. Historically, delimiting different species was performed predominantly on the basis of morphological features (Wilkins, 2009). Another important concept in defining species has been the biological species concept, which defines species as a group of individuals that can breed among each other, but not with other groups (Mayr, 1942, 1969). More recently, the access to sequencing technologies has led to the development of standard DNA markers for classifying extant, and even extinct, species (e.g. Miller *et al.*, 2009; Dalén *et al.*, 2017).

Mitochondrial DNA (mtDNA) is the most widely used of DNA regions for phylogenetic analyses among insects and animals in general, because it has some desirable properties, such as high copy number, simple genetic structure and a maternal mode of inheritance (Avise *et al.*, 1987; Zhang and Hewitt, 1996). Among nuclear loci, only ribosomal DNA (rDNA) genes have been widely used due to their higher copy number and the ease of amplification and sequencing, which is similar to mtDNA (Caterino *et al.*, 2000). For both mitochondrial cytochrome *c* oxidase 1 (mtCOI) and 16S rDNA genes, it is also relatively straightforward to use universal primers across taxa (Caterino *et al.*, 2000; Cameron, 2014).

The popularity of mtDNA sequences was further enforced by the development of the molecular technique for species cataloguing and assignment called DNA barcoding (Hebert *et al.*, 2003a) and the Barcode of Life Data System (BOLD) platform (Ratnasingham and Hebert, 2007). The process involves sequencing of a standardised region in the organism's DNA and comparing it to the DNA barcode database to enable specimen identification. For insects, the region used for DNA barcoding is a partial sequence of the mtCOI gene (Wilson, 2012), which has since become a near-exclusive data source for species identification and even delimitation (Cameron, 2014).

Putative species within the *B. tabaci* complex have been delineated predominantly on the basis of their partial mtCOI sequence, because these species lack reliable morphological

distinguishing features (Brown *et al.*, 1995b; Rosell *et al.*, 1997). Attempts have been made to establish a cut-off threshold for species delimitation, calculated as a nucleotide distance within the mtDNA marker. Dinsdale *et al.* (2010) identified a threshold at 3.5% nucleotide distance in a 657 bp partial sequence at the 3' end of the mtCOI gene. This threshold has subsequently been adopted by other authors in classification of field-collected samples and announcing new putative species, if the sample was over 3.5% divergent from other putative species. This approach has led to the identification of 46 putative species within the *B. tabaci* complex to date (Dinsdale *et al.*, 2010; Hu *et al.* 2011; Alemandri *et al.*, 2012; Chowda-Reddy *et al.*, 2012; Parrella *et al.*, 2012; Esterhuizen *et al.*, 2013; Firdaus *et al.*, 2013; Tahiri *et al.*, 2013; Hu *et al.*, 2014, 2017; Mugerwa *et al.*, 2018) (Table 2.2).

It is, however, possible that the number of putative species is over-estimated due to the inclusion of pseudogene sequences in the barcoding databases (Song *et al.*, 2008; Tay *et al.*, 2017a). Pseudogenes are sequences that originated as copies of functioning genes, but have since accumulated mutations and become dysfunctional (Jacq *et al.*, 1977). In the case of the mtDNA, they are called nuclear mitochondrial DNA segments (NUMTs) and originated from transposition of a portion of mtDNA into nuclear DNA (Lopez *et al.*, 1994). Among insect orders, NUMTs have been reported in the Orthoptera, Hymenoptera and Hemiptera (Bensasson *et al.*, 2001). Specifically, hemipteran species with reported NUMTs include aphids from the genus *Sitobion* (Sunnucks and Hales, 1996) and *B. tabaci* MEAM1 (Tay *et al.*, 2017a). The NUMT segments can be similar enough to mtDNA to be picked up undetected while obtaining mtDNA marker sequences, but different enough to provide false signals about the sequence divergence and evolutionary histories (Zhang and Hewitt, 1996; Song *et al.*, 2008).

A subset of the 46 putative species of the *B. tabaci* complex has been studied in the context of the biological species concept (Xu *et al.*, 2010; Liu *et al.*, 2012b). Reciprocal crossing experiments were performed to ascertain whether the groups delimited on the basis of the 3.5% threshold represented biological species that are reproductively isolated from one another. In most cases, usually between the invasive MEAM1 or MED and an indigenous putative species, a reproductive incompatibility was observed (Demichelis *et al.*, 2005; Zang and Liu, 2007; Wang *et al.*, 2010; Sun *et al.*, 2011; Wang *et al.*, 2011a). However, some putative species were only partially compatible (De Barro and Hart, 2000; Delatte *et al.*, 2006; Qin *et al.*, 2016).

The MED putative species displayed one of the highest intraspecific genetic variabilities within the *B. tabaci* complex (Lee *et al.*, 2013). A phylogenetic analysis of the 3' partial mtCOI sequences revealed four intra-species groups called Q1, Q2, Q3 and African Silver-leafing (ASL) within the MED putative species (Gueguen *et al.*, 2010). These groups harboured specific combinations of endosymbiotic bacteria and, typically, occupied particular regions in the Mediterranean basin and sub-Saharan Africa (Gueguen *et al.*, 2010). The bacterial endosymbionts are known to influence the fitness and reproductive system of their insect hosts and these changes can be both advantageous as well as disadvantageous (Kikuchi, 2009).

Reproductive compatibility amongst the MED groups is not yet fully understood, but has been studied to a limited extent by field population genetic studies. Hybrids between the Q1 and Q2 groups from western and eastern Mediterranean Basin, respectively, were reported (Gauthier *et al.*, 2014; Terraz *et al.*, 2014; Thierry *et al.*, 2015), while no evidence of gene flow was found for the Q1 and ASL populations living in sympatry in West Africa (Mouton *et al.*, 2015). In addition, indications of differences in the host-plant range among the MED groups were also reported (Sseruwagi *et al.*, 2005; Parrella *et al.*, 2014; Malka *et al.*, 2018; Mugerwa, 2018).

The genetic and biological heterogeneity within the MED putative species raises questions concerning the species status of MED and the accuracy of the species delimitation threshold in the partial mtCOI sequences. There are gaps in the knowledge of evolutionary relationships between populations of the MED putative species, which is now present on almost every continent in the world and threatens food production. The data on reproductive compatibility among MED groups are incomplete and there is little information about the differences in their host-plant range. Such knowledge gaps in the genetics, biology and ecology of MED slow down the development of efficient pest management strategies (Stansly and Naranjo, 2010).

An overall aim of this PhD study was to generate the molecular, biological and ecological data in an integrative approach to evaluating the species status of mtCOI groups within the putative MED species.

The research questions addressed in chapters to follow were:

1. How closely are the MED groups related at the mtDNA level? (Chapter 3)
2. How reliable are the commonly used barcoding methods for identifying biological species? (Chapter 3)

3. Are the MED groups reproductively compatible? (Chapter 4)
4. Do endosymbionts play a role in the reproductive relationships among the MED groups? (Chapter 4)
5. Do the MED groups share the same host-plant range and do they perform equally well on the host-plants? (Chapter 5)

The genetic relationships among the MED groups were studied by (i) comparing their mitogenomes, which were assembled from high-throughput sequencing (HTS) data to avoid the inclusion of NUMTs, and (ii) by a phylogenetic analysis of multiple mitochondrial genes. The reliability of the current barcoding method was addressed by comparing the nucleotide divergences within and between key regions in the mitogenome, as well as by carrying out a quality assessment of published mtCOI sequences of the MED groups. The reproductive compatibility among MED groups was studied in reciprocal crossing experiments coupled with rigorous molecular verification of hybrid individuals and testing the fertility of the hybrid F_1 offspring. Bacterial endosymbionts infecting our model populations were detected to facilitate formulation of hypotheses about their potential effect on their host's life and reproduction. Lastly, the host-plant ranges and performances were evaluated in bioassays measuring multiple traits on a range of plants and comparing these among the MED populations.

2. Literature review

2.1 Systematics of the *B. tabac*i species complex

Bemisia tabaci refers to a complex of species (Frohlich *et al.*, 1999; Xu *et al.*, 2010; Dinsdale *et al.*, 2010; Boykin *et al.*, 2012) belonging to the family Aleyrodidae of the superfamily Aleyrodoidea. The superfamily comprises 1556 species of whiteflies belonging to 161 genera (Martin and Mound, 2007). Together with families of aphids, psyllids and scale insects, whiteflies belong to the suborder Sternorrhyncha in the order Hemiptera. The *B. tabaci* species complex, or some of its members, are also known under common names such as sweet potato whitefly, tobacco whitefly, cotton whitefly, silverleaf whitefly, poinsettia whitefly or cassava whitefly (CABI, 2018).

The species within the *B. tabaci* complex are morphologically indistinguishable (Brown *et al.*, 1995b; Rosell *et al.*, 1997), which poses a significant challenge to resolving their systematics. Since the first species description, the systematics of *B. tabaci* has changed several times (see section 2.1.2). Our understanding of how many species there are within the complex and how they can be identified accurately is still ongoing.

2.1.1 Species conceptualisation and species delimitation

An overview of species concepts and delimitation

Species is the basic unit in many fields of biology, particularly in taxonomy (Zachos, 2015). The definition of species as a taxonomic category, however, has triggered many long and heated debates in the past and remains unresolved (Zachos, 2018). There are at least 34 different species concepts (SCs) (Zachos, 2016) and different fields have different needs and preferences for defining species. There are several basic ideas that underlie the diverse species concepts, such as reproductive isolation, ecological isolation, evolutionary history, and the convenience of biologists (Wilkins, 2009; Morrison, 2011).

The fundamental concept of species generally accepted among biologists was formulated by the evolutionary SC, which regards a species as *"a single lineage of ancestral descendant populations of organisms which maintains its identity from other such lineages and which has its own evolutionary tendencies and historical fate"* (Wiley, 1978). This is a primary SC, and majority of other SCs agree in the primary property of species to be "separately evolving meta-population lineages"; however, they disagree in the secondary species criteria (e.g. reproductive isolation, diagnosability or monophyly) (De Queiroz, 2007). This dispute will likely never be fully resolved because of the inherent limitation of taxonomy, which is trying to fit discrete categories of species to the

continuous process of evolution (Zachos, 2018). This means that a completely objective species definition will never be achieved and we must accept the “fuzzy” boundaries in nature (Zachos, 2018).

The species concept debate is a rather philosophical dispute about the concept of the species category. It is not, however, the same as species delimitation, which is a methodological problem for inferring the numbers of species and the boundaries between them (De Queiroz, 2007; Boykin *et al.*, 2012). Species delimitation is also highly relevant to areas outside science, such as conservation policy, management and environmental legislation (Zachos, 2015). A unified SC has been proposed, in which only the primary property is necessary to define a species, while all the secondary criteria originating from alternative SCs serve as different lines of evidence of the lineage separation and thus also segregating the two problems of species conceptualisation and species delimitation (De Queiroz, 2007). Indeed, several authors have proposed that an ideal approach for an accurate delimitation of animal species would integrate multiple lines of evidence, such as reproductive compatibility, morphology, ethology, ecology and molecular markers (Perring *et al.*, 1993; Smith *et al.*, 2008; Leaché *et al.*, 2009).

Biological species concept

During the modern synthesis, in which Darwin’s theory of evolution and Mendelian genetics were connected, the biological SC was introduced (Mayr, 1942) and became dominant for about half a century (Groves and Grubb, 2011). This concept defines species as a group of individuals that can breed together, but cannot with other groups, *i.e.* the group is isolated reproductively and genetically from other groups (Mayr, 1942, 1969).

Mayr viewed species as discrete, reproductively isolated entities, but this view was later criticised for the lack of acknowledgement of the continuity between varieties and species (Mallet, 2008). There are many examples of species boundaries that are “porous” and lineages that diverged while maintaining some level of gene flow (Harrison, 1990; Wu, 2001; Harrison and Larson, 2014). Due to introgressive hybridisation, the observed patterns of gene exchange are inconsistent across genome, which makes the permeability of species boundaries a function of genomic region (Harrison and Larson, 2014). It has also been proposed that those genomic regions that consistently show low levels of introgression probably contain genes that contribute to creating reproductive barriers (Teeter *et al.*, 2010; Harrison and Larson, 2014) and species separation.

The biological SC is not universally applicable, as it can only be applied for sexually reproducing organisms in synchrony (Zachos, 2018). Despite these limitations, the biological SC has a strong position in studying the origin of species, seen as the genesis of reproductive isolative mechanisms which helps in understanding the formation of biological diversity (Coyne, 1994). Overall, mate recognition, interbreeding and genetic isolation between lineages play a large role in multiple SCs, not only the biological SC (Harrison, 1998). Furthermore, the biological SC makes it possible to explain the existence of morphologically identical species or sexual dimorphism (Coyne, 1994), both of which exist in the *B. tabaci* complex.

Phylogenetic species concept

In more recent years there has been a shift from the biological SC to various versions of phylogenetic SC, which have been widely promoted since 1989 (Isaac *et al.*, 2004). The phylogenetic SCs are based on the idea that species are monophyletic (sharing a common ancestor) (Hennig, 1966; Mallet, 2007) and possess fixed heritable and diagnosable traits (Groves and Grubb, 2011). These species are also genetically differentiated; however, it does not necessarily mean that they are also reproductively isolated (Groves and Grubb, 2011).

Advantages of this SC are that it deals with allopatric populations that could not be studied within the framework of the biological SC, and it is data-driven without the need of extrapolation and speculation (Groves and Grubb, 2011). Because the phylogenetic SC characterises species as the least inclusive phylogenetic unit (Groves and Grubb, 2011), the numbers of described species have risen since taxonomists started using this approach. For example, a meta-analysis comparing the number of species delimited across wide range of animal, plant and fungal taxa showed that the phylogenetic SC recognised 48.7% more species than non-phylogenetic SCs (Agapow *et al.*, 2004). The rapid growth in species diversity, also called “taxonomic inflation”, has been criticised (Agapow *et al.*, 2004; Isaac *et al.*, 2004; Zachos, 2015). This trend has serious consequences for biodiversity conservation, as it can change the number of endangered species and hampers the decision making about what should be protected (Agapow *et al.*, 2004). In addition, another line of criticism highlighted that the applied degrees of difference between phylogenetic species are subjective, which makes the biological significance of such species questionable (Mayr, 2002; Xu *et al.*, 2010).

Phylogenetic analyses of molecular data

Despite the limitations of the phylogenetic SC, molecular data remain an invaluable source of information for distinguishing between different populations (which make up

species in every SC; Groves and Grubb, 2011) and quantifying levels of their divergence. Amino acid sequences of protein molecules have been used since the 1960's, and nucleotide sequences of DNA since the 1980's. DNA sequences have since become predominant, because they are easier to prepare, allow for investigation of both coding and non-coding regions, and contain more phylogenetic information due to the presence of synonymous mutations that do not result in amino acid changes (Brown, 2002). Protein sequences are still used for phylogenetics, e.g. for resolving the phylogenies of very distant organisms, but for more closely related species, DNA sequences are more suitable (Baldauf, 2003).

Mitochondrial DNA has been a favoured molecule for these purposes, given its desirable properties as a molecular marker. These properties include its nearly ubiquitous distribution, high copy number, simple genetic structure, effective haploidy, maternal mode of inheritance and the mosaic structure composed of both quickly and slowly evolving regions (Avise *et al.*, 1987; Zhang and Hewitt, 1996). Within the mitochondrial genome, protein-coding genes were considered more useful than the rDNA genes, because the latter commonly contain insertions and deletions (Hebert *et al.*, 2003a). Out of the 13 mitochondrial protein-coding genes, mtCOI offers the advantage of a higher mutation rate and so a higher potential for providing phylogenetic signals (Hebert *et al.*, 2003a). In *B. tabaci*, the nucleotide diversity of mtCOI sequence was five-fold larger than that of nuclear internal transcribed spacer 1 (ITS1), or even ten-fold in the case of the MED putative species (Gueguen *et al.*, 2010). Studies of such variable regions are useful for resolving relationships among recently diverged species (Campbell *et al.*, 1994). In contrast, however, a comparison of the evolutionary rates of 13 mitochondrial protein-coding genes across Hemiptera revealed that mtCOI had the lowest rate of evolution (Wang *et al.*, 2015). Nevertheless, the mtCOI gene has become widely used as it was deemed to provide satisfactory power in species resolution across many invertebrate taxa in the barcoding approach (Hebert *et al.*, 2003a, 2003b). The development of universal primers and the popularity of this region lead to the accumulation of mtCOI sequences in public databases, making it available for phylogenetic analyses.

A phylogeny is "*a branching tree diagram showing the course of evolution in a group of organisms*" (Felsenstein, 1983) and these diagrams represent our attempts to reconstruct the evolutionary history of life (Huelsenbeck and Ronquist, 2001). The trees are inferred from multiple alignments of sequences of the same region from different samples. Using mathematical algorithms, the samples are grouped and organised into trees representing

the information gathered from the alignment. Two general methods for calculating the trees are (i) distance-matrix or clustering method, e.g. neighbor-joining, and (ii) discrete data or tree searching methods, such as parsimony, maximum likelihood or Bayesian inference (Page and Holmes, 1998; Baldauf, 2003). The distance-matrix method is relatively simple and fast, but does not contain more information other than the nucleotide distance between samples, while the discrete data methods are slower, but also more robust and information-rich, because each column of the alignment is examined separately (Baldauf, 2003). Parsimony is a principle that the hypothesis about evolutionary relationships that requires the smallest number of character (nucleotide) changes is most likely to be correct, and so the maximum parsimony tree representing the smallest number of changes in all sequence positions is identified (Mount, 2008). Maximum likelihood method uses all the sequence data to formulate a probabilistic model of evolution and chooses the tree with the highest likelihood as the estimate of phylogeny (Felsenstein, 1981). The Bayesian approach combines the prior probability of a phylogeny with the tree likelihood and chooses the maximum posterior probability tree (Rannala and Yang, 1996). The maximum parsimony approach is best suited for a small amount of similar sequences (Mount, 2008). Maximum likelihood and Bayesian methods have been reported to produce more accurate tree reconstructions than neighbor-joining (Ogden and Rosenberg, 2006), and to produce roughly similar results to each other (Huelsenbeck and Ronquist, 2001).

To accommodate the variation in substitution patterns among different sites, the sequence data analysed in phylogenetic analysis can be partitioned into groups (Buckley *et al.*, 2001). Common partitions include separating third codon positions from first and second codon positions, as they evolve at a different rate (Bofkin and Goldman, 2007). Another option is to partition the sequences into individual genes or elements within genes (e.g. exons and introns in protein-coding genes, or stems and loops in rDNA genes) (Leavitt *et al.*, 2013; Kainer and Lanfear, 2015). No partitioning of the data, as well as over-partitioning, can have a negative effect on the accuracy of the resulting phylogeny (Fenn *et al.*, 2008; Kainer and Lanfear, 2015).

2.1.2 A historical review of *B. tabaci* systematics and taxonomy

Description and synonymisation (1889–1957)

The first formal description of *B. tabaci* dates back to 1889, when Panayiotis Gennadius collected a specimen of a tobacco pest in Greece and named it *Aleyrodes tabaci* (Gennadius, 1889). A similar species was later found on sweet potato in the New World and described as *Aleyrodes inconspicua* (Quaintance, 1900). Both *A. inconspicua* and *A. tabaci* were later classified in the genus *Bemisia* (Quaintance and Baker, 1914; Takahashi, 1936). Based on the morphological characters of the fourth nymph instar, numerous *Bemisia* species names were created around the world, 22 of which were synonymised with *B. tabaci* (Gennadius) as one species (Russell, 1957; Mound and Halsey, 1978) (Table 2-1). The reason for this synonymisation was the recognition of the high plasticity in nymph morphology, which was dependent on the host plant, rather than being species-specific (Russell, 1957; Mound, 1963). In addition, the geographic isolation and low exchange of information between taxonomists of the time probably also contributed to the confusion (De Barro *et al.*, 2011).

The synonymised *B. tabaci* species (*sensu* Russell, 1957), however, were not uniform in their biological attributes, so the differences observed between different *B. tabaci* populations led to the naming of various host races and biotypes.

Table 2-1: List of species synonymised as *B. tabaci* (Russell, 1957). From Mound and Halsey (1978).

Name	Described by	Country	Year
Aleurodes tabaci	Gennadius	Greece	1889
Aleurodes inconspicua	Quaintance	FL, US	1900
Bemisia emiliae	Corbett	Sri Lanka	1926
Bemisia inconspicua	(Quaintance) Quaintance & Baker	FL, US	1914
Bemisia costa-limai	Bondar	Brazil	1928
Bemisia signata	Bondar	Brazil	1928
Bemisia bahiana	Bondar	Brazil	1928
Bemisia gossypiperda	Misra & Lamba	India, Pakistan	1929
Bemisia achyranthes	Singh	India	1931
Bemisia hibisci	Takahashi	Taiwan	1933
Bemisia longispina	Priesner & Hosny	Egypt	1934
Bemisia gossypiperda var. *mosaicivectura*	Ghesquière	Congo	1934
Bemisia goldingi	Corbett	Nigeria	1935
Bemisia nigeriensis	Corbett	Nigeria	1935
Bemisia rhodesiaensis	Corbett	Zimbabwe	1936
Bemisia tabaci	(Gennadius) Takahashi	Mariana Islands	1936
Bemisia manihotis	Frappa	Madagascar	1938
Bemisia vayssierei	Frappa	Madagascar	1939
Bemisia (Neobemisia) hibisci	Takahashi; Visnya	Taiwan	1941
Bemisia (Neobemisia) rhodesiaensis	Corbett; Visnya	Zimbabwe	1941
Bemisia lonicerae	Takahashi	Japan	1957
Bemisia minima	Danzig	Georgia	1964
Bemisia miniscula	Danzig	Georgia	1964

Host races, strains and biotypes of B. tabaci (1957–1994)

A host race is defined as *"a population of a species that is partially reproductively isolated from other conspecific populations as a direct consequence of adaptation to a specific host"* (Diehl and Bush, 1984). Different populations of *B. tabaci* have been reported to show specific host plant ranges. For example, in Puerto Rico two different host races of *B. tabaci* were distinguished. One was strongly associated with *Jatropha gossypifolia*, transmitting *Jatropha mosaic virus* exclusively to and from *J. gossypifolia* (Bird, 1957) and the other was a polyphagous race named "Sida" that transmitted various geminiviruses to and from numerous hosts (Bird and Maramorosch, 1978). A similar phenomenon was observed in Ivory Coast, where one *B. tabaci* race specialised on cassava and aubergine

(the “cassava biotype”), while the other (the “okra biotype”) fed on multiple plant species excluding cassava (Burban *et al.*, 1992). The terms “host race” and “biotype” have sometimes been used interchangeably.

The naming system for *B. tabaci* populations was greatly affected by the concept of biotypes and at least 36 of them were described from around the world (Liu *et al.*, 2012b). However, the biotype concept in general has been criticised by multiple authors. While the definitions of “host race” differ among authors, it assumes at least some mechanism of cohesion (*i.e.* adaptation to a specific host) (Downie, 2010). In contrast, “biotype” was adopted by applied biologists to refer to populations distinguished by a very wide range of criteria (Diehl and Bush, 1984). Some described “biotype” as a taxonomic concept mostly used by non-taxonomists (Eastop, 1973), or even a pseudo-taxonomic category that led to simplistic and misleading perceptions of the variation within and among populations (Downie, 2010). In entomology, “biotype” was defined as “*an individual or population that is distinguished from the rest of its species by criteria other than morphology, for example a difference in parasite ability*” (Maxwell and Jennings, 1980). However, this definition confuses the variability between individuals and variability between populations and is too broad to have any biological significance (Claridge and Hollander, 1983).

Biotypes of *B. tabaci* (*sensu* Russell, 1957) were described primarily by using non-specific esterase bands visualised by polyacrylamide gel electrophoresis (Costa and Brown, 1991; Bedford *et al.*, 1992; Byrne and Devonshire, 1993). The esterase banding patterns were assigned a letter, which was then used to describe the whitefly population with that pattern (e.g. biotype A). The esterase patterns were, however, not always consistent in individuals within one biotype and they varied with the level of resistance to agrochemicals, which is why authors advised not to use exclusively this method as a taxonomic tool (Byrne and Devonshire, 1993; Bedford *et al.*, 1994; Byrne *et al.*, 1995a). Other allozymes that sometimes showed diagnostic differences among populations or biotypes of *B. tabaci* were acetylcholinesterase (in its kinetic activity) and phosphoglucomutase, phosphate-glucose isomerase and aconitase (in isoelectric focusing electrophoresis) (Perring *et al.*, 1992, 1993; Byrne and Devonshire, 1993; Byrne *et al.*, 1995b; Brown, 2000). Among DNA-based methods, populations were distinguished by random amplification of polymorphic DNA (RAPD), which also produces band patterns visualised by polyacrylamide gel electrophoresis (Gawel and Bartlett, 1993; De Barro and Driver, 1997; Guirao *et al.*, 1997; De Barro *et al.*, 1998; Maruthi et al., 2001, 2002a), and by sequence comparison of specific genes.

The first two *B. tabaci* biotypes were described after an outbreak of whiteflies that infested greenhouse ornamentals in the southern United States and the Caribbean Basin in the 1990s. The indigenous population was named A-biotype, while the new population infesting greenhouses was designated as B-biotype (Costa and Brown, 1991). In some cases, these biotypes were referred to as "strains" (Byrne and Miller, 1990; Bethke *et al.*, 1991; Campbell *et al.*, 1993) e.g. the B-biotype was also called a Poinsettia strain (Perring *et al.*, 1992). The A- and B- biotypes were characterised by their esterase banding pattern and the ability of the B-biotype to induce squash silver-leaf disorder (Yokomi *et al.*, 1990; Costa and Brown, 1991). Further differences between them were found using allozyme isoelectric focusing electrophoresis (Perring *et al.*, 1992) and RAPD profiles (Gawel and Bartlett, 1993; De Barro and Driver, 1997). In addition, several morphological characters of the fourth nymphal instar were identified; the absence of a dorsal seta, the width of the thoracic tracheal folds and the width of the wax extrusions from the tracheal folds (Bellows *et al.*, 1994).

Description of a new Bemisia species (1994)

The genetic, ecological, and morphological differences between A- and B-biotype, together with evidence of their reproductive isolation (Perring *et al.*, 1992; Costa *et al.*, 1993a) led to the promotion of the B-biotype to a species (Perring *et al.*, 1993; Bellows *et al.*, 1994). It was given the binomial name, *B. argentifolii* (Perring and Bellows), and the common name, the silverleaf whitefly. This promotion triggered a debate within the scientific community. The criticism was related to the classification based on whitefly morphology, because it was previously shown to be plastic and able to vary with different hosts (Mound, 1963), leaf surface topology (Mohanty and Basu, 1986; Bedford *et al.*, 1994), temperature and humidity (Mohanty and Basu, 1986).

Campbell *et al.* (1993) argued that the decision to promote the B-biotype to a species level was premature. Furthermore, De Barro *et al.* (2005) in their phylogenetic analysis using sequences of ITS1 and mtCOI identified six major "races" of *B. tabaci* and suggested that there were not enough molecular, biological, and behavioural data available to raise any of the races to a species level. On the other hand, Rosell *et al.* (1997) and Frohlich *et al.* (1999) reasoned that insects, e.g. mosquitoes (Narang *et al.*, 1991), are known for evolving morphologically indistinct species and they proposed that *B. tabaci* represents a complex of cryptic sibling species.

In the light of current knowledge, raising of the B-biotype to a species level was valid as it corresponds to the MEAM1 putative species (Dinsdale *et al.*, 2010). However, during this reassignment the A-biotype was erroneously designated as the true *B. tabaci*

(Bellows *et al.*, 1994). It was later shown that the *B. tabaci* syntype corresponded to MED (Tay *et al.*, 2012), while the A-biotype belonged to the putative species called New World (NW) (Dinsdale *et al.*, 2010).

Molecular identification of putative species within the B. tabaci complex

The absence of reliable morphological differences among populations of *B. tabaci* makes the use of molecular methods necessary. One of the first DNA markers examined in *B. tabaci* was the nuclear 18S rDNA (Campbell, 1993). The 1039 bp region examined was, however, 99.8% identical between NW and MEAM1 (Campbell, 1993) which meant that the marker lacked the variability required to differentiate species within the *B. tabaci* complex.

Other markers that have been more successful at distinguishing between species were segments of mitochondrial genes, such as 16S rDNA and mtCOI. The maximum parsimony analysis of 16S rDNA and mtCOI haplotypes of whiteflies collected from the New World, India, the Middle East and North Africa confirmed the different origin of NW and MEAM1 species by providing the first molecular evidence of MEAM1 coming from the Old World (Frohlich *et al.*, 1999). Based on its position in the phylogenetic tree and several biological traits, such as the ability to induce silver-leafing, host range and long-distance flying *en masse*, MEAM1 was considered to have roots in Sahel-like regions of the Middle East (Frohlich *et al.*, 1999).

A similar study was performed with 31 populations of *B. tabaci* from Asia, Oceania, New World, Mediterranean basin, Middle East and sub-Saharan Africa using ITS1 sequences (De Barro *et al.*, 2000). The resulting phylogeny, combining parsimony and maximum likelihood methods, revealed strong geographic patterns among clades, which was congruent with the findings of Frohlich *et al.* (1999). De Barro *et al.* (2000) provided evidence for the monophyly of the American clade (A-, C- and R-biotypes from US, Costa Rica and Columbia, respectively, all currently considered to be the NW putative species) and of MEAM1, samples of which were collected from various continents. The analysis also provided strong support for the clade consisting of Q-biotype from Spain, L-biotype from Sudan and J-biotype from Nigeria, all currently considered MED (Dinsdale *et al.*, 2010).

The phylogenies based on the mitochondrial (mtCOI, 16S rDNA) and nuclear (ITS1, microsatellites) markers were compared and shown to produce similar groupings to each other (Frohlich *et al.*, 1999; De Barro *et al.*, 2005; Delatte *et al.*, 2006; Gueguen *et al.*, 2010). However, the use of mtCOI partial sequences for species identification became

predominant due to the growing popularity of DNA barcoding (Hebert *et al.*, 2003a). The 5' region of the mtCOI gene has been the most commonly used region for reconstructing phylogenies (Hebert *et al.*, 2003a). The whitefly research community has, however, adopted the use of 3' segment instead due to this region being the one first used most widely by different researchers globally (Frohlich *et al.*, 1999; Brown, 2000; Dinsdale *et al.*, 2010; Boykin *et al.*, 2012; Firdaus *et al.*, 2013; Mugerwa *et al.*, 2018). For *B. tabaci*, the mtCOI proved to be more informative than 16S rDNA and suitable for examining the relationships between species of the *B. tabaci* complex (Frohlich *et al.*, 1999; Brown, 2000). Moreover, advances in phylogenetic methods and the growing number of accessible partial mtCOI sequences from *B. tabaci* samples worldwide allowed increasingly large-scale global Bayesian phylogenetic analyses to be performed and increasingly more robust phylogenies to be produced (Dinsdale *et al.*, 2010; Boykin *et al.*, 2013; Lee *et al.*, 2013; Mugerwa *et al.*, 2018).

The first Bayesian analysis of 366 whitefly mtCOI sequences resulted in 12 well-resolved genetic groups with monophyletic origin (Figure 2-1) (Boykin *et al.*, 2007). The grouping of sequences into clades was based broadly on the geographic origin of the samples, except the "Mediterranean/Asia Minor/Africa invasive" and "Mediterranean invasive" clades, members of which showed very broad geographic range (Frohlich *et al.*, 1999; De Barro *et al.*, 2000; Boykin *et al.*, 2007). Sequences from "Africa non-silver-leafing" and "Uganda sweet-potato" whiteflies were placed in the most basal clades in the tree, suggesting that sub-Saharan Africa was most likely the place of origin of the *B. tabaci* complex (Boykin *et al.*, 2007). The Q-related biotypes were split into two clades: the "Mediterranean invasive" and "Sub-Saharan Africa silver-leafing". Despite deliberately avoiding the term "species", the authors suggested that the study represented a transition from the inconsistent biotype classification to a more systematic approach (Boykin, 2014).

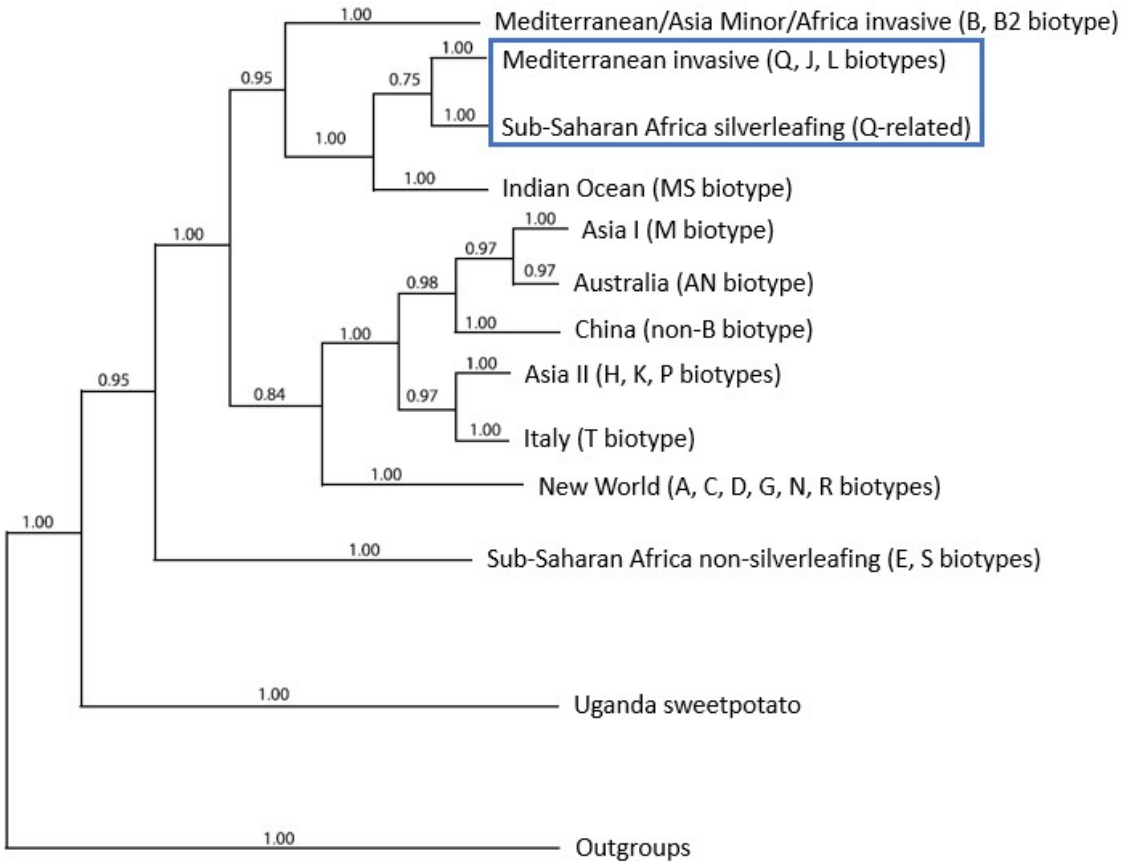


Figure 2-1: Schematic illustration of the phylogenetic tree from Boykin *et al.* (2007). Numbers shown above branches are posterior probabilities from the Bayesian inference analysis. Blue box highlights the division of Q-related biotypes.

The global analysis by Boykin *et al.* (2007) was later reproduced and refined by Dinsdale *et al.* (2010). Authors of this key study applied more rigorous standards on the data involved in the analysis. Sequences with unresolved bases, insertions/deletions, stop codons and duplicate haplotypes were excluded, leaving 202 unique haplotypes used in the analysis. Moreover, the authors attempted to establish the level of nucleotide distance in the 3' mtCOI partial sequence that would enable the identification of species boundaries. A frequency histogram of the pairwise genetic distances among all haplotypes showed a multimodal distribution with two distinct gaps at 3.5% and 11% (Dinsdale *et al.*, 2010). The authors suggested that these gaps represented natural breaks in genetic diversity and indicated the presence of multiple species. At the >3.5% divergence, 24 putative species were resolved within the *B. tabaci* complex (Figure 2-2). These were named Asia I, Asia II-1 to Asia II-8, Australia, Australia/Indonesia, China 1, China 2, Italy, Sub-Saharan Africa 1 to 4, Uganda, New World, Mediterranean, Middle East-Asia Minor 1 and 2, and Indian Ocean. The 24 putative species were assigned to 11 larger genetic groups separated by the nucleotide distance exceeding 11% (Dinsdale *et al.*, 2010).

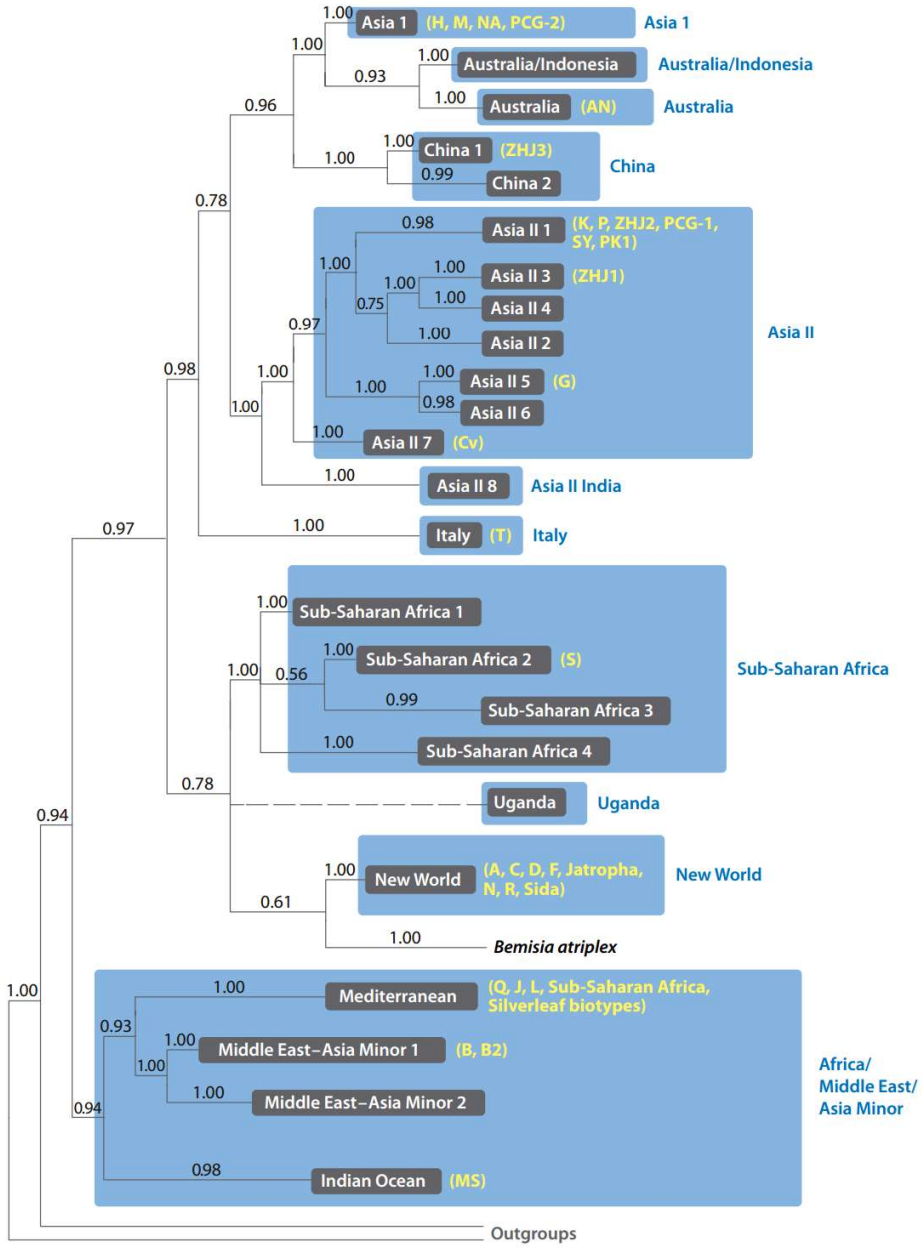


Figure 2-2: Schematic illustration of the phylogenetic tree from Dinsdale *et al.* (2010). Numbers shown above branches are posterior probabilities from the Bayesian inference analysis. Black boxes indicate 24 species-level groups separated by 3.5% mtCOI sequence divergence; blue boxes and blue text mark 11 genetic groups differing by more than 11%. Biotypes assigned to particular species are designated in yellow brackets besides the species name. The figure is reproduced from De Barro *et al.* (2011).

The 3.5% divergence threshold has been adhered to in subsequent analyses, which has led to the identification of 25 new putative species in addition to the 24 proposed by Dinsdale *et al.* (2010) (Table 2-2). However, two of the 24 putative species were later shown to cluster with outgroup species (e.g. *B. atriplex* and *B. afer*) outside the *B. tabaci* complex, suggesting that they are probably non-*tabaci Bemisia* species. These two

species were Japan 2 or *JpL* (Ueda *et al.*, 2009; Firdaus *et al.*, 2013; Lee *et al.*, 2013; Tay *et al.*, 2017b) and “Uganda” collected from sweet potato (De Barro, 2012; Boykin *et al.*, 2013; Firdaus *et al.*, 2013). In addition, despite the efforts to eliminate pseudogene sequences from the dataset, Dinsdale *et al.* (2010) reported the MEAM2 putative species, which was later shown to be an artefact based on pseudogenes from MEAM1 nuclear DNA (Tay *et al.*, 2017a; Elfekih *et al.*, 2018). A “MEAM2” population was recently also reported in Uganda (Mugerwa *et al.*, 2018), however, this is a new putative species that was named on the basis of 97.1% similarity (*i.e.* within the 3.5% species delimitation threshold) to a published MEAM2 pseudogene sequence and should not be confused with the MEAM2 pseudogene clade. Excluding Japan 2, Uganda and MEAM2 (*sensu* Dinsdale *et al.* 2010), the current number of putative species in the *B. tabaci* complex is 46 (Table 2-2).

Adopting the 3.5% mtCOI sequence divergence as a species delimitation threshold, however, has resulted in some incongruities. A statistical parsimony network analysis of mtCOI haplotypes, for example, identified six out of the 24 putative species which did not form a single network, but rather resulted in multiple networks and/or unconnected haplotypes due to the genetic distance being too high (De Barro and Ahmed, 2011). These six putative species were Asia I, Asia II-7, Asia II-6, MED, MEAM and New World (De Barro and Ahmed, 2011). A similar observation was made by Lee *et al.* (2013), who also found six putative species in which the intraspecies genetic variability exceeded 3.5%. The list of species, however, differed between these two studies, as Lee *et al.* (2013) reported Australia and Sub-Saharan Africa 1 (SSA1) instead of MEAM1 and Asia II-6 in De Barro and Ahmed (2011). Pairwise genetic distances of 31 putative species revealed a gap at 4% lying between the distributions of intra- and interspecific divergences, which led to the suggestion to increase the genetic threshold for species delimitation to 4% (Lee *et al.*, 2013). The 4% threshold, however, has not been adopted in subsequent studies.

Table 2-2: Putative species within the *B. tabaci* complex identified to date based on the 3.5% divergence threshold (Dinsdale *et al.*, 2010). The crossed-out names were shown to be non-*tabaci Bemisia* species (Boykin *et al.*, 2013; Firdaus *et al.*, 2013; Tay *et al.*, 2017b) or a clade identified based on a pseudogene (Tay *et al.*, 2017a; Elfekih *et al.*, 2018).

No.	Putative species	n	Divergence from the closest relative	mtCOI primers used	Reference
1	Asia I	16			
2	Australia/Indonesia	4			
3	Australia	1			
4	China 1	3			
5	China 2	1			
6	Asia II-1	5			
7	Asia II-2	1			
8	Asia II-3	2			
9	Asia II-4	1			
10	Asia II-5	2			
11	Asia II-6	1			
12	Asia II-7	5			
13	Asia II-8	4	"Divergence greater than 3.5%"	C1-J-2195 TL2-N-3014	Dinsdale *et al.* (2010)
14	Italy	7			
15	SSA 1	20			
16	SSA 2	11			
17	SSA 3	1			
18	SSA 4	8			
	~~Uganda~~	1			
19	New World (1)	6			
20	Mediterranean	49			
21	Middle East-Asia Minor 1	44			
	~~Middle East-Asia Minor 2~~	1			
22	Indian Ocean	4			
23	Asia II-9	2			
24	Asia II-10	2	"Exceeds 3.5%"	C1-J-2195 TL2-N-3014	Hu *et al.* (2011)
25	Asia III	3			
26	China 3	1			
27	New World 2	6	3.6–4.3% from NW	C1-J-2195 L2-N-3014	Alemandri *et al.* (2012)
28	Ru	5	10.7% from Italy	C1-J-2195 TL2-N-3014	Parrella *et al.* (2012)
29	Asia I-India	1	6.2% from Asia I	CO10 CO12	Chowda-Reddy *et al.* (2012)
30	SSAF-5	6	5.6–6.5% from SSAF-1	C1-J-2195 TL2-N-3014	Esterhuizen *et al.* (2013)

31	Asia II-11	2			
32	Asia II-12	1			
33	Japan 1 (former China)	3			
	~~Japan 2 (former *JpL*)~~	7	"At least 3.5% divergent"	C1-J-2195 L2N3014	Firdaus *et al.* (2013)
34	Asia IV	4			
35	African (Cameroon)	1			
36	SuBSahAf5	1			
37	Morocco	2	5.5–6.4% from Italy	C1-J-2195 L2-N-3014	Tahiri *et al.* (2013)
38	China 4	1	9.6% from China 3	C1-J-2195 TL2-N-3014	Hu *et al.* (2014)
39	China 5	2	4.9% from China 3	C1-J-2195 TL2-N-3014	Hu *et al.* (2017)
40	Asia V	1	11.6% from China 4		
41	SSA 9	3	6.5–7.2% from "UgCsNm3"		
42	SSA 10	11	14–14.7 from NW		
43	SSA 11	3	15.2% from NW		
44	SSA 12	18	5.9–6.1% from "Bangalore 2"	2195Bt CO12/Bt-sh2	Mugerwa *et al.* (2018)
45	SSA 13	2	5.3–5.4% from "Morocco Ag1" (MED)		
46	Middle East-Asia Minor 2 (Uganda)	7	2.9% from "MEAM2" from La Réunion		

n = number of haplotypes in the referenced publication
SSA = Sub-Saharan Africa.

Caution must be taken, however, when reporting new putative species based exclusively on a short segment of DNA. As shown in Table 2-2, the number of haplotypes used to suggest the existence of a new putative species was often low. In 23 cases out of 49 (47%), the new putative species was suggested on the basis of only one or two haplotypes. There is an associated risk with this approach, because such a low amount of sequencing data could represent erroneous sequences or pseudogenes. Inclusion of such sequences in analyses can hinder the elucidation of *B. tabaci* systematics by inflating the genetic divergence values, overestimating the species richness (Song *et al.*, 2008) and confounding phylogenetic analyses (van der Kuyl *et al.*, 1995; Thalmann *et al.*, 2004), as well as create confusion and inconclusive results in the research community.

Although the notion of *B. tabaci* being a complex of cryptic species is now widely accepted, there are ongoing challenges in resolving the systematics within this species complex. Firstly, the systematics based on a short sequence of a single molecular marker has been criticised (DeSalle *et al.*, 2005; De Barro, 2012; Hsieh *et al.*, 2014). It is instead advised to use multiple genes or genome-wide data for constructing robust phylogenies and combining the sequence-based approaches with other methods, such as reciprocal mating experiments and biological assays, to form a more integrative approach. Secondly, the new putative species need to be formally described and given a Latin binomial. However, it is possible that some of the newly recognised species have already been named and subsequently synonymised with *B. tabaci* (Table 2-1), which would make any new binomials invalid. The solution would be to apply the barcoding approach for specimen classification and sequence the partial mtCOI gene from the syntypes of the original *Bemisia* species deposited in a museum collection. Sequence alignment with mtCOI sequences from a set of currently recognised list of putative species would enable linking our current knowledge to historical samples and using correct scientific names. To date, the links between modern putative species and museum specimens were achieved for MED (*B. tabaci*), New World 2 (*B. inconspicua*), Asia II-7 (*B. emiliae)* and Asia II-8 (*B. gossypiperda*) (Tay *et al.*, 2012, 2013, 2017b).

2.1.3 Systematics of the Mediterranean putative species

The MED putative species was first named as Q biotype on the Iberian Peninsula where it coexisted with MEAM1 (Guirao *et al.*, 1997). Based on the esterase electromorph pattern it was initially designated as Q biotype (Rosell *et al.*, 1997). It was later found that its geographic location was not restricted to the Iberian Peninsula but it is widely spread across the whole Mediterranean basin, including European, African and Middle-Eastern countries (Chu *et al.*, 2008). Following the proposal of 3.5% partial mtCOI sequence divergence as the threshold for putative species, the Q biotype was synonymised as MED together with biotypes J and L found in Nigeria and Sudan, respectively (Brown *et al.*, 1995a; Dinsdale *et al.*, 2010).

Together with MEAM1 and Indian Ocean (IO) putative species, MED belongs to the “Africa/Middle East/Asia Minor” clade (Dinsdale *et al.*, 2010). Phylogenetic studies suggested that MED is a more recently diverged sister species of MEAM1 (Boykin *et al.*, 2007; Lee *et al.*, 2013; Mugerwa *et al.*, 2018), while IO is basal to the clade (Dinsdale *et al.*, 2010; De Barro, 2012; Lee *et al.*, 2013). The relatedness of MED to IO, which is found in East Africa and south-western Indian Ocean islands, and to African silver-leafing clades led to the hypothesis that MED originated in sub-Saharan Africa and later spread

throughout the Mediterranean basin (Boykin *et al.*, 2007). The genetic networking analysis of De Barro and Ahmed (2011) suggested the presence of two home ranges of the MED species; a sub-Saharan and a Mediterranean range. The data indicated that MED evolved in sub-Saharan Africa and then spread through Turkey/Syria/Israel to the European Mediterranean countries. Furthermore, the presence of haplotypes from the western Mediterranean region in African countries suggests that MED migrated back to sub-Saharan Africa more recently as a part of the global invasion (De Barro and Ahmed, 2011).

Even though MED is considered to be a single putative species, it shows a complex inner genetic structure at both mitochondrial and nuclear levels (McKenzie *et al.*, 2009; De Barro and Ahmed, 2011; Terraz *et al.*, 2014). Three independent studies have found initially that MED formed two distinct subclades (Tsagkarakou *et al.*, 2007; Chu *et al.*, 2008; Ahmed *et al.*, 2009). The first subclade, named Q1 or MedBasin1, consisted of populations predominantly from the western Mediterranean basin (Spain, Portugal, France, Greece, Morocco and Algeria), while the other designated as Q2 or MedBasin2 was apparently indigenous to the eastern Mediterranean countries, such as Israel and Cyprus (Chu *et al.*, 2008; Ahmed *et al.*, 2009). Seven different measures of genetic variation, such as the number of polymorphic sites, the number of haplotypes or the nucleotide diversity, were used to compare these two subclades. All of them showed that Q1/MedBasin1 had a higher genetic variability than Q2/MedBasin2 (Chu *et al.*, 2008; Ahmed *et al.*, 2009).

The partition of MED putative species into two mtCOI subclades was later expanded to four (Gueguen *et al.*, 2010). In addition to Q1 and Q2, two sub-Saharan groups were reported: Q3 from Burkina Faso and ASL group, corresponding to the one previously reported by Boykin *et al.* (2007) (Figure 2-1), collected in Burkina Faso, Ivory Coast and Cameroon (Figure 2-3 A) (Gueguen *et al.*, 2010). Each of these four subclades was associated with a specific combination of secondary endosymbionts (Gueguen *et al.*, 2010). Furthermore, another phylogenetic study distinguished five different subclades, named Q1–Q5 (Figure 2-3 B) (Chu *et al.*, 2012a). In this study, the subclades Q1 and Q2 were congruent with previous analyses (Chu *et al.*, 2008; Gueguen *et al.*, 2010); haplotypes of Q1 mostly originated from the western Mediterranean countries, but were also found in more northern European, sub-Saharan African, Asian and American countries, while the members of Q2 subclade were collected in Israel, Syria, Cyprus, Turkey, Egypt, and the US. The subclade Q4 comprised samples from sub-Saharan Africa (Ivory Coast, Burkina Faso, Uganda, Nigeria and Ghana) and the last two

subclades were only identified in single countries (Q3 in Croatia and Q5 in Burkina Faso) (Chu *et al.*, 2012a). However, the authors did not state how Q3, Q4 and Q5 corresponded to the subclades described in previous literature.

It should be noted that MED haplotypes named Q1, Q2 and Q3 were also reported in Florida (McKenzie *et al.*, 2009). This terminology, however, pre-dates the reports of Q1–ASL and Q1–Q5 groupings of MED. This study was conducted in a single American state, so it is likely that the Q1–Q3 names are unrelated to the Mediterranean/African populations and denote unique haplotypes within the invasive MED groups instead. Therefore, multiple independent naming of different "Q" haplotypes leads to inconsistencies of species description, subsequently making it challenging to compare the data from different studies.

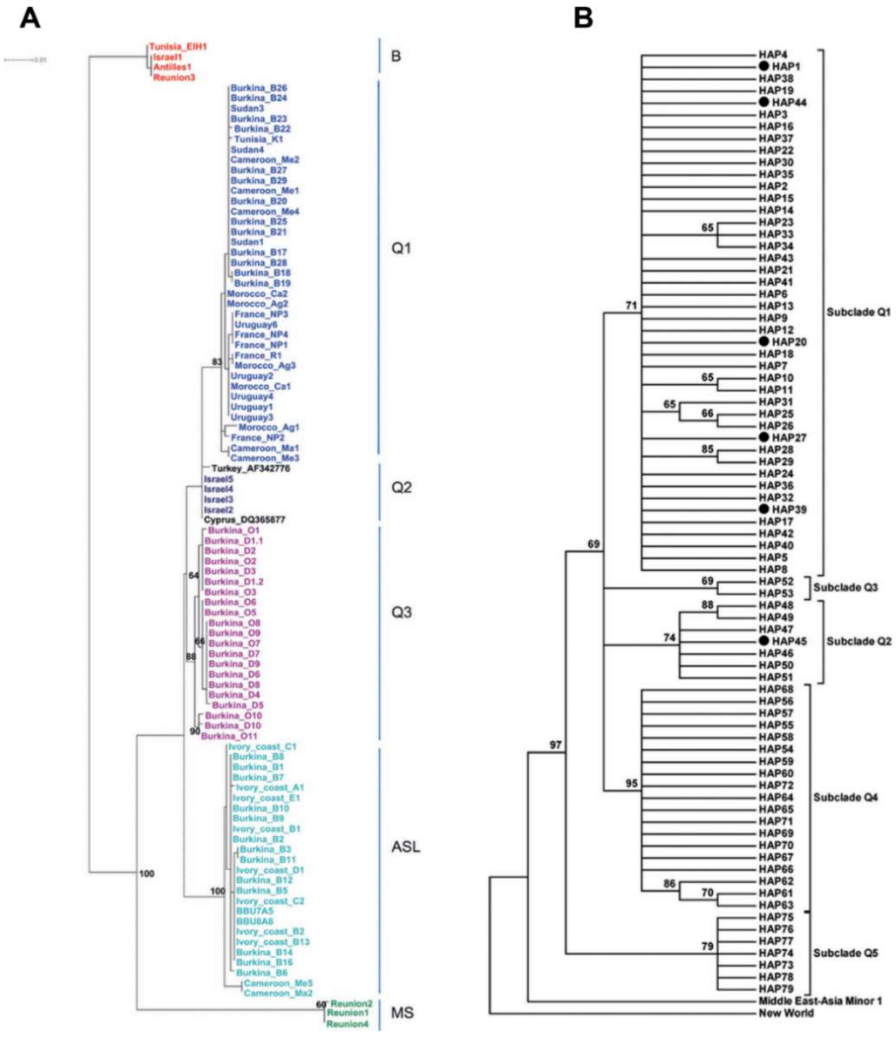


Figure 2-3: Phylogenetic trees based on partial mtCOI sequences showing subclades identified within the Mediterranean putative species of the *B. tabaci* complex. **A:** Maximum-likelihood tree showing Q1, Q2, Q3 and ASL groups of MED, as well as MEAM1 (B) and IO (Ms) species (Gueguen *et al.*, 2010), **B:** Neighbor-joining tree showing Q1–Q5 groups of MED, as well as MEAM1 and NW species (Chu *et al.*, 2012a).

2.2 Reproductive relationships within the *B. tabaci* species complex and the role of bacterial endosymbionts

2.2.1 Integrating the biological species concept with the molecular phylogeny of *B. tabaci*

The biological SC is an important concept in delimiting *B. tabaci* species complex, as there are no reliable diagnostic morphological characters and the molecular data can be equivocal when used in isolation. It has been suggested that phylogenetic analyses studying evolutionary relationships between populations of the *B. tabaci* complex are useful by providing a framework against which the existence of biological species can be tested (Suatoni *et al.*, 2006).

The methods used to study the reproductive compatibilities among *B. tabaci* populations can be divided into two groups. The first one involves crossing experiments with whiteflies, exploiting their haplodiploid sex-determination system in which diploid females can only be produced from fertilised eggs, while males are haploid (Byrne and Devonshire, 1996). Therefore, the occurrence of a successful mating between individuals in the parental generation can be identified by the presence of females in their F_1 progeny. The crossing trials consist of control crosses, where females and males of the same group are allowed to mate, and of reciprocal crosses, in which females from one group are put together with males from the other group and vice versa (Liu *et al.*, 2012b; Xu *et al.*, 2010). These experiments can be set up between one female and one male, one female and multiple males or as a group mating of multiple females and multiple males.

The second group of methods includes using molecular data to detect hybrids between populations or species, usually from samples collected in the field. These methods involve detecting species-specific polymorphisms, either at random by RAPD (Moya *et al.*, 2001; Sun *et al.*, 2011), or locus-specific (e.g. sequence characterised amplified regions, SCAR, or cleaved amplified polymorphic sequences, CAPS) (Khasdan *et al.*, 2005). Another locus-specific method is the analysis of microsatellites, which are generally neutral and highly polymorphic markers linked to loci under strong selection (Delatte *et al.*, 2006; Thierry *et al.*, 2015). More recently, an evidence of gene flow between putative species has also been demonstrated by tracing patterns in a genome-wide analysis of single nucleotide polymorphisms (SNPs) (Elfekih *et al.*, 2018).

Bedford *et al.* (1994) were the first to attempt a systematic comparison between different *B. tabaci* biotypes and their capacity to interbreed. Crossing experiments were performed between MEAM1 (biotypes B and B2), Asia II-1 (biotype K), Asia I (biotype M) and NW

(biotype D) species. Biotypes B and B2, now considered the same putative species MEAM1, were readily interbreeding, while all other crosses showed a lack of gene flow due to the production of only male progeny (Bedford *et al.*, 1994).

Due to the invasive nature of MED and MEAM1, studying the mating compatibility between them and the indigenous *B. tabaci* species has been a priority. Complete reproductive isolation was reported from crossing experiments between MEAM1 and Asia II-1, Asia II-3, Asia II-7 and China 1 (Zang and Liu, 2007; Sun *et al.*, 2011; Wang *et al.*, 2011a). MED whiteflies were reported to be reproductively incompatible with Italy, Asia II-1, Asia II-3, Asia II-7 and China 1 (Demichelis *et al.*, 2005; Wang *et al.*, 2010; Wang *et al.*, 2011a).

In the last 20 years, the chance of an encounter of MED and MEAM1 whiteflies in the field have been higher than for any other *B. tabaci* species, because both MED and MEAM1 were spreading invasively around the world. The consequences of a potential interbreeding between MED and MEAM1 could be particularly disastrous, given the high fecundity of MEAM1 and insecticide resistance in MED. It was, therefore, important to assess the level of gene flow between these two species. Studies analysing sympatric field samples from countries around the Mediterranean basin have not detected any hybrid individuals of MED and MEAM1 (Moya *et al.*, 2001; Khasdan *et al.*, 2005; Simón *et al.*, 2007; Dalmon *et al.*, 2008; Saleh *et al.*, 2012), with the exception of Tahiri *et al.* (2013) who identified two hybrid individuals among 346 samples collected in Morocco. In laboratory crossing trials, MED and MEAM1 populations from Israel were reported to be unable to produce female progeny (Horowitz *et al.*, 2003b; Elbaz *et al.*, 2010) and no copulation was observed between them (Elbaz *et al.*, 2010). In contrast, different authors reported some level of reproductive compatibility between MED and MEAM1. Viable F_1 females were reported from a MED x MEAM1 cross in Spain (Ronda *et al.*, 2000), but it was later demonstrated that these females were sterile as the F_2 generation failed to develop (Ma *et al.*, 2004). Mating between Chinese populations of MED and MEAM1 resulted in 0.2–1.9% of females in the progeny, most of which were confirmed as hybrids by RAPD (Sun *et al.*, 2011). However, the authors noted that the inter-species copulation was only observed after placing an individual in a dense cohort of the opposite sex of different species for more than three days. In addition, all the females in the F_1 generation were sterile (Sun *et al.*, 2011). It is possible that in more natural conditions with a free mate choice the probability of interspecies mating would be further reduced (Liu *et al.*, 2012b), which could explain the absence of hybrids in the field reported by majority of studies. Taken together, these findings show that MED and MEAM have incompletely

isolated mate recognition systems, but strong pre-zygotic and post-zygotic reproductive barriers are present.

While many of the putative species delimited by 3.5% mtCOI sequence divergence also showed reproductive incompatibility (Xu *et al.*, 2010; Liu *et al.*, 2012b), exceptions were also reported. Only partial reproductive isolation was demonstrated in one direction of the cross between Asia II-3 and Asia II-9, as shown by the percentage of females produced (17.1%, contrasted with the 47.6% produced in the opposite direction that was fully compatible) (Qin *et al.*, 2016). This incomplete isolation was surprising, as the partial mtCOI sequence divergence between these two species was 4.47% and this value was similar to the divergence between MED and MEAM1 (4.72%) that were reproductively incompatible in the same study (Qin *et al.*, 2016). Inter-species hybridisation was also documented between MEAM1 and IO in La Réunion, which differ by 7–8% in the mtCOI partial sequence (Delatte *et al.*, 2006). In this case it was asymmetrical and locus-specific introgression detected in the field via analysis of microsatellite markers (Delatte *et al.*, 2006; Thierry *et al.*, 2011) and supported by genome-wide SNP analysis (Elfekih *et al.*, 2018). The latter study also found evidence of introgression between IO and MED in La Réunion (Elfekih *et al.*, 2018), which was not picked up by microsatellite analysis in an earlier study from that region (Thierry *et al.*, 2015). Lastly, an incomplete reproductive isolation was also reported between MEAM1 and Australia, as the reciprocal crosses yielded 0.7–17% female progeny whose hybrid identity was confirmed by RAPD (De Barro and Hart, 2000). The 3' partial mtCOI sequence divergence between these two species was as high as 14.3% (Vyskočilová *et al.*, 2018).

2.2.2. Inter-species competition and interference

When reproductively isolated species with overlapping niches live in sympatry, a competition is likely to occur between them (Reitz and Trumble, 2002). Interspecific competition occurred in 76% of studied interactions between species of phytophagous insects, particularly if they were closely related, introduced (*i.e.* non-native), sessile and aggregative (Denno *et al.*, 1995). The mechanisms of such competition can be indirect through their potential to reproduce and efficiency in acquiring resources, or it can involve direct contact in the form of physical contest, intraguild predation and reproductive interference (Reitz and Trumble, 2002).

In *B. tabaci*, as well as many other arthropod species, the competition was most commonly observed between an exotic species and a native species in the invaded region (Reitz and Trumble, 2002). The event of a species prevailing over the formerly

established species to the point of its removal is called competitive displacement (Reitz and Trumble, 2002). Within the *B. tabaci* complex, cases of competitive displacement were observed in numerous locations in the field, where the invasive MEAM1 prevailed over an indigenous species (Brown *et al.*, 1995b; Lima *et al.*, 2000; Liu *et al.*, 2007). The competitive advantage of MEAM1 over indigenous *B. tabaci* species has also been studied in laboratory conditions, in which it displaced Asia II-2 and Australia in four to five generations (Crowder *et al.*, 2010a) and Asia II-1 in three to ten generations (Luan *et al.*, 2012). In a field cage experiment, MEAM1 dominated over Australia within three generations (De Barro and Hart, 2000). At a regional scale, long-term field observations in China and Australia reported displacement of indigenous species by MEAM1 within three to five years (Liu *et al.*, 2007).

The ability of MEAM1 to sometimes displace other species within the *B. tabaci* complex was attributed to its high fecundity, the ability to adjust the sex ratio towards more females in the progeny and to mating interference. Males of MEAM1 were actively courting females of different species (Asia II-3, Australia, MED, NW) and thus interfering with the courtship of the competing species (Perring *et al.*, 1994; Pascual, 2006; Liu *et al.*, 2007). MEAM1 males spent more time courting MEAM1 and even MED females and initiated more courtships with both types of females than MED males (Pascual, 2006; Crowder *et al.*, 2010b). The frequency of courting by MEAM1 males increased when a higher number of males was available to court a female, regardless of whether the extra males were MEAM1 or a different species, while this trait was not observed for the indigenous males (Liu *et al.*, 2007; Zang and Liu, 2007). Females of MEAM1 accepted the courtship from their conspecific males more readily if MED males were also present, while MED females were equally likely to accept a MED male in mixed and non-mixed cultures (Crowder *et al.*, 2010b). The percentage of females in the progeny was consistently higher for MEAM1 than MED and rose even higher in mixed cultures, while the female progeny of MED decreased in mixed culture with MEAM1 and with fewer MED males available (Pascual and Callejas, 2004; Crowder *et al.*, 2010b). This means that when MEAM1 population mixes with a population of a different species, it interferes with the mating behaviour of the competing species while intensifying its own mating efforts, which leads to more fertilisation and more female offspring.

After an invasive species establishes in the region, it can be challenged by another invasive species. That has been the case for the MEAM1 and MED species. In laboratory experiments, MEAM1 had a competitive advantage over MED under most conditions. When reared with MED in a mixed colony, the percentage of MEAM1 rose from 50% to

79% in a single generation (Pascual and Callejas, 2004). Other laboratory studies reported that MEAM1 completely displaced MED in five to seven generations (Sun *et al.*, 2011), or even two to three generations in the case of Q1 and Q2, respectively (Rogan, 2012). The competitive advantage of MEAM1, however, applied only when feeding on a suitable host that was free of insecticides. Multiple studies have shown that MED became fitter than MEAM1 under insecticide treatments (Pascual, 2006; Crowder *et al.*, 2010a; Pan *et al.*, 2011; Sun *et al.*, 2013) or on a plant-host more suitable for MED, such as pepper (Tsueda and Tsuchida, 2011; Chu *et al.*, 2012b; Sun *et al.*, 2013). It took two generations for MED to displace MEAM1 on pepper, and seven and two generations on cotton treated with 12.5 and 50.0 mg/l of imidacloprid, respectively (Sun *et al.*, 2013). The displacement of MEAM1 by MED in the field was observed in China, where the displacement took one to four years (Chu *et al.*, 2010a, 2010b).

Some evidence also exists for intra-species competition, specifically within MED. The Q2 group of MED prevailed over Q1 in a mixed colony, as it reached 95% in three generations (Rogan, 2012). This phenomenon was also observed in the field, as Q2 populations became prevalent (70.5%) over Q1 in Italy between 2008 and 2013 (Parrella *et al.*, 2014). Such events have been attributed to the potential effects of endosymbiotic bacteria, which can alter the fitness or the reproductive system of their whitefly hosts (Himler *et al.*, 2011; Parrella *et al.*, 2014).

2.2.3 Bacterial endosymbionts of whiteflies and their role

Insects, often rated as the most diverse animal group, show various types of endosymbiotic associations ranging from obligate mutualism to facultative parasitism (Kikuchi, 2009). Studies of such associations suggested that they are common in insects that depend on diets containing an excess of one class of nutrients, but are deficient in other essential nutrients (Buchner, 1965; Dadd, 1985). Phloem-sucking insects, such as whiteflies or aphids, often lack essential amino acids in their diets, which is thought to be compensated for by the endosymbionts synthesising nutrients for their host (Douglas, 1998; Baumann *et al.*, 2006). The endosymbiotic bacteria often occur within specialised insect cells called bacteriocytes that are aggregated into a bacteriome (Baumann, 2005). These bacteria are then transmitted vertically through the maternal line in the process of transovarial transmission, in which whole bacteriocytes migrate to the female's ovaries and enter the eggs (Costa *et al.*, 1996).

There are two major types of endosymbionts: the primary (obligate) and the secondary (facultative). Primary endosymbionts are indispensable to their hosts and therefore fixed

in the populations (Zchori-Fein and Bourtzis, 2011). They are restricted to the bacteriocytes and have a close co-evolutionary relationship with their host (Thao and Baumann, 2004; Baumann, 2005; Wilson and Duncan, 2015). All members of the Aleyrodinae subfamily, including *B. tabaci*, possess the primary endosymbiont called *Candidatus* Portiera aleyrodidarum ("*Portiera*" hereafter) (Thao and Baumann, 2004). *Portiera* provides some essential amino acids and carotenoids for its host (Nováková and Moran, 2012; Santos-Garcia *et al.*, 2015). It is a pleomorphic bacterium residing within host-derived vesicles in bacteriocytes (Costa *et al.*, 1993b). *Portiera* was originally thought to lack the cell wall (Costa *et al.*, 1993b), but this notion has later been rebutted (Santos-Garcia *et al.*, 2014). Comparison of phylogenetic trees between *Portiera* and its whitefly hosts suggested that this endosymbiont was acquired in a single infection of a whitefly ancestor, followed by a subsequent co-speciation of these two organisms (Thao and Baumann, 2004). Analysis of the full *Portiera* genome showed that, similar to other primary endosymbionts, its genome has been severely reduced (Santos-Garcia *et al.*, 2012). *Portiera* lineages in Aleurodicinae and Aleyrodinae subfamilies evolved a small and stable genome with 319 genes, out of which 280 code for proteins (Santos-Garcia *et al.*, 2015). Genome of *Portiera* from *B. tabaci*, however, showed signs of extensive genome rearrangements and gene losses since its divergence from *Trialeurodes* (Sloan and Moran, 2012, 2013). Some of the lost genes were components of synthetic pathways for histidine, phenylalanine, tryptophan, glycine, and in the case of MED and MEAM1 also arginine, histidine, arginine and lysine (Santos-Garcia *et al.*, 2015). The authors hypothesised that the missing components for amino acid synthesis are complemented by the host.

The secondary endosymbionts have much more diverse representation among whitefly species and populations, because their presence is not necessary for the survival of their host (Baumann, 2005). The localisation of secondary endosymbionts is not restricted to bacteriomes, but they are also inherited maternally (Baumann *et al.*, 2006). However, unlike the primary endosymbiont, their evolutionary history is not fully congruent with their hosts, indicating that they were acquired more recently and via multiple events of horizontal transmission (Thao and Baumann, 2004; Ahmed *et al.*, 2013, 2015). Among the secondary endosymbionts of *B. tabaci* are bacteria from genera *Wolbachia*, *Fritschea*, *Rickettsia*, *Hamiltonella*, *Cardinium*, *Arsenophonus* and *Candidatus* Hemipteriphilus asiaticus (Zchori-Fein and Brown, 2002; Thao *et al.*, 2003; Gottlieb *et al.*, 2006; Bing *et al.*, 2013; Marubayashi *et al.*, 2014).

The role of secondary endosymbionts is more difficult to understand. Their relationship with the host can range from mutualistic association to reproductive manipulation (Zchori-Fein and Bourtzis, 2011). The advantages provided to the host by secondary endosymbionts generally relate to fitness advantage in particular environments, such as heat, plant defence chemicals or natural enemies (Clark *et al.*, 2010). For example, MEAM1 whiteflies infected with *Rickettsia* showed increased thermotolerance to temperatures up to 40°C (Brumin *et al.*, 2011) and a decrease in mortality after infection with the bacterium *Pseudomonas syringae* (Hendry *et al.*, 2014). A generally higher fitness was also observed in *Rickettsia*-infected MEAM1 whiteflies in the US, as demonstrated with the increased fecundity, survival and faster developmental rate (Himler *et al.*, 2011; Asiimwe *et al.*, 2014). In contrast, Israeli MEAM1 whiteflies were not significantly affected by *Rickettsia* in any of the parameters above, except for the developmental time, which was shorter in *Rickettsia*-positive individuals (Chiel *et al.*, 2009). Overall higher fitness was also observed in *Wolbachia*-infected Chinese MED, shown by decreased developmental time, increased nymphal developmental rate and increased adult life span (Xue *et al.*, 2012). Elimination of *Wolbachia* from MED led to a smaller body size of 4th instar nymphs, which suggested *Wolbachia*'s potential role as a nutritional mutualist (Xue *et al.*, 2012), as is the case in bed bugs (Hosokawa *et al.*, 2010). A certain level of protection against the parasitoid wasp *Encarsia bimaculata* in *Wolbachia*-infected MED was also observed (Xue *et al.*, 2012). The infection with *Hamiltonella* was advantageous for Chinese MED, because higher fertility, nymphal survival, faster developmental time and larger adult body size were recorded in comparison to the uninfected whiteflies (Su *et al.*, 2013a).

Some secondary endosymbionts can have negative effects on the fitness of their hosts. For example, the elimination of *Arsenophonus* from Asia II-1 led to increased fecundity, nymphal survival and adult life span relative to the *Arsenophonus*-positive control (Raina *et al.*, 2015). *Cardinium*-negative MED whiteflies in China had a faster developmental rate, higher nymphal survival and reproduction rate than the *Cardinium*-positive (Fang *et al.*, 2014).

The composition of endosymbiotic bacteria has also been correlated negatively with the capacity to detoxify insecticides and plant-defence chemicals (Kontsedalov *et al.*, 2008; Ghanim and Kontsedalov, 2009; Pan *et al.*, 2013). *Rickettsia*-positive MEAM1 whiteflies were 15-times more susceptible to pyriproxyfen and to a lesser extent to acetamiprid, thiamethoxam and spiromesifen, when compared with *Rickettsia*-negative MEAM1 (Kontsedalov *et al.*, 2008). Israeli MED whiteflies with double infections *Arsenophonus*-

Rickettsia and *Arsenophonus*-*Wolbachia* showed higher mortality after treatment with thiamethoxam, imidacloprid, pyriproxyfen and spiromesifen compared with their counterparts with a single *Arsenophonus* infection (Ghanim and Kontsedalov, 2009). The relative density of particular endosymbionts also appeared to play a role in the detoxification. Thiamethoxam-susceptible MEAM1 whiteflies contained a higher relative amount of *Portiera* and *Hamiltonella*, while the thiamethoxam-resistant strain contained more *Rickettsia* (Pan *et al.*, 2013).

Interestingly, the bacterial endosymbionts do not interact exclusively with their whitefly hosts, but they can also affect the efficiency with which the whiteflies transmit plant viruses. The presence of *Hamiltonella*, for example, increased the efficiency of *Tomato yellow leaf curl virus* (TYLCV) transmission in MEAM and Chinese MED (Gottlieb *et al.*, 2010; Su *et al.*, 2013b). The interaction occurred between the coat protein of TYLCV and the chaperone protein GroEL of *Hamiltonella*, but not of *Wolbachia* or *Arsenophonus* (Gottlieb *et al.*, 2010). *Rickettsia* infection of the midgut of MEAM1 whiteflies was also associated with an increased TYLCV transmission efficiency (Kliot *et al.*, 2014). A similar interaction was also observed between the GroEL protein of *Arsenophonus* and the coat protein of a cotton leaf curl virus (species unknown) in Asia II whiteflies (Rana *et al.*, 2012).

Besides providing nutritional and environmental advantage, endosymbiotic bacteria can also interfere with the reproduction of their hosts (Engelstädter and Hurst, 2009). Endosymbiotic bacteria are inherited via female reproductive tract; therefore, males represent a “dead end” that stops the spread of their endosymbionts (Cosmides and Tooby, 1981). The incongruence between the host and the bacterium in the means of maximised bacterial gene propagation creates an intra-genomic conflict (Cosmides and Tooby, 1981). The result of this conflict is the evolution of multiple types of reproductive parasitism, all of which are aimed at enhancing the production of infected females and promoting the spread of the bacterium (Werren *et al.*, 2008). These types of manipulation include feminisation (producing phenotypic females from genetic males), parthenogenetic induction (development of females from unfertilised eggs), male killing (death of males during embryogenesis or larval development) and sperm-egg incompatibility (also called cytoplasmic incompatibility) between infected males and uninfected females, or with male and female infected with different bacterial strains (Stouthamer *et al.*, 1999; Werren *et al.*, 2008; Engelstädter and Hurst, 2009).

The most widely recognised reproductive manipulator is *Wolbachia* (Werren *et al.*, 2008). This is a highly diverse bacterial genus that has been divided into 14 supergroups A–O,

except for G, which is a combination of A and B (Lo *et al.*, 2002, 2007; Baldo and Werren, 2007; Haegeman *et al.*, 2009; Ros *et al.*, 2009; Augustinos *et al.*, 2011). Members of supergroups C and D were found in nematodes, while the remaining ones were found in arthropods, most commonly supergroups A and B (Werren *et al.*, 2008). Another endosymbiont recognised as a reproductive manipulator is *Cardinium*, which was found to cause parthenogenesis, increased fecundity and cytoplasmic incompatibility in predatory mites and *Encarsia* parasitoids (Weeks *et al.*, 2001; Zchori-Fein *et al.*, 2001; Hunter *et al.*, 2003; Weeks and Stouthamer, 2004). *Cardinium* is the only symbiont besides *Wolbachia* known to cause cytoplasmic incompatibility (Mann *et al.*, 2017), but is less prevalent in arthropods than *Wolbachia* (Zchori-Fein and Perlman, 2004).

B. tabaci harbours *Wolbachia* from supergroups A, B and O (Ahmed *et al.*, 2010a; Chu *et al.*, 2011; Bing *et al.*, 2014). The supergroup B was reported to be more prevalent than A (Nirgianaki *et al.*, 2003; Ahmed *et al.*, 2010a; Bing *et al.*, 2014), except for one report from China, in which the A supergroup was dominant (Chu *et al.*, 2011). *Wolbachia* strains from supergroups A and B were also detected in co-infections (Ahmed *et al.*, 2010a; Chu *et al.*, 2011). The supergroup O was a newly identified supergroup found in *B. tabaci* putative species Asia II-1, and to a lesser extent in MED, in China (Bing *et al.*, 2014). Overall, supergroup B infections appeared much more common in *B. tabaci* from Asian, Australian and New World clusters than the African/Mediterranean/Middle East/Asia Minor clade, which led to the suggestion that *Wolbachia* may have played a role in the speciation of the *B. tabaci* complex (Chu *et al.*, 2011).

In interactions between members of the *B. tabaci* complex, *Wolbachia* was suggested to cause partial cytoplasmic incompatibility between MEAM1 and Australia whiteflies (De Barro and Hart, 2000). In addition, *Wolbachia* infection of Chinese MED led to a higher proportion of females in the offspring (Xue *et al.*, 2012). The potential role of *Cardinium*, or double infection *Cardinium*-*Arsenophonus*, in reproductive relationships between *B. tabaci* species was proposed after the observation of non-random hybridisation between MEAM1 and IO (Thierry *et al.*, 2011). In contrast, crossing experiments between *Cardinium*-positive and *Cardinium*-negative MED in China showed no evidence of reproductive manipulation (Fang *et al.*, 2014). Additionally, despite *Rickettsia* not usually being reported as a reproductive manipulator, *Rickettsia*-infected MEAM1 also produced more female-biased progeny compared with *Rickettsia*-free whiteflies (Himler *et al.*, 2011; Asiimwe *et al.*, 2014). No evidence of reproductive manipulation was observed for *Hamiltonella* (Su *et al.*, 2013a).

Even though the secondary endosymbionts play such a wide variety of roles, the mutualism and reproductive parasitism roles are not mutually exclusive (Zchori-Fein and Bourtzis, 2011). For example, both *Wolbachia* in MED and *Rickettsia* inMEAM1 increased the overall fitness of their hosts, while also manipulating the sex-bias in the host's progeny (Himler *et al.*, 2011; Xue *et al.*, 2012). More research is required to gain a better understanding of the ways the endosymbionts affect their hosts' reproduction and biology.

2.3 *B. tabaci* as an agricultural pest

2.3.1 Life cycle, mating and development

Whiteflies *B. tabaci* are small insects that feed on the phloem sap of predominantly herbaceous plant species (De Barro *et al.*, 2011). The adults are 1–2 mm long (Basu, 1995), yellowish in colour with red eyes and white wings. *B. tabaci* usually feeds on the abaxial side of leaves, penetrating the phloem tissue with its mouthpart (stylet) and extracting the phloem sap from minor veins (Cohen *et al.*, 1996; Jiang and Walker, 2007). This diet is rich in sugars and relatively low in amino acids and minerals, so a large amount of sap needs to be ingested to accumulate a sufficient quantity of nutrients. The excess sugar is excreted as honeydew droplets (Byrne and Miller, 1990).

Life cycle

Females of some whitefly species, such as *B. tabac*i and the greenhouse whitefly *Trialeurodes vaporariorum* (Westwood), lay eggs on the abaxial side of leaf surface, secured with a pedicel anchored in a slit punctured in the leaf epidermis by the ovipositor (Figure 2-4). In contrast, many other whitefly species, e.g. the woolly whitefly *Aleurothrixus floccosus* (Maskell) or the orange spiny whitefly *Aleurocanthus spiniferus* (Quaintance), insert the pedicel directly into the plant stomata (Paulson and Beardsley, 1985). In close proximity to the pedicel inside the egg a single bacteriome is positioned, which is a specialised organ containing maternally inherited endosymbiotic bacteria (see section 2.2.3) (Costa *et al.*, 1996).

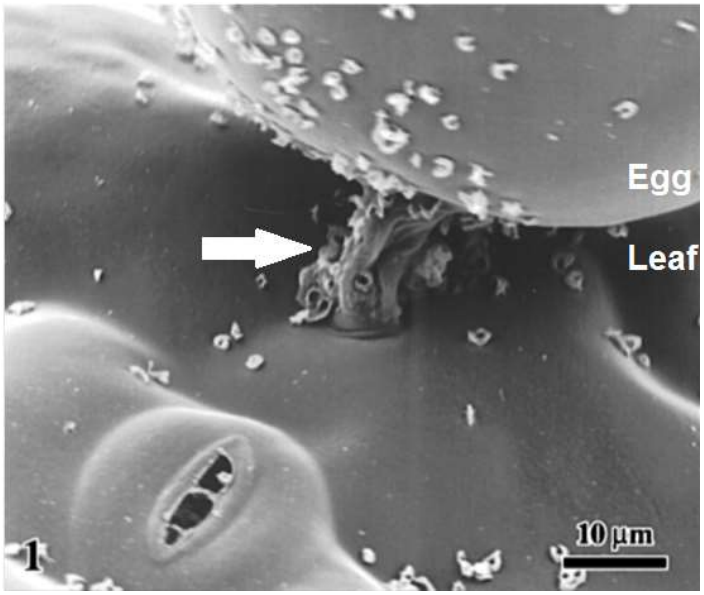


Figure 2-4: Egg of *B. tabaci* (MEAM1) anchored in the epidermis of a cotton leaf by the pedicel (white arrow) (Buckner *et al.*, 2002).

The egg stage is followed by four nymphal instars. The early first instar, or ‘’crawler’’, is the only mobile immature stage in whitefly development (Walker *et al.*, 2010). After hatching from the egg, the crawler moves over the leaf surface in search of a suitable place to insert its stylet. Upon finding such place, the larva starts to feed and usually remains sessile for the rest of its development through second (Figure 2-5 A) to fourth (Figue 2-5 B) instars (Walker *et al.*, 2010). The late stage of the fourth instar, sometimes also called the ‘’pupa’’, is characterised by the enlargement of the eyes which can be observed as two dark red oval spots (Gelman *et al.*, 2002). At this stage, feeding stops until after emergence as an adult (Costa *et al.*, 1999).

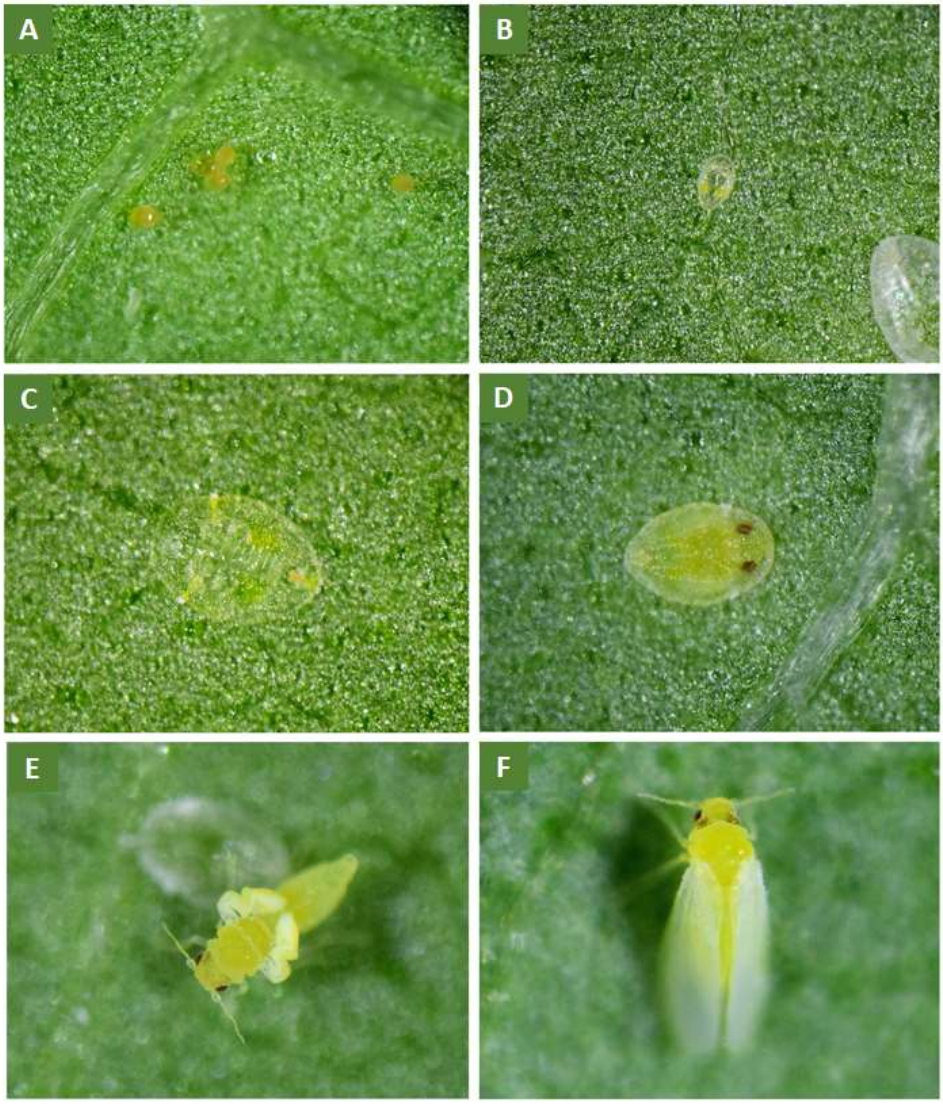


Figure 2-5: Developmental stages of *B. tabaci* at 80x magnification (ASL population from Uganda on an okra leaf). **A:** Eggs, **B:** second instar nymph. **C:** third instar nymph, **D:** fourth instar nymph with enlarged red eyes, **E:** newly emerged adult (<1 min) with curled wings and the empty nymphal case in the background, **F:** young adult with fully expanded transparent wings. Photos by S. Vyskočilová.

Adults emerge from the nymphal case via a T-shaped rupture (Figure 2-6) and their wings are initially curled (Figure 2-5 C). Adults then cover their body and wings with wax particles produced by wax plates on their abdomen, giving the wings their white colour (Byrne and Hadley, 1988) after initially being transparent (Figure 2-5 D). The wax is composed of 65–75% triacylglycerols and 3–7% hydrocarbons and it is probable that its function is to protect the insects from water loss, pathogens and UV radiation (Byrne and Hadley, 1988).

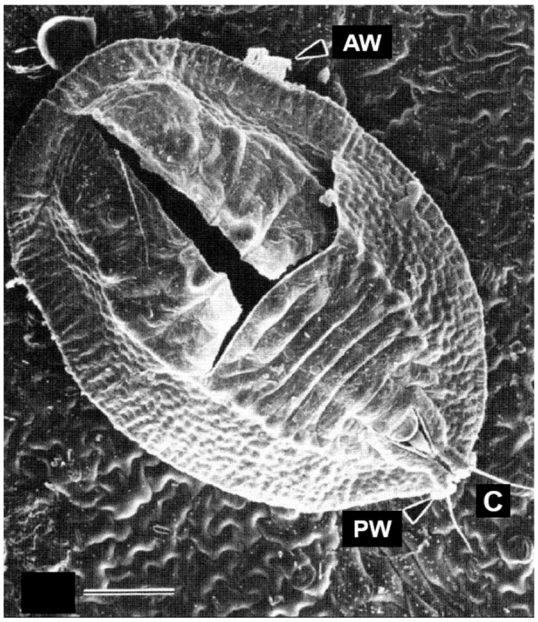


Figure 2-6: Empty nymphal case after adult emergence (MED). AW = anterior wax margin, PW = posterior wax margin, C = caudal setae. Bar = 100 µm (Rosell *et al.*, 1997).

Many studies on the bionomics of *B. tabaci* (e.g. Coudriet *et al.*, 1985; Gerling *et al.*, 1986; Powell and Bellows, 1992) were conducted before it was accepted that *B. tabaci* is a complex of multiple species. Traits such as mating behaviour, developmental time, longevity and fecundity are likely to differ between species. Authors of the studies did not always state the origin of whiteflies used in the study, and even with the knowledge of the geographical location it was usually impossible to deduce the species identity with certainty. In the remaining part of this section, the origin of whiteflies used in the studies is indicated by geographical location or biotype/species identification with a focus on the MED, where possible.

Mating behaviour

B. tabaci whiteflies reproduce by arrhenotokous parthenogenesis, in which fertilised eggs develop into diploid females and unfertilised eggs become haploid males (Byrne and Devonshire, 1996). The mating behaviour of *B. tabaci* whiteflies was studied in detail using video camera recordings. Studies of MEAM1 and Asia II-3 whiteflies

reported that the first mating can occur two to six hours after emergence (Luan *et al.*, 2008). They mate multiple times throughout their lives (Zang and Liu, 2007) and up to four times in the first 72 hours after emergence (Luan *et al.*, 2008). Polygamy (a male mating with more than one female) and polyandry (a female mating with more than one male) do occur (Basu, 1995). Depending on the environment, *B. tabaci* produce 11–15 generations per year (Brown *et al.*, 1995b).

Courtship and mating behaviour were studied using individuals from the NW (Li *et al.*, 1989) and MEAM1 (Perring and Symmes, 2006; Ruan *et al.*, 2007; Zang and Liu, 2007; Luan *et al.*, 2008). The mating process of *B. tabaci* was divided into three major parts: (i) searching for a mating partner and first contact, (ii) courtship behaviour and (iii) copulation.

Both sexes have distinct roles in the mating behaviour. Males were shown to play the active role in the courtship initiation (Zang and Liu, 2007). The male searched for a female by moving randomly across the leaves until he was within two to three millimetres from her (Li *et al.*, 1989). The male then contacted the female, usually on her wing, using his antennae or tarsi and subsequently positioned himself in parallel to her body (Li *et al.*, 1989; Perring and Symmes, 2006). The courting continued by the male starting to “drum” the medial segment of the nearest female antenna with the medial segment of his own antenna (‘’antennal drumming’’ phase, Figure 2-7 A) (Li *et al.*, 1989; Perring and Symmes, 2006; Zang and Liu, 2007).

In the next stage, the male either moved his abdomen up and down in synchrony with the antennal drumming, gradually increasing the speed (Li *et al.*, 1989), or ‘’male abdominal undulation’’ occurred after the termination of antennal drumming (Perring and Symmes, 2006). In 15% of cases, the ‘’body pushing’’ was observed in NW whiteflies, in which the male hastened the tempo of antennal drumming and pushed the female with his side, resulting in a semi-circular movement of the posterior parts of their bodies (Li *et al.*, 1989). However, this behaviour was not observed in MEAM1 (Perring and Symmes, 2006).

In the final stage of courtship, the male raised all four wings (Perring and Symmes, 2006) or the pair of wings closest to the female (Li *et al.*, 1989) and covered her (Figure 2-7 B). Simultaneously, he positioned his abdomen beneath hers, bent it upwards and initiated the contact of copulatory organs (Li *et al.*, 1989). The copulation itself was observed as two to three minutes (Zang and Liu, 2007) or two to four minutes (Ruan *et al.*, 2007) of continuous body overlap.

In contrast, females played a more passive role in the mating behaviour. During the beginning of the courtship, the NW females were usually passive or flew away from the male; however, in the stage of male positioning prior to copulation, 70% of observed females rejected the males by flapping their wings or pushing the male's abdomen with their legs (Li *et al.*, 1989). Males also sometimes terminated the courtship, but it typically occurred during the first contact with the female (Li *et al.*, 1989). In MEAM1, male rejection was observed more frequently throughout the whole courtship, ranging from 25% during the first contact to approximately 50% during the male positioning, accounting for an average of 32.7% of all courtship sequences aborted by males. The proportion of couples reaching successful copulation resulting from initial courtship ranged from 7.7% in NW to 12% in MEAM1 (Li *et al.*, 1989; Perring and Symmes, 2006).

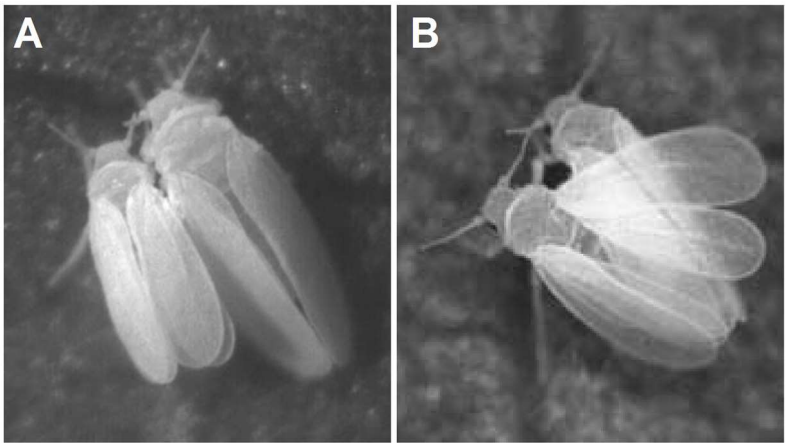


Figure 2-7: Mating behaviour of *B. tabaci* (MEAM1). **A:** Antennal drumming phase (Zang and Liu, 2007). **B:** Male covering female with his wings prior to copulation (Ruan *et al.*, 2007).

Life history traits

The total developmental time of *B. tabaci* from oviposition to adult emergence can vary considerably. The large effect of temperature on the developmental time of insects is well known (Ratte, 1985; Régnière *et al.*, 2012), and it is the case for *B. tabaci* as well (Gerling *et al.*, 1986; Basu, 1995; Bonato *et al.*, 2007). Extremes ranged from 11 days in summer (Pruthi and Samuel, 1942) to 107 days in winter (Husain and Trehan, 1933), both recorded in India. For a Spanish MED population reared on pepper, the development time varied from 16 days at 33°C to 42.7 days at 17°C (Muñiz and Nombela, 2001). Similar results were obtained with a French MED population reared on tomato, in which the development from egg to adult took 20 days at 30°C and 56 days at 17°C (Bonato *et al.*, 2007). The developmental time was negatively correlated with temperatures between 17 and 30°C, while no significant difference was observed between 30 and 35°C (Bonato

et al., 2007). The developmental times for each immature stage separately and total developmental times at different temperatures are summarised in Table 2-3.

Table 2-3: Developmental times (in days) of life stages of MED whiteflies at different temperatures on pepper[(MN)] and tomato[(B)].

		Nymph instar				
Temp.	**Egg**	**1st**	**2nd**	**3rd**	**4th**	**Egg to adult**
17°C	11.2 [MN] 21.5 [B]	5.2 [MN] 9.4 [B]	4.8 [MN] 7.5 [B]	9.7 [MN] 8.0 [B]	11.9 [MN] 9.4 [B]	42.7 [MN] 55.8 [B]
20°C	8.9 [MN]	4.0 [MN]	3.9 [MN]	7.8 [MN]	9.5 [MN]	34.1 [MN]
21°C	14.0 [B]	7.1 [B]	4.1 [B]	8.8 [B]	5.6 [B]	39.6 [B]
23°C	6.4 [MN]	3.1 [MN]	3.5 [MN]	5.3 [MN]	6.1 [MN]	24.4 [MN]
25°C	10.4 [B]	4.3 [B]	3.7 [B]	3.8 [B]	3.4 [B]	25.6 [B]
26°C	5.1 [MN]	2.9 [MN]	2.9 [MN]	4.7 [MN]	5.1 [MN]	20.7 [MN]
30°C	4.1 [MN] 7.7 [B]	2.3 [MN] 3.2 [B]	2.6 [MN] 3.3 [B]	3.7 [MN] 3.5 [B]	3.5 [MN] 2.5 [B]	16.1 [MN] 20.2 [B]
33°C	4.0 [MN]	2.1 [MN]	2.6 [MN]	3.9 [MN]	3.8 [MN]	16. [MN]
35°C	4.7 [MN] 6.5 [B]	2.4 [MN] 3.9 [B]	2.9 [MN] 3.3 [B]	4.2 [MN] 3.5 [B]	4.0 [MN] 3.3 [B]	18.1 [MN] 20.5 [B]

MN = Muñiz and Nombela (2001)
B = Bonato *et al.* (2007)

Another factor influencing the developmental time is the food supply from the host, *i.e.* the plant species and the plant physiological state. This means that stressors such as low light, high temperature and extreme humidity can influence the whitefly development both directly and indirectly through the host (Gerling *et al.*, 1986). According to Coudriet *et al.* (1985), mean development times of a *B. tabaci* population from California varied from 18.6 to 29.8 days when reared on different plant hosts at the same temperature (25–27°C).

Adult longevity is also influenced by temperature, as adults live longer in winter or at lower temperature in laboratory conditions (Basu, 1995). In MED, female longevity decreased exponentially with increasing temperature; 39.6 days at 17°C decreased to 8.5 days at 35°C (Bonato *et al.*, 2007). In addition, females lived longer than males by 1.6–4.2 fold in winter and 2–4.3 fold in summer (Basu, 1995). The greater longevity of *B. tabaci* females than males occurs regardless of the whitefly species, host plant and temperature (Powell and Bellows, 1992; Lü *et al.*, 2014; Khan and Wan, 2015).

Female fertility followed a similar negative correlation with respect to temperature as female longevity, with the exception of 17°C (Bonato *et al.*, 2007). During their lifespan, MED females laid on average 105.3 eggs at 21°C, and only 41 eggs at 35°C (Bonato

et al., 2007). When MED females were exposed to a heat shock of -12°C or 45°C, this led to a significant and non-significant reduction of the oviposition rate, respectively (Lü *et al.*, 2014). Interestingly, the oviposition rate increased significantly if females were exposed to high-temperature heat shock prior to the low-temperature heat shock (Lü *et al.*, 2014). The opposite cross-stress treatment did not change the number of eggs laid (Lü *et al.*, 2014).

Key attributes in traits such as mating behaviour, fecundity, insecticide resistance and heat tolerance have been linked to the mechanisms underlying the invasive characteristics of MED and MEAM1 (Pascual, 2006; Liu *et al.*, 2007; Cui *et al.*, 2008; Wan *et al.*, 2009; Chu *et al.*, 2010b; Yu *et al.*, 2012).

2.3.2 Geographic spread and losses caused by the invasive *B. tabaci* species

Spread of the invasive B. tabaci species

Whiteflies *B. tabaci* are distributed worldwide, predominantly in tropical and subtropical regions, but they also infest greenhouses in temperate climates (Mound and Halsey, 1978; Basu, 1995). Members of the species complex can be found on all continents except Antarctica (Oliveira *et al.*, 2001). While most *B. tabaci* species are indigenous within a particular geographic region, two of them (MEAM1 and MED) have spread invasively from their home range in the Mediterranean and Middle East regions around the world (De Barro *et al.*, 2011).

The pest status of *B. tabaci* became more serious after it caused a crisis in cotton farming in the Sudanese Gezira Irrigation Scheme in the 1970s (Eveleens, 1983). Arguments explaining this outbreak related to favourable changes in agricultural practices (e.g. intensification and increased fertilisation) and to the use of insecticides, either directly through promoting the development of insecticide resistance, or indirectly through eliminating natural enemies (Eveleens, 1983; Dittrich *et al.*, 1990; Castle, 1999). Another prominent outbreak occurred in the 1980s in southern United States. *B. tabaci* whiteflies grew to excessive densities in cotton fields in the south-western US and on poinsettia in greenhouses in Florida (Toscano *et al.*, 1998). Since then the infestations have spread into multiple southern states and onto a range of vegetable and field crops (Toscano *et al.*, 1998). During the 1990s, damaging infestations by *B. tabaci* were also reported from Mexico, Central America and the Caribbean Islands, Brazil, Mediterranean Basin, Middle East, China and Australia (Oliveira *et al.*, 2001).

The whitefly outbreaks in the 1980s and 1990s represented a key moment for identification of different biotypes/species within *B. tabaci*. After the heavy infestation

in the US, the new invasive whitefly was identified as the B-biotype/MEAM1 species to distinguish it from the indigenous A-biotype/NW species (see section 2.1.2). Since then, the MEAM1 species has spread into at least 54 countries in six continents (De Barro *et al.*, 2011).

The devastating invasions of MEAM1 around the world have led to a higher awareness and more routine field screening of *B. tabaci* populations and the level of their insecticide resistance (Dennehy *et al.*, 2005; McKenzie *et al.*, 2009). Importantly, it also led to the development of safety measures. Not long after the threat of MEAM1 was suppressed with novel insect growth regulators (buprofezin, pyriproxyfen) and other management measures (Dalton, 2006), the field surveys started detecting samples of another invasive whitefly, the MED.

The MED was discovered on the Iberian Peninsula where it coexisted with MEAM1 (Guirao *et al.*, 1997), but eventually became predominant (Simón *et al.*, 2007). It was later found that the geographic location of MED was not restricted to the Iberian Peninsula, but that it was widely spread across the whole Mediterranean basin. It was reported in Greece (Tsagkarakou *et al.*, 2007), France (Dalmon *et al.*, 2008), Tunisia (Chermitti *et al.*, 1997; Bel-Kadhi *et al.*, 2008), Morocco, Canary Islands (Monci *et al.*, 2000), Italy (Demichelis *et al.*, 2000), Israel (Horowitz *et al.*, 2003a), Crete (Roditakis *et al.*, 2009), Croatia (Žanić *et al.*, 2005), Egypt and Sudan (Ahmed *et al.*, 2010b).

Populations of MED outside the assumed home range in the Mediterranean Basin were first detected in China on morning glory and poinsettia in 2003 (Zhang *et al.*, 2005), in the US on poinsettia in 2004 (Dennehy *et al.*, 2005), in Japan on pumpkin, melon and tomato in 2004 (Ueda and Brown, 2006), in Mexico and Guatemala on poinsettia in 2005 (Martinez-Carrillo and Brown, 2007; Bethke *et al.*, 2009), in New Zealand on poinsettia and capsicums in 2006 (Scott *et al.*, 2007), in Burkina Faso and Togo on cotton and tomato in 2007 (Gnankiné *et al.*, 2013), in Argentina and Uruguay on pepper and melon in 2010 (Grille *et al.*, 2011) and in Brazil on pepper and sweet potato in 2013 (Barbosa *et al.*, 2015). In the UK, *B. tabaci* whiteflies have been intercepted at nurseries since 1987 (Bartlett, 1992), most often on poinsettia from the Netherlands and Israel (Cuthbertson *et al.*, 2010). At first the intercepted species was MEAM1, but in the early 2000's a shift to MED was detected (Powell *et al.*, 2012). Certain areas in Europe are still MED-free, such as Finland, Sweden, Ireland and the UK (Cuthbertson and Vänninen, 2015).

A major worry about the new invader MED is its exceptionally low susceptibility to insecticides, particularly to neonicotinoids, which were introduced to market in 1991

(Nauen and Denholm, 2005). MED strains from Spain, Italy and Germany were shown to have low susceptibility to multiple types of neonicotinoids with cross-resistance (Cahill *et al.*, 1996; Elbert and Nauen, 2000; Nauen *et al.*, 2002; Rauch and Nauen, 2003). The MED population identified in Arizona was practically unaffected by pyriproxyfen and showed a reduced susceptibility to buprofezin, mixtures of fenpropathrin and acephate, as well as to the neonicotinoids acetamiprid, imidacloprid, and thiamethoxam (Dennehy *et al.*, 2005). The rate with which MED developed insecticide resistance was higher than that of MEAM1 (Nauen and Denholm, 2005). MED populations from Spain and Israel were shown to be more resistant to neonicotinoids or pyriproxyfen than MEAM1 (Nauen *et al.*, 2002; Horowitz *et al.*, 2005). Both MEAM1 and MED were resistant to organophosphates, carbamates and pyrethroids, but MED was in addition less susceptible to several neonicotinoids and insect growth regulators (Scott *et al.*, 2007). This higher efficiency of overcoming the insecticide pressure is assumed to be a key competitive advantage of MED over MEAM1, as shown by the displacement of MEAM1 by MED in Mediterranean countries and China (Guirao *et al.*, 1997; Scott *et al.*, 2007; Chu *et al.*, 2010b). In insecticide-free conditions, however, MEAM1 outcompetes MED, probably due to its higher fecundity and via mating interference (Pascual and Callejas, 2004; Horowitz *et al.*, 2005) (see section 2.2.2).

The main source of the MEAM1 and MED outbreaks was attributed to international trade of ornamental plants (Dennehy *et al.*, 2005; Dalton, 2006; Bethke *et al.*, 2009; De Barro *et al.*, 2011; Rao *et al.*, 2011), which is also evident in the frequency with which the invasive species were found on poinsettias. From molecular diagnostic data, all mtCOI haplotypes of MEAM1 from the invaded regions connected to the home range via Israel, which has a very developed industry of exporting ornamental plants (De Barro and Ahmed, 2011). The numbers of MEAM1 and MED haplotypes associated with invasion outside their home ranges were very small; only one MEAM1 and three MED haplotypes made up 80% of GenBank records from the invaded regions (De Barro and Ahmed, 2011).

Economic damage caused by B. tabaci

Whiteflies damage their plant-hosts in direct and indirect ways. The direct damage is caused by withdrawing of the phloem sap by immature and adult whiteflies and causing chlorotic spots at the feeding sites that, at high pest population density, result in wilting and premature shedding of the leaves (Cock, 1986). In addition, the excreted honeydew accumulates on the plant and promotes the growth of bacteria and sooty mould, which has a negative impact on photosynthesis (Cock, 1986; Thompson, 2011). Feeding by

some *B. tabaci* species (MEAM1 and IO) also causes phytotoxic disorders, such as silver-leafing in squash (Yokomi *et al.*, 1990; Costa *et al.*, 1993a; Delatte *et al.*, 2005) or irregular ripening in tomato (Schuster *et al.*, 1990; Hanif-Khan *et al.*, 1998). In general, the direct damage from whitefly feeding weakens the plant, slows down its growth and reduces the infested crop's yield (Mound, 1965).

The indirect damage is caused by the transmission of plant viruses causing diseases. Members of the *B. tabaci* complex are recognised vectors of begomoviruses, criniviruses, carlaviruses, ipomoviruses and torradoviruses that infect various plant species and cause large economic losses (Jones, 2003; Seal *et al.*, 2006; Navas-Castillo *et al.*, 2011). Of these, begomoviruses are reported as economically the most important in agriculture, as they cause substantial yield losses, ranging from 20% to 100% (Brown and Bird, 1992; Cathrin and Ghanim, 2014). The global "success" of this genus is related to its great diversity (Navas-Castillo *et al.*, 2011) and there are over 200 species of begomoviruses that are exclusively transmitted by *B. tabaci* (Luan *et al.*, 2014).

Whiteflies affect a plethora of economically important crops, such as cotton, tomato, soya bean or squash, as well as staple crops important in subsistence farming, such as cassava. The whiteflies spread cassava mosaic disease (CMD) and cassava brown streak disease (CBSD), which are among the most significant threats to food security. These plant viral diseases cause devastating losses in sub-Saharan Africa (Holt and Colvin, 2001; Maruthi *et al.*, 2002a, 2002b). It was estimated that cassava-infesting whiteflies cause an annual loss of 1–2 billion USD, with an average yield loss about 50% in affected countries (Fauquet and Fargette, 1990; Legg *et al.*, 2014). The more abundant populations of *B. tabaci* were associated with the spread of a CMD pandemic through East and Central Africa (Legg, 1999). Estimated losses caused by CMD were 1.2–2.4 billion USD across sub-Saharan Africa in the 1990s (Thresh *et al.*, 1997), about 60 million USD in Uganda in the 1990s (Otim-Nape and Thresh, 1998) and over 10 million USD in Kenya in 1998 alone, causing severe food shortages and abandonment of cassava cultivation (Legg, 1999). The scale and impact of the CMD epidemic makes it one of the most damaging plant virus diseases in the world (Legg and Fauquet, 2004). The second cassava disease, CBSD, has been spreading at a fast rate in East and Central Africa since 2004, posing even more threat to food security of millions of farmers (Alicai *et al.*, 2007; Legg *et al.*, 2011). The tuber in the ground can be damaged by the disease without the plant showing noticeable symptoms, and the yield losses can reach 70–100% (Hillocks *et al.*, 2001; Pennisi, 2010). CBSD was estimated to cause annual losses of 75 million USD in East and Central Africa (Manyong *et al.*, 2012); however, this

figure was considered significantly underestimated (Manyong *et al.*, 2012; Patil *et al.*, 2015). Another estimate of annual losses for East and South Africa was 726 million USD (Maruthi, unpublished data in Mohammed *et al.*, 2017).

Cotton is also a crop seriously affected by whitefly infestation. The major damage by whiteflies was done by spreading the cotton leaf curl disease (CLCuD), causing outbreaks in Nigeria, United States, Pakistan, India, and China (Rahman *et al.*, 2017). This disease is caused by five different virus species around the world, namely *Cotton leaf curl Multan virus*, *Cotton leaf curl Bangalore virus*, *Cotton leaf curl Kokharan virus*, *Cotton leaf curl Allahabad virus*, and *Cotton leaf curl Gezira virus* (Muhire *et al.*, 2014; Rahman *et al.*, 2017). In addition, the honeydew deposited by whiteflies leads to lint stickiness and mould growth, which causes problems in the lint processing (Miller *et al.*, 1994; Ellsworth *et al.*, 1999). Cotton growers in the south-western US spent 154 million USD to control *B. tabaci* populations between 1994 and 1998 (Ellsworth *et al.*, 1999). CLCuD has decimated cotton yields in Pakistan and north-western India in two waves of epidemics, one in the 1990s and the current one which started in 2001 (Zubair *et al.*, 2017). Between 1992 and 1999 the estimated losses for cotton in Pakistan were 5 billion USD (Mansoor *et al.*, 1999). More recently, an epidemic of whitefly incidence and CLCuD symptoms occurred on over 90% of Bt cotton crop grown in north-western India, which led to widespread protests and even cases of suicide among the cotton farmers (Jadhav and Nair, 2015; Kranthi, 2015; Vasudeva, 2015).

Tomato crops around the world are particularly susceptible to many different begomoviruses (Czosnek, 2007; Gaur *et al.*, 2013). As many as 88 species of tomato begomoviruses are recognised by the International Committee on Taxonomy of Viruses (King *et al.*, 2018). Tomato yellow leaf curl disease (TYLCD) is one of the most devastating plant diseases worldwide (Hanssen *et al.*, 2010; Péréfarres *et al.*, 2012). Infection during early growth stages can lead to a total loss of the crop (Nakhla *et al.*, 1994; Picó *et al.*, 1996). Outbreaks of TYLCD have usually been associated with large populations of whiteflies (Cohen and Berlinger, 1986; Czosnek, 2007). The *Tomato yellow leaf curl virus* (TYLCV) associated with the invasive MEAM1 populations spread from the Middle East to the Americas, Mediterranean, Japan, North Africa and La Réunion (Seal *et al.*, 2006). The control of TYLCV is difficult, expensive and not always successful (Polston *et al.*, 1999; Czosnek, 2007). For example, the yield losses in Israel were often reported to be 100% despite almost daily insecticide treatments (Cohen and Antignus, 1994). In the US the MEAM1 invasion harmed tomato crops in two ways; by causing irregular tomato ripening (Schuster *et al.*, 1990) and by transmitting *Tomato mottle virus*

(Kring *et al.*, 1991). The estimated cost of the yield loss and control measures for tomatoes in Florida in 1990–1991 was 141 million USD (Schuster *et al.*, 1996).

B. tabaci infestations also caused large losses in Latin America after the widespread implementation of soyabean cultivation. *Bean golden mosaic virus* and *Bean dwarf mosaic virus* have devastated soyabean yields in Brazil, Argentina and Bolivia (Morales and Anderson, 2001). In China, vegetable and ornamental crops in at least 22 provinces were threatened by whiteflies and the viruses they transmit, namely TYLCV, *Tomato leaf curl virus* and *Squash leaf curl virus* (Ren *et al.*, 2001).

The most devastating outbreaks worldwide have usually been associated with invasive, rather than indigenous species of the *B. tabaci* complex. For example, the invasive MED and MEAM1 species were both about twice as efficient at transmitting TYLCV when compared with an indigenous Asia II-1 species in China (Li *et al.*, 2010). The MEAM1 was called a “supervector” and a “superbug”, given the number and diversity of viruses it transmits, its polyphagy, fertility, dispersal and resistance to insecticides (Barinaga, 1993; Gilbertson *et al.*, 2015). The estimated economic loss caused by the MEAM1 infestation in the US in 1991 was over 500 million USD (Perring *et al.*, 1993). In the Californian Imperial Valley alone, the annual loss between 1991 and 1995 was over 100 million USD (Oliveira *et al.*, 2001). In 1998, the reported cost of losses and pest control expenses since the introduction of MEAM1 into the US was over two billion USD (Toscano *et al.*, 1998).

2.3.3 Pest management strategies

To stop the spread of viral diseases devastating the crops, it is important to target the vector organism responsible for transmitting the plant viruses. In many cropping systems, whitefly infestations have primarily been controlled by insecticides (Sharaf, 1986; Naranjo *et al.*, 1998; Zheng *et al.*, 2013). One constraint on this approach is that not all farmers can afford or have access to these chemicals. Other important fundamental issues with using insecticides are the contamination and collateral damage they cause to non-pest insects, the environment and the pest’s ability to develop resistance. Because of the high capacity of MED and MEAM1 to develop insecticide resistance and the environmental reasons, it is vital to combine multiple, compatible approaches and minimise the use of chemical control. This approach is called integrated pest management (IPM) and it has four major cornerstones: (i) chemical control, (ii) biological control, (iii) cultural practices and (iv) host plant resistance (Hilje *et al.*, 2001).

Chemical control

Chemical control remains an important part of IPM, but the use of insecticides should be judicious and selective in order to extend the longevity and efficiency of this method, as well as minimise the unwanted impact on the environment (Naranjo *et al.*, 1998; Naranjo and Ellsworth, 2009). This is achieved by using action thresholds (*i.e.* the number of adult *B. tabaci* per leaf at which the treatment should be applied) and insecticide resistance management (*i.e.* measures taken to forestall resistance development by conserving the pest's susceptibility) (Phillips *et al.*, 1989; Naranjo *et al.*, 1998).

An alternative to conventional insecticides with active ingredients is to use environmentally safer chemicals that kill the insects by their physical mode of action. For example, a glasshouse trial of the refined petroleum oil Tri-Tek caused 100% mortality of MED eggs and adults and 69% mortality of MED 2nd instars (Cuthbertson *et al.*, 2012; Cuthbertson and Collins, 2015). Such chemicals can be used in combination with other methods, such as biological control agents, to increase the overall efficiency of the pest control (Cuthbertson *et al.*, 2012; Cuthbertson and Collins, 2015).

Biological control

The biological control is a significant aspect of pest management that employs or augments the pests' natural enemies. It can exploit the indigenous fauna or import exotic species and can be used both indoors and outdoors (Gerling *et al.*, 2001). Biocontrol organisms can include predatory and parasitoid insects or arachnids (Arthropoda), pathogenic species from a different phylum (Nematoda) or a different kingdom (Fungi).

There are at least 113 species from 9 orders of arthropods listed as predators of *B. tabaci*, most of which are beetles (family Coccinellidae), true bugs (Miridae, Anthocoridae), lacewings (Chrysopidae, Coniopterygidae), mites (Phytoseiidae) and spiders (Araneae) (Gerling *et al.*, 2001). By the year 2000, 19 species were commercially available for the biocontrol of whiteflies (Gerling *et al.*, 2001). The efficiency of whitefly predators is understudied and their role is often undervalued (Naranjo, 2001). Examples of successful applications in greenhouses are the mirid bug *Nesidiocoris tenuis* and the predatory mite *Amblyseius swirskii*. *N. tenuis* reduced *B. tabaci* population by 90% on tomato, and *A. swirskii* by about 70% on cucumber (Calvo *et al.*, 2009, 2011). Larvae of the beetle *Serangium parcesetosum* are also voracious feeders with high preference for *B. tabaci* nymphs (Legaspi *et al.*, 1996). Mean cumulative predation on whitefly eggs and nymphs by these larvae in their lifetime was 2,586 at 20°C and 4,909 at 30°C (Legaspi *et al.*, 1996), or 1000 in the case of the undescribed *Serangium* sp. observed in cassava field in Uganda (Asiimwe *et al.*, 2007). The release of *S. parcesetosum* suppressed the

whitefly infestation by 97–98% in cage experiments (Kutuk *et al.*, 2008) and by 86.5–90.7% in greenhouse conditions (Al-Zyoud *et al.*, 2007).

Whitefly parasitoids are small wasps (Hymenoptera), females of which lay eggs into the nymphs of their hosts and the hatched eggs continue their development inside the whitefly nymph until eclosion of the adult wasp (Godfray, 1994). About 34 *Encarsia* and 12 *Eretmocerus* species parasitise *B. tabaci*, out of which 11 and 8 species, respectively, were evaluated in biological studies (Gerling *et al.*, 2001). *En. formosa* successfully parasitised both MEAM1 and MED whiteflies in China (Liu *et al.*, 2016). A combined release of multiple natural enemies, however, can affect the predation/parasitism efficiencies of individual species. For example, in a greenhouse trial with Q1 MED on tomato, the parasitism rate of *En. formosa* and *En. sophia* was increased by the presence of the predator *Harmonia axyridis*, but the adult parasitoid emergence was decreased (Tan *et al.*, 2016). In addition, the predator showed a strong preference for non-parasitised whitefly nymphs (Tan *et al.*, 2016).

Other sources of natural enemies can be found amongst entomopathogenic nematodes from Steinernematidae and Heterorhabditidae, which can penetrate and ultimately kill many economically important insect pests (Poinar, 1990). The infective nematode juveniles can be sprayed on plants as an alternative to chemical insecticides (Head *et al.*, 2004; Garcia-del-Pino *et al.*, 2018), although the addition of adjuvants is needed to increase their mobility and prevent desiccation (Head *et al.*, 2004; Qiu *et al.*, 2008). Promising species for the control of *B. tabaci* are *Steinernema feltiae* and *S. carpocapsae*, which most efficiently attacked the 2nd and 3rd nymphal instars (Cuthbertson *et al.*, 2003a, 2007a). While some variation in mortality was observed on different plants (Head *et al.*, 2004), laboratory and greenhouse experiments demonstrated 60–90% whitefly mortality on various host plants including tomato, verbena, cucumber, poinsettia and chrysanthemum (Cuthbertson *et al.*, 2003a, 2007a, 2007b). After combining *S. feltiae* with an adjuvant, the whitefly mortality reached 86% on collard and 90% on hibiscus (Qiu *et al.*, 2008). A field trial in Thailand using another nematode species *S. thailandensis* suppressed the whitefly population on cilantro by 35–46% (Pangnakorn and Chuenchooklin, 2016). The compatibility of nematodes with some insecticides enhances the potential for their use in IPM (Cuthbertson *et al.*, 2003b, 2008).

A rich source of biopesticides can also be found among fungi. Many fungal species are entomopathogens, which means that their conidia (spores) attach to the insect body, germinate and penetrate the cuticle and their continued growth inside the host's body ultimately results in the host's death (McCoy *et al.*, 1988). Over 20 fungal species infect

whiteflies (Fransen, 1990; Lacey *et al.*, 1996; Steenberg and Humber, 1999). Some of these are commercially available as whitefly biocontrol agents, such as *Verticillium lecanii*, *Beauveria bassiana* and *Paecilomyces fumosoroseus* (Xia *et al.*, 2013). The search for novel fungal entomopathogens is ongoing. For example, a field survey in China isolated *Paecilomyces lilacinus*, *Lecanicillium psalliotae*, *Aspergillus ustus*, *Isaria fumosorosea* and *Metarhizium anisopliae* var. *anisopliae* (Dong *et al.*, 2016). In laboratory assays, all these isolates induced 30–90% mortality in 2nd instar of MEAM1 whiteflies, with the most virulent isolate identified as *P. lilacinus* (Dong *et al.*, 2016). In greenhouse conditions, *B. bassiana*, *V. lecanii* and *M. anisopliae* caused 35–46% mortality four days after treatment and reached 90.5–100% seven days after treatment (Abdel-Raheem and Al-Keridis, 2017). Another fungal species *Lecanicillium muscarium* tested in laboratory conditions caused up to 83% mortality of 2nd instar MEAM1 nymphs eight days after treatment (Ali *et al.*, 2017). Mortality increased to 95% when used in synergistic action with the botanical insecticide matrine (Ali *et al.*, 2017). A synergistic effect was also observed when combining *B. bassiana* and Tri-Tek oil, as the mortality of 2nd instars of MED rose from 65–69% when used in isolation to 95.5% when combined (Cuthbertson and Collins, 2015).

It is apparent that a great variety of organisms are available that can potentially serve as biological control. Furthermore, the ability to combine multiple biological agents with each other, and/or with aspects of chemical control, increases their potential to be an efficient strategy for controlling *B. tabaci* whiteflies. A major challenge, however, is to find the right balance among multiple environmental factors affecting the biological control agents, as well as ensuring their specificity for the targeted pest organism.

Cultural control

The cultural control includes a diverse group of practices that are implemented with the aim of making the environment less favourable for reproduction and survival of the pest (Hilje *et al.*, 2001; Ren *et al.*, 2001). Such practices include physical barriers blocking the pest from accessing the crop, for example mesh, plastic, wood or living barriers consisting of tall plant species (Hilje *et al.*, 2001). Another strategy is to manipulate the pest's behaviour by interfering with its visual or olfactory cues, which can be achieved by using mulches and ground covers (Hilje *et al.*, 2001). These methods can have a repellent or distracting effect on the pest (Antignus *et al.*, 1996; Cradock *et al.*, 2002). The distraction can also have the form of "trap crop", which is a highly favourable plant for the pest, but often sprayed with an insecticide to prevent the pests from migrating back onto the main crop (Schuster, 2004; Lin *et al.*, 2015). Apart from trap crops, intercropped plant species

can also be chosen with the aim of supporting populations of natural whitefly enemies (Hilje *et al.*, 2001).

Other cultural measures for limiting pest outbreaks involve decisions in crop choice, timing and location. An early or late planting date and avoiding a continuous overlap of susceptible crops (ideally including crop-free periods) can suppress the pest population (Hilje *et al.*, 2001; Ren *et al.*, 2001). For example, during the 1980s in the south-western US, large numbers of whiteflies migrated from a terminated cotton field directly onto a newly planted lettuce and melon fields, which almost decimated the local production (Duffus and Flock, 1982; Blua *et al.*, 1994; Nuessly *et al.*, 1994). This cycle was broken by introducing short-season cotton and delaying the planting of vegetables and melons, which reduced the whitefly populations and virus incidence (Blua *et al.*, 1994; Nuessly *et al.*, 1994). The virus incidence can be further reduced by limiting the vector's access to sources of the virus, e.g. by removal of the infected plants and all residues after harvest and by controlling the adjacent non-crop areas containing weeds, which can provide overwintering sites for whiteflies and serve as virus reservoirs (Muñiz, 2000; Ren *et al.*, 2001; Simmons *et al.*, 2008). Lastly, a crop can be made less suitable for the pest by modifying the irrigation regime. Water-stressed cotton plants were shown to be more susceptible to whiteflies and carry higher numbers of nymphs (Mor, 1987; Flint *et al.*, 1996). The method of the irrigation can also play a role, as watering via sprinklers led to a significant reduction of immature whitefly density compared with furrow irrigation (Castle *et al.*, 1996).

Host resistance

The crop choice can involve resistant crop varieties, which are less suitable for the pest and/or more tolerant to the pest damage (Ren *et al.*, 2001). Traits often associated with whitefly resistance are leaf hair density, leaf thickness, glandular trichomes and production of defensive chemicals, contributing to a higher mortality of whiteflies attempting feeding and/or reproduction on the plant (Williams *et al.*, 1980; Butter and Vir, 1989; De Ponti *et al.*, 1990; Channarayappa *et al.*, 1992; Jauset *et al.*, 2000; Parsa *et al.*, 2015). The advantage of this pest management method is that is does not require complete eradication of the pest in order to prevent crop damage (Ren *et al.*, 2001).

The relative tolerance of a crop to a pest or a pathogen can be ascertained in bioassays, which enable the identification of susceptible and resistant varieties. The resistant plants can then be used as a genetic source of resistance in breeding programmes (Bas *et al.*, 1992; Carabalí *et al.*, 2010a). For example, a wild tomato *Lycopersicon hirsutum* f. *glabratum* containing a 72-fold higher concentration of the natural insecticide

2-tridecanone than the cultivated tomato *L. esculentum* (Williams *et al.*, 1980) was used to breed whitefly-resistant hybrid lines (Bas *et al.*, 1992). Resistance to whiteflies (MEAM1) has also been identified in wild tomato relatives *Solanum pennellii*, *S. habrochaites*, *S. chilense*, *S. pimpinellifolium* and *S. galapagense* (Fancelli and Vendramim, 2002; Toscano *et al.*, 2002; Muigai *et al.*, 2003; Baldin *et al.*, 2005; Firdaus *et al.*, 2012). Mechanisms of whitefly resistance in tomato have been correlated with the density of glandular trichomes and the content of defence chemicals (Firdaus *et al.*, 2012). In cassava (*Manihot esculenta*), screening of 5000 clones for resistance to whitefly *Aleurotrachelus socialis* identified a range of insect mortality from 25% (on M Bra 12) to 72.5% (on M Ecu 72) (Bellotti and Arias, 2001). Breeding-crosses made between these two clones resulted in a new cassava variety Nataima-31, which was resistant to *A. socialis* (Arias *et al.*, 2004). Resistance to *A. socialis* was also identified in wild cassava *M. flabellifolia* and *M. tristis* (Carabalí *et al.*, 2010a, 2010b). Resistance to the pest, however, does not equal resistance to the viral disease. Screening of 22 cassava land races for resistance to both *B. tabaci* and CMD revealed that CMD-resistant varieties were highly susceptible to *B. tabaci*, which contributed to the whitefly population build-up during the CMD pandemic (Omongo *et al.*, 2012).

In addition to exploiting natural sources of resistance and using traditional breeding techniques, resistant crops can also be generated by targeted genetic modifications. A higher tolerance to *Tobacco mosaic virus* (TMV) was achieved by transforming tobacco with the gene for the TMV coat protein (Powell Abel *et al.*, 1986). The transformants showed a delayed development of tobacco mosaic disease or remained symptomless throughout the experiment (Powell Abel *et al.*, 1986). Plant resistance to the insect pest can be acquired by inserting a gene that, when expressed, causes insect mortality after it feeds on the plant.

In tomato, resistance to *B. tabaci* MED and MEAM1 was achieved by transforming the plants with *Mi-1.2* gene (Nombela *et al.*, 2003), which encodes a protein conferring resistance to *B. tabaci*, potato aphids and rootknot nematodes (Roberts and Thomason, 1986; Rossi *et al.*, 1998). Another strategy against the insect pest is the expression of a double-stranded RNA (dsRNA) designed against vital insect genes, which leads to their silencing via RNA interference (RNAi). This was achieved with transgenic tobacco expressing dsRNA targeted to silence insect gene for vacuolar ATPase enzyme, subunit A, which caused 34–83% mortality in a *B. tabaci* population form India (Thakur *et al.*, 2014). The RNAi could also target key detoxification genes, as was done with the glutathione S-transferase gene BtGSTs5 (Eakteiman *et al.*, 2018). Whiteflies feeding on

the transgenic *Arabidopsis thaliana* plants showed a decreased fecundity and prolonged developmental time, which would result in 5-fold population decrease during a growing season in Israel (Eakteiman *et al.*, 2018).

The key lesson from previous whitefly-control experience is that no single control method remains effective forever. The fecundity, polyphagy and insecticide resistance in the invasive *B. tabaci* populations call for an IPM approach to achieve efficient and sustainable pest control.

2.4 Host-plant relations within the *B. tabaci* complex

2.4.1 Host range of the *B. tabaci* species complex

Many reports, in which the host range of the *B. tabaci* complex has been mentioned, stated that it reportedly colonises over 600 species of host plants (Sharma *et al.*, 2008; Ahmed *et al.*, 2010b; Fontes *et al.*, 2012; Cuthbertson, 2014; Cuthbertson and Vänninen, 2015; Shah *et al.*, 2015). The host range statements in these papers usually referred to a previous review by Oliveira *et al.*, (2001), in which authors referred to the work of Greathead (1986), Mound and Halsey (1978) and Secker *et al.* (1998). Mound and Halsey (1978) collated data on the geographical location and host plants of whiteflies from collections of the Natural History Museum and reported 302 plants as hosts for *B. tabaci*. This list was later expanded by Greathead (1986) to 506 plants representing 74 families. In an abstract of a conference paper, Secker *et al.* (1998) stated that the number of plants was over 600, however, they attributed this number only to the MEAM1 putative species. In fact, several other authors used the figure of more than 600 plant hosts to describe the host range of just MEAM1 as well (Markham *et al.*, 1994; Xie *et al.*, 2011; Shah and Liu, 2013), while referring to Oliveira *et al.* (2001) who stated the information for *B. tabaci* collectively. In some cases, the figures were as high as 700 plants from 86 families (Fekrat and Shishehbor, 2007) or even 4200 plants from 74 families (Ren *et al.*, 2001), both citing the work of Greathead (1986). Lastly, Abd-Rabou and Simmons (2010) stated that the *B. tabaci* complex has over 1000 reproductive hosts, but cited works of multiple authors and organisations.

It is clear that some ambiguity exists in assessing the host range of *B. tabaci*. A major source of confusion was probably the changing notion about the systematics of this species complex. Before it was recognised as a complex of cryptic species, the host range was regarded for *B. tabaci* as one species. However, the views on the systematics within the *B. tabaci* complex appear to change faster than those on its host range. Albeit morphologically indistinguishable, it should not be assumed that individual species of the *B. tabaci* complex use the same sets of host plants and this ambiguity should therefore be resolved. It is very likely that the host ranges of individual species are smaller and partially overlapping subsets of the reported 506–1000 plant species. Indeed, a literature survey by Malka *et al.* (2018) revealed that among 16 putative species within the *B. tabaci* complex, only MEAM1 was a true generalist, other four putative species (Asia I, IO, MED and SSA1) had “extended” host ranges, while the rest exhibited relatively restricted host ranges.

Plant families with the largest percentages of species listed as *B. tabaci* hosts were Malvaceae, Cucurbitaceae, Euphorbiaceae and Convolvulaceae, which together with Brassicaceae involves most crops on which *B. tabaci* is a pest (Stansly and Naranjo, 2010). More recently, a detailed literature survey examining the host-plant data recorded for 16 putative species (or clusters of related putative species) identified nine families most often reported as hosts for species of the *B. tabaci* complex, *i.e.* Solanaceae (13/16), Cucurbitaceae (12/16), Lamiaceae (12/16), Convolvulaceae (11/16), Euphorbiaceae (11/16), Malvaceae (11/16), Asteraceae (10/16), Fabaceae (8/16) and Brassicaceae (6/16) (Malka *et al.*, 2018). The largest number of plant families reported as hosts were associated with MEAM1 (49/55 families), followed by Asia I (18/55), MED (17/55), "SSA" cluster (16/55), IO (15/55) and Asia II (10/55) (Malka *et al.*, 2018).

It is worth noting that there are different ways of collecting the data concerning host-plants. The host-plant ranges of individual putative species have primarily been assessed based on the field collection data, *i.e.* from which plant species the samples were collected (Anderson *et al.*, 2001; Sseruwagi *et al.*, 2005; Saleh *et al.*, 2012; Parrella *et al.*, 2012, 2014; Quintela *et al.*, 2016). Host-plant range assessments can also be studied in a more controlled environment by conducting experiments in a greenhouse or a laboratory. For example, Simmons *et al.* (2008) evaluated the host range of MEAM1 putative species in a greenhouse with 120 different plant species. The authors reported that 117 species from 18 genera had some level of whitefly infestation, out of which 49 species and 11 genera had not been reported as hosts before (Simmons *et al.*, 2008). Controlled experiments also allow for evaluating the effect of a particular variable in whitefly behaviour on different hosts, such as temperature (Muñiz and Nombela, 2001; Nava-Camberos *et al.*, 2001; Tsueda and Tsuchida, 2011) or relative humidity (Chu *et al.*, 2012b). Detailed observations were also made using the electrical penetration graph (EPG) technique, which provides information about the probing and feeding behaviour and so the suitability of a host plant by analysing waveform patterns (Jiang *et al.*, 1999; Liu *et al.*, 2012a; Liu *et al.*, 2013).

Most studies concerned with detailed examination of host performance and/or preference were focused on the invasive putative species MEAM1 and MED. The MEAM1 was studied alone (Nava-Camberos *et al.*, 2001; Simmons *et al.*, 2008; Xu *et al.*, 2012; Shah and Liu, 2013; Xie *et al.*, 2014), or in comparison with MED (Muñiz, 2000; Muñiz and Nombela, 2001; Nombela *et al.*, 2001; Iida *et al.*, 2009; Tsueda and Tsuchida, 2011; Liu *et al.*, 2012a; Jiao *et al.*, 2012, 2013; Sun *et al.*, 2013; Xu *et al.*, 2014). Occasionally, MEAM1 was studied in comparison to an indigenous putative species, such as Australia,

Asia II-1, and Asia II-3 (De Barro *et al.*, 2006; Zang *et al.*, 2006; Xu *et al.*, 2011; Ahmed *et al.*, 2014; Xu *et al.*, 2015). Such comparisons shed more light on the nature and dynamics of interactions between the indigenous species and the invader, or between the two invasive species.

2.4.2 Host range of the Mediterranean putative species

The performance of whiteflies from the MED putative species was usually studied in comparison to the MEAM1 putative species, as both have been spreading globally and, in some regions, occurred in sympatry. Plant hosts included in the analyses were mostly vegetable and fibre crops (tomato, pepper, cabbage, cucumber, aubergine, beans and cotton), occasionally ornamentals (poinsettia) and weeds.

On three summer weeds (*Malva parviflora*, *Capsella bursa-pastoris* and *Brassica kaber*), the fecundity of MED was significantly greater than that of MEAM1, while it was the opposite for *Lactuca serriola* (Muñiz, 2000). Among winter weeds, both putative species preferred *Datura stramonium* and *Solanum nigrum* to *Amaranthus retroflexus*, *Chenopodium album* and *Echinochloa crus-galli* and MED developed significantly faster than MEAM1 on the two weeds (Muñiz, 2000). Findings like these are relevant to the pest management strategies because weeds can serve as reservoirs supporting the whitefly expansion in the field (Gerling, 1984; Muñiz, 2000).

Tomato, cucumber and aubergine appeared to be similarly suitable hosts for both putative species, although some discrepancies occurred between studies. For example, the larval survival of MEAM1 on tomato was superior to MED in Iida *et al.* (2009), while the opposite was reported by Jiao *et al.* (2012). In addition, the EPG data suggested that MED fed better on tomato, while cucumber was more suitable for MEAM1 (Jiang *et al.*, 1999; Liu *et al.*, 2012a). Cotton was used as a similarly good host for both MED and MEAM1 in Sun *et al.* (2013), however, previous studies reported the superiority of MED on cotton based on the oviposition rate, nymph survival, adult longevity and EPG data (Liu *et al.*, 2012a; Jiao *et al.*, 2012, 2013). MEAM1 had a significantly shorter developmental time on cotton (Jiao *et al.*, 2012), but the fast development appeared to be a general characteristic of this putative species (Iida *et al.*, 2009; Jiao *et al.*, 2012). Cabbage was unambiguously more suitable for MEAM1 than for MED (Iida *et al.*, 2009; Liu *et al.*, 2012a; Jiao *et al.*, 2013), while poinsettia was a better host for MED than MEAM1 according to the EPG data and observations on oviposition preference and development (Liu *et al.*, 2012a; Jiao *et al.*, 2013). In a choice experiment with poinsettia, cotton and cabbage, MED females preferentially oviposited on poinsettia (about 60% of eggs compared with about 20% laid

by MEAM1 females) and the survival from egg to adult on poinsettia was also significantly higher in MED (about 75%) compared with MEAM1 (about 45%) (Jiao *et al.*, 2013).

Other plants on which MED performed better include beans and pepper. In the analysis including five cultivars of beans, MEAM1 was unable to develop on four of them, while MED successfully reproduced on all five cultivars (Iida *et al.*, 2009). On pepper, MED had a significantly shorter developmental time, with a higher nymph survival, higher fecundity and a higher rate of adult emergence (Muñiz and Nombela, 2001; Iida *et al.*, 2009; Chu *et al.*, 2012b). However, the results can vary between cultivars as MEAM1 outperformed MED on the 'Piquillo' cultivar of pepper (Muñiz and Nombela, 1997).

Studies concerned with the MED putative species rarely distinguished between different MED subgroups and without publishing their mtCOI sequence it was not possible to assign the population to Q1, Q2, Q3 or ASL subgroups. In some cases, it was possible to infer that the population under study probably belonged to the Q1 group, such as Spanish, Chinese or Japanese MED populations, because the Q1 MED is predominant in these countries (Chu *et al.*, 2008; Fujiwara *et al.*, 2015). Only a handful of studies have stated which MED subgroups were studied.

A Q1 population in Tunisia was reported on the ornamental plant *Lantana camara*, and also on tomato, aubergine and courgette (Saleh *et al.*, 2012). Q1 whiteflies were also found on various hosts in Italy, including pepper, tomato, radish, passion flower, vine, Chinese hibiscus, mint, black nightshade, blue morning glory and horseweed (Parrella *et al.*, 2012). Several years later, a spread of the Q2 and decline in the Q1 population was documented in the same region of southern Italy (Parrella *et al.*, 2014). The Q2 population was predominant on hosts from the family Solanaceae (tomato, aubergine, and black nightshade), while Q1 was collected from blue morning glory, confederate rose and gerbera, belonging to Convolvulaceae, Malvaceae and Asteraceae, respectively (Parrella *et al.*, 2014). In Israel, a Q2 population was shown to be able to feed on *Brassica* spp. due to its ability to circumvent the plant chemical defence (Malka *et al.*, 2016).

In sub-Saharan Africa, a difference in the host range was reported for Q1 and ASL, as Q1 was predominantly found on *Nicotiana tabacum*, while ASL (or "Ug4") occurred on *Cucurbita moschata*, *Sida acuta*, *Ipomoea batatas*, *Pavonia urens*, *Cucumis sativus*, *Cucurbita pepo*, *Leonotis nepetifolia* and *S. melongena* (Sseruwagi *et al.*, 2005; Mugerwa, 2018). A distinction between MED and ASL has also been made by Malka *et al.* (2018), who reported that ASL colonised plants from four families (Cucurbitaceae,

Lamiaceae, Malvaceae and Solanaceae), all of which overlapped with the 17 plant families utilised by MED.

Differences in the host-plant ranges and performances among whitefly populations point towards their differences at the molecular level, more specifically in the detoxification molecular machinery that is needed to cope with plant defence responses.

2.4.3 Whitefly detoxification mechanisms in response to plant defence

A damage or stress caused to a plant, such as an attack by an herbivore, triggers an induced response, which can increase the plant resistance to further attack (Karban and Myers, 1989). This response can involve an accumulation of defensive compounds (allelochemicals) that influence the herbivores' feeding behaviour, reproductive success or host plant selection (Inbar and Gerling, 2008). The herbivores, however, can develop mechanisms allowing them to cope with the plant defence. These mechanisms can take place before ingestion, which involves deactivation of the compound, or after ingestion, which includes prompt excretion, sequestration or metabolisation of the allelochemical (Després *et al.*, 2007).

An example of compound deactivation can be found in the Q2 population feeding on *Brassica* spp. The whitefly was able to cleave the sulfate group off intact glucosinolates, which are plant defence chemicals in order Brassicales (Malka *et al.*, 2016). The altered chemical (desulfoglucosinolate) cannot be broken down by myrosinase, which is a plant enzyme breaking glucosinolates into bioactive products (Malka *et al.*, 2016).

Metabolising plant allelochemicals after the ingestion requires detoxification systems, which were divided into two phases. Phase I involves enzymes like cytochrome P450 monooxygenases (P450s) or carboxylesterases (CarEs) that add polar groups to the substrates through oxidation, hydrolysis or reduction reactions (Lindroth, 1989; Schuler, 2011). These activated metabolites are then substrates for enzymes considered as phase II, which includes glutathione-S-transferases (GSTs), sulfotransferases and UDP-glucosyltransferases (UGTs) that conjugate them with glutathione, sulfate or glycosyl group, respectively, which increases the efficiency of their excretion (Cermak, 2008). These enzymes are important not only for dealing with plant allelochemicals, but are also closely linked to detoxification of other chemicals, such as synthetic insecticides (Byrne *et al.*, 2000; Li *et al.*, 2007; Liang *et al.*, 2007; Xie *et al.*, 2011). The plant allelochemicals and insecticides were together referred to as xenobiotics (Li *et al.*, 2007).

Insect resistance to xenobiotics typically involves an increase in the metabolic efficiency of the detoxification mechanisms, as well as a decrease in sensitivity of the site targeted by the xenobiotic compound (Li *et al.*, 2007). The change in enzymatic activity can be caused by a structural change of the protein through mutation in the coding DNA sequence (Claudianos *et al.*, 1999; Amichot *et al.*, 2004). Alternatively, the metabolic capacity can be increased by overexpression of some enzymes through mutations in regulatory elements (Maitra *et al.*, 2000), gene duplication (Wang *et al.*, 1991) or epigenetic control (Field, 2000).

The relationship between the transcriptional regulation of detoxification genes and the herbivore's feeding strategy is not yet fully understood (Halon *et al.*, 2015). In general, genes can be expressed constitutively independent of the plant defence, or their transcription can be inducible by encountering the allelochemicals (Li *et al.*, 2007). According to the feeding strategy, organisms can be roughly divided into generalists and specialists. While generalist herbivores are able to develop on a wide range of plants, usually spanning several families, the specialists are restricted to one plant or a group of closely related plants that share a similar defence system (Farrell and Mitter, 1990; Chew and Renwick, 1995; Becerra *et al.*, 2009; Hopkins *et al.*, 2009). Generalists typically have the ability to detoxify a broad range of allelochemicals, although less efficiently than specialist herbivores (Krieger *et al.*, 1971; Li *et al.*, 2004). In contrast, the specialists possess highly efficient enzymes with high specificity to the defensive compound in the narrow range of their hosts (Mao *et al.*, 2006). The constitutive expression is considered to be typical for specialists, while the induced resistance is speculated to be more widespread among generalists (Després *et al.*, 2007; Wen *et al.*, 2009).

Within the *B. tabaci* species complex, the MED and MEAM1 putative species are considered generalists (Halon *et al.*, 2015). Some authors proposed that MED has the capacity to adapt to a broader range of host plants than MEAM1 (Muñiz, 2000; Muñiz and Nombela, 2001; Iida *et al.*, 2009; Tsueda and Tsuchida, 2011; Jiao *et al.*, 2012). In contrast, results from a recent literature survey suggested that it was MEAM1 that had a substantially broader reported host range than MED (Malka *et al.*, 2018). A study comparing enzymatic activities of CarE, P450 and GST proteins before and after a host shift from cucumber to 55 different plants from 29 families reported that most values of CarE and P450 were significantly higher in MED compared with MEAM1 (Xu *et al.*, 2014). However, in the response to 7-ethoxycoumarin the activity of CarE was significantly (0.71-fold) lower in MED compared with MEAM1 (Guo *et al.*, 2014).

At the transcriptional level, expression of 43 out of 65 P450 genes was higher in MED than MEAM1, out of which eight were overexpressed by over 50-fold in MED (Guo *et al.*, 2014). Neither the enzymatic activity nor the expression level of GST differed between the putative species in these analyses (Guo *et al.*, 2014; Xu *et al.*, 2014). Another transcriptional analysis comparing constitutive expression among six putative species identified 105 detoxification genes with significantly different expression in at least one species, out of which nine were differentially expressed in MEAM1 and 16 in MED-Q1 (Malka *et al.*, 2018). Among genes that were plastically expressed in response to a host shift, a significant enrichment of genes from the P450 and UGT families was observed (Malka *et al.*, 2018). The level of gene expression was also compared between transcriptomes from the guts of MED and MEAM1 whiteflies, as the gut is likely to play a major role in detoxification of harmful substances in phloem during digestion and assimilation (Ye *et al.*, 2014). Among 58 differentially expressed orthologous genes, three P450 genes and one GST gene were significantly upregulated in MED (Ye *et al.*, 2014). In addition to the differences at the RNA and protein levels mentioned above, some of the genes involved in metabolism of xenobiotics (e.g. P450 4C1, GST) were found to be highly divergent between MED and MEAM1 putative species, and the high rate of non-synonymous mutations suggested that the genes evolved under positive selection (Wang *et al.*, 2011b; Ye *et al.*, 2014).

The relationships between plant defence chemicals and metabolic response in whiteflies were studied in more detail by comparing the expression levels of 18 detoxification genes between MED and MEAM1 after feeding on an artificial diet with and without six phytochemicals, including alkaloids, flavonoids and glucosinolates (Halon *et al.*, 2015). In response to these compounds, relatively small number of genes had a significantly different expression, which might reflect the simplified single-toxin diets (Halon *et al.*, 2015). However, differences were found between the putative species. MED showed a higher level of constitutive expression of detoxification genes than MEAM1, which was interpreted as indicating different defence strategies of these putative species (Halon *et al.*, 2015). It is possible that because MEAM1 allocates less energy into constitutive defence, it exhibits greater fitness under favourable conditions but is more severely affected under stress. On the other hand, the fitness of MED might be more stable due to the constitutive expression, but its fitness is relatively higher when confronted with plant defences (Halon *et al.*, 2015).

The evolution of different degrees of generalist and specialist feeding strategy and its potential effect on the diversification of *B. tabaci* species remains to be understood (Malka

et al., 2018). Understanding the molecular mechanisms underpinning the insect-plant interactions facilitate the identification of target genomic regions and devising specific pest management strategies.

3. Molecular characterisation of *B. tabaci* MED populations based on mtDNA[1]

3.1 Introduction

Molecular data obtained from DNA or protein sequences provide an invaluable source of information about organisms under study. It enables large-scale classification of specimens, which would be slow to achieve using morphological characters, or in the case of the *B. tabaci* cryptic species complex, it would not be possible. In addition, because the sequences were shaped by evolutionary processes, the data can be used to infer the phylogeny of the studied organisms and study the evolutionary relationships among them.

MtDNA has been a favoured molecule for these purposes, given its desirable properties as a molecular marker outlined in section 2.1.1. More recently, especially with the decreasing costs of HTS, sequences of the full mitochondrial genomes became available for many organisms. Among insects, about 500 species have complete, or nearly complete mitogenome sequence available and these include representatives from all 28 insect orders (Cameron, 2014).

Barcoding of animal DNA has also developed around mtDNA markers (Hebert *et al.*, 2003b). The practice of submitting DNA sequences generated for barcoding purposes to public data repositories such as GenBank (Clark *et al.*, 2016) has made them readily available for further studies including phylogenetic analyses. For *B. tabaci*, the biotype classification based on biological or protein biochemical characteristics (e.g. Costa and Brown, 1991; Burban *et al.*, 1992) has been superseded by molecular analyses based on the 657 bp sequence of the 3' region of the mtCOI gene used for whitefly barcoding (Frohlich *et al.*, 1999; Boykin *et al.*, 2007; Dinsdale *et al.*, 2010; Boykin *et al.*, 2012). Furthermore, adoption of the 3.5% threshold in nucleotide distance measured in the mtCOI barcoding region (Dinsdale *et al.*, 2010) led to the identification of at least 46 putative species within the *B. tabaci* complex (see Table 2-2).

Although the barcoding approach is attractive, simple and relatively fast, it has some fundamental problems, including: (i) the overlap between intra- and inter-species sequence divergences (Meier *et al.*, 2006; Lee *et al.*, 2013), (ii) the arbitrary approach to selecting a 'preferred' barcoding gene region, *i.e.* either the 5' region of mtCOI (Hebert *et al.*, 2003a) or its 3' region (Frohlich *et al.*, 1999), (iii) the large number of misidentified

[1] This chapter formed a part of publication Vyskočilová *et al.* (2018); page 231.

and erroneous sequences in online databases (Harris, 2003; Vilgalys, 2003; Balakirev *et al.*, 2017; Vyskočilová *et al.*, 2018), and (iv) difficulties differentiating NUMTs from mtDNA genes (Tay *et al.*, 2017a; Vyskočilová *et al.*, 2018).

Avoiding the inclusion of NUMTs in molecular analyses can be achieved by obtaining the mitogenome sequence via HTS. The output of an HTS run are millions of short reads that can subsequently be assembled *in silico* into larger contigs and/or genomes. The large scale of the process means that all sequences in the sample are likely to be represented (Morey *et al.*, 2013). Sequencing errors can thus be identified based on their low frequency, and high-copy mitochondrial sequences can also be distinguished from the low-copy nuclear pseudogenes (Tay *et al.*, 2017a). A limitation of the HTS approach is that short sequences are not reliable for assembling low-complexity and repetitive regions (Metzker, 2010; Morey *et al.*, 2013).

Several mitogenomes from members of the *B. tabaci* complex have been published (see Table 3-2). The first published *B. tabaci* mitogenome was the one of New World species, and it was generated exclusively by PCR and Sanger sequencing (Thao *et al.*, 2004). The mitogenome of MED from China was generated by a combined approach, in which gaps between HTS-derived transcriptome sequences of mitochondrial genes were filled by PCR and Sanger sequencing (Wang *et al.*, 2013). The remaining mitogenomes were generated by the HTS method only (Tay *et al.*, 2016, 2017a, 2017b; Vyskočilová *et al.*, 2018).

The aim of this chapter was to analyse the genetic relationships among the four MED populations and evaluate the reliability of the commonly used barcoding methods. The research objectives of this chapter were to: (1) assemble mitogenome sequences from HTS data, (2) reconstruct phylogenetic tree from available mitogenome sequences, (3) calculate nucleotide distances within mitochondrial DNA regions among our populations and more distant species and evaluate the accuracy of the partial mtCOI gene regions to represent multiple mitochondrial genes, and (4) evaluate the quality of published sequences used for study of the intraspecific structure of MED.

3.2 Material and Methods

3.2.1 Growing plants and rearing *B. tabaci* colonies

All plants were grown in a whitefly-free room at 28°C and relative humidity 60% with 14:10 h light:dark periods. Seeds were sown into a 1:1 mixture of loam-based compost (J. Arthur Bower's John Innes No. 2, UK) and coir/perlite 70/30 mix compost (Jiffy Tref Propagation Compost, UK). For experimental purposes, 10–15 cm plantlets were used; for maintaining whitefly colonies the plants were larger (10–30 cm). Plants were watered twice a week.

Colonies of *B. tabaci* populations (Table 3-1) were reared on aubergine plants (*Solanum melongena* 'Black Beauty') at 28 ± 2°C, 30% relative humidity and 14:10 L:D. Core colonies were maintained in rectangular cages (BugDorm, US) made out of nylon mesh (three sides and top) and plastic (bottom and front) with dimensions 45 x 44 x 44 cm (Figure 3-1 A).

Table 3-1: Summary of populations from which colonies of *B. tabaci* used in this study were established with associated GenBank accessions of their partial mtCOI sequences.

Country	Species	Group	Plant	Year	Source	GenBank
Spain	MED	Q1	Melon	2013	Jesús Navas-Castillo	MH357339
Sudan	MED	Q1	Cotton	1978	Rothamsted research	MH357340
Israel	MED	Q2	Cotton	2003	Shai Morin	MH357341
Uganda	MED? [a]	ASL	Okra	2013	Habibu Mugerwa	MH357342

[a] The inclusion of the ASL group in MED species has been reassessed (Mouton *et al.*, 2015; this study).

Experimental plants were transplanted into the John Innes No. 2 compost and enclosed in whitefly-proof Lock&Lock cages (Wang *et al.*, 2011a) made from two cylindrical food containers (Hanacobi, Korea) (Figure 3-1 A). The upper 1.8 L container had three circular openings (7–8 cm in diameter) covered by 180 µm metal mesh or 160 µm nylon mesh for ventilation. The lower 700 ml container had a 5–8 mm hole to enable watering. The lids were screwed together and had a 3–6 cm hole in the middle for the plant stem. Plantlets were placed into the Lock&Lock cage by removing their roots from soil, leading the roots through the hole, repotting them in soil inside the lower container, securing the stem with cotton wool (Figure 3-1 B) and covering with the upper container (Figure 3-1 C).

Figure 3-1: Cage types used for whitefly rearing. **A:** Aubergine plants enclosed in BugDorm cage for rearing core whitefly colonies. **B:** Components of the Lock&Lock cage. **C:** Aubergine plantlet rooted in soil in the lower container of Lock&Lock cage and secured with cotton wool. **D:** Aubergine plantlet enclosed in Lock&Lock cage ready to be used.

3.2.2 DNA extraction, PCR and Sanger sequencing

Samples of adult whiteflies were collected from core colonies using an aspirator and stored in 90% ethanol at -20°C. DNA extractions were carried out using the Chelex method (Walsh *et al.*, 1991) using Chelex 100 resin (Bio-Rad, US). Whiteflies were put individually into 1.5 ml tubes and 25 µl of stirred 10% Chelex solution was added. Samples were homogenised using individual sterilised plastic pestles and another 25 µl of stirred 10% Chelex solution was added. The mixture was incubated at 56°C for 20 min, subsequently at 95°C for 5 min and centrifuged for 5 min at 13,520 *g*. Tubes with the extracted DNAs were stored at -20°C and thawed prior to being used as a template for PCR amplification.

PCRs were performed using reSource Taq DNA polymerase (Source Bioscience, UK), or DreamTaq DNA Polymerase (Thermo Scientific, UK), following the manufacturers' protocols. A negative control, in which the DNA template was substituted by an equal volume of molecular grade water (Sigma-Aldrich, UK), was included in every PCR run. Primers were designed using Primer 3 software (http://primer3.ut.ee/, Rozen and Skaletsky, 2000) and synthesised by Sigma-Aldrich (UK). The primers and PCR conditions used to amplify the 3' region of mtCOI were 2195Bt (5'-TGRTTTTTTGGTCATCCRGAAGT-3') and C012/Bt-sh2 (5'-TTTACTGCACTTTCTGCC -3') with an expected amplicon size of 862 bp (Mugerwa *et al.*, 2018). The PCR program used was 2 min at 94°C, 30 cycles of 20 s at 94°C, 30 s at 50°C and 1 min at 72°C, and a final extension of 5 min at 72°C. The PCR products were purified on a column using reSource PCR purification kit (Source Bioscience, UK) or GeneJET PCR purification kit (Thermo Scientific, UK). Sanger sequencing was carried out by Source BioScience (Nottingham, UK) or GATC Biotech (Cologne, Germany).

3.2.3 Species identification based on partial mtCOI sequence

Partial 3' mtCOI DNA sequences were aligned in Geneious version 10.0.8 (Kearse *et al.*, 2012), trimmed to 657 bp and compared to (i) consensus sequences for 24 putative species of *B. tabaci* (Dinsdale *et al.*, 2010) to verify the MED putative species, and (ii) a set of 48 MED, four MEAM1 and three IO mtCOI haplotypes (Gueguen *et al.*, 2010) to place the populations in a recognised naming framework (Q1, Q2, Q3, ASL populations). The 48 MED haplotypes were acquired by reducing all MED sequences from Gueguen *et al.* (2010) to a set of unique haplotypes in FaBox DNA Collapser version 1.41 (Villesen, 2007). Neighbor-joining trees were generated in Geneious to visualise the similarities between our and reference sequences. A sequence from *B. afer* (GU220055) was used as outgroup for the whole *B. tabaci* complex, and IO was used as outgroup for the MED and MEAM1 sequences.

3.2.4 High-throughput sequencing from single males and quality control

Total DNA from single male whiteflies representing each laboratory colony was isolated using the silica spin-column method of Thangaraj *et al.* (2016) and quantified using a Qubit 2.0 Fluorometer (Invitrogen). The DNA (~30 ng in 100 µl of TE buffer) was sheared for 10 min using an ultrasonic cleaner (VGT-1620QTD) and size-selected (300–500 bp) using BluePippin (Sage Science). Sequencing libraries were prepared separately using each size-selected DNA pool using the NEBNext Ultra DNA Library Prep Kit for Illumina with NEBNext Multiplex Oligos for Illumina (New England BioLabs). Individual libraries

were checked by capillary microchip electrophoresis (MultiNA, Shimadzu), pooled in equimolar amounts, and then purified using Agencourt AMPure XP Beads (Beckman Coulter). The multiplexed library was sequenced on a single lane of an Illumina HiSeq 4000 platform (Novogene Bioinformatics Institute, Beijing, China) and 150 bp paired-end reads were generated.

Quality control of the reads was carried out with FastQC version 0.11.5 (Andrews, 2014) before and after trimming the reads with Skewer version 0.2.2 (Jiang *et al.*, 2014). The reads trimmed with a quality threshold of 40 and minimum length of 18 bp were used in subsequent analyses.

3.2.5 Mitogenome assembly and annotation

Draft mitogenomes of the Q1, Q2 and ASL males were assembled by iterative reference-guided assembly. The trimmed reads were mapped to the reference MED mitogenome JQ906700 (Wang *et al.*, 2013) using Geneious mapper set to medium-low sensitivity with up to 5 iterations. The draft mitogenomes were annotated in MITOS2 (Bernt *et al.*, 2013) and manually checked for start and stop codons in all protein-coding genes by translating into amino acid sequences using the invertebrate mtDNA genetic code in Geneious.

3.2.6 Calculating divergences in mitochondrial genes

Concatenations of 13 protein-coding genes and two rDNA genes were generated from 12 mitogenomes (Table 3-2). These 15 genes were extracted from the mitogenomes and orientated in 5'–3' direction. All 12 sequences of each gene were aligned by MUSCLE alignment tool in Geneious, trimmed to uniform length and subsequently concatenated for each population to form a 12,595 bp sequence. The *Bemisia* "JpL" sample (Tay *et al.*, 2017b) was not used due to the incomplete mitogenome sequence.

Table 3-2: Mitogenomes used for generating concatenations of 15 genes and for identifying pseudogenes (see section 3.2.9).

No.	Whitefly population	Accession	Reference
1	Spain Q1	MH205752	This study
2	Sudan Q1	MH714535	This study
3	Israel Q2	MH205753	This study
4	Uganda ASL	MH205754	This study
5	China MED Q1	JQ906700	Wang *et al.* (2013)
6	Burkina Faso MED Q1	KY951447	Tay *et al.* (2017a)
7	Peru MEAM1	KY951450	Tay *et al.* (2017a)
8	Australia I	KY951451	Tay *et al.* (2017a)
9	Indian Ocean	KY951448	Tay *et al.* (2017a)
10	*B. emiliae* (former Asia II-7)	KX714967	Tay *et al.* (2017b)
11	Asia I	KJ778614	Tay *et al.* (2016)
12	New World	AY521259	Thao *et al.* (2004)
13 [a]	*Bemisia* "JpL"	KX714968	Tay *et al.* (2017b)

[a] Only used for pseudogene identification (see section 3.2.9).

The 3' and 5' partial mtCOI sequences were obtained from mitogenomes by extracting the annotated sequences of full mtCOI gene, aligning them with reference partial sequences (Dinsdale *et al.*, 2010) or the LCO1490 and HCO2198 primers (Hebert *et al.*, 2003a), respectively, and trimming them to 657 bp. The nucleotide positions within the full mtCOI gene were 782–1439 for the 3' partial region, and 17–673 for the 5' region.

Uncorrected nucleotide pairwise distances (*p* distances) were used to express nucleotide sequence divergence as suggested previously (Collins *et al.*, 2012; Srivathsan and Meier, 2012; Collins and Cruickshank, 2013). The values were calculated using the alignment and distance matrix in Geneious.

The similarity among nucleotide divergences in the three mitochondrial regions (3', 5' and full mtCOI sequences), the 15 mitochondrial genes and a 657 bp concatenate of the first 11 tRNA genes in the mitogenome were compared. Nucleotide distances among 12 populations were calculated for each region in Geneious, transformed into distance matrices and compared by Mantel test in R using the vegan package (Oksanen *et al.*, 2018). The test returned a correlation coefficient and a statistical significance.

3.2.7 Sliding window analysis

Intra- and inter-specific sequence divergence levels within the Africa/Middle East/Asia Minor clade (Dinsdale *et al.*, 2010), excluding MEAM2 as a pseudogene artefact (Tay *et al.*, 2017a), were compared across the full length of mtCOI gene (1,542 bp) in a sliding window analysis in DnaSP 5.10.01 (Librado and Rozas, 2009), implementing the Jukes and Cantor (JC) nucleotide substitution model (Jukes and Cantor, 1969). DnaSP offers only the JC model for calculating nucleotide sequence divergence in the sliding window analysis. However, using the "best close match" criteria (Meier *et al.*, 2006), JC model ranked second after the *p* distance (Collins *et al.*, 2012), and so the use of the JC model by DnaSP for calculating nucleotide divergence represents an acceptable alternative. The size of commonly used partial mtCOI sequence (657 bp) was used for the window size, sliding along the sequence in 5 bp steps while plotting the nucleotide divergence (K) of JC-total between the compared samples. Sequences included in the sliding window analysis were Spain Q1, Israel Q2, Uganda ASL, Burkina Faso MED Q1 (KY951447), Peru MEAM1 (KY951450) and Indian Ocean (KY951448) (Tay *et al.*, 2017a).

3.2.8 Mitogenome phylogeny

The multiple sequence alignment of 12 concatenations of 15 mitochondrial genes (section 3.2.6) in fasta format was submitted to IQ-TREE software (http://iqtree.cibiv.univie.ac.at/, Nguyen *et al.*, 2015; Trifinopoulos *et al.*, 2016), along with a manually created nexus file indicating the data partitioning. Three different data partitioning schemes were compared in order to detect the effect of variable evolutionary histories of different loci in our dataset on the resulting phylogeny (Leavitt *et al.*, 2013). The schemes were: (i) no data partitioning, (ii) partitioning the protein-coding genes into two groups (1^{st} + 2^{nd} codon positions and 3^{rd} codon positions), and rDNA genes in the third group; and (iii) partitioning into 15 individual genes. The best-fitting substitution model for each partition was identified using ModelFinder (Kalyaanamoorthy *et al.*, 2017) and the maximum-likelihood tree was inferred using the edge-linked partition model (Chernomor *et al.*, 2016). Branch support was performed with the Ultrafast Bootstrap (Hoang *et al.*, 2018) with 1,000 replications. The tree was visualised in Dendroscope 3 (Huson and Scornavacca, 2012).

3.2.9 Pseudogene identification

Published 3' partial mtCOI sequences of MED groups Q1, Q2, Q3, Q4, Q5 and ASL (Chu *et al.*, 2008; Ahmed *et al.*, 2009; Gueguen *et al.*, 2010; Chu *et al.*, 2012a) were downloaded from GenBank (accessed 14th November 2017) and reduced to unique haplotypes in FaBox DNA Collapser 1.41. The haplotypes were then mapped to the full

mtCOI genes extracted from the mitogenomes of Q1, Q2 and ASL from this study to identify errors in the downloaded sequences.

The criteria used to categorise a partial mtCOI sequence as a potential pseudogene included either: (i) the presence of INDELs, (ii) the presence of premature stop codons inside the protein-coding sequence, (iii) anomalous polymorphisms (*i.e.* clusters of non-synonymous mutations or haplotypes with an outlying sequence divergence compared with the intra-group divergence, and/or (iv) non-synonymous mutations resulting in amino-acid substitutions in positions conserved across other *Bemisia* species. For the fourth criterion, only sequences that passed the first two criteria were analysed. Such DNA sequences were translated and aligned by Geneious alignment tool using Blosum62 matrix to a reference set of mtCOI amino acid sequences translated from 9 mitogenomes (numbers 5–13, Table 3-2). The likelihood of the non-synonymous substitutions occurring in MED haplotypes was estimated by comparing the intra- and inter-species variability in respective triplet positions. Sequences with high nucleotide divergence from the group were compared against the NCBI database using BLASTn (Chen *et al.*, 2015) to investigate the origin of the sequence.

3.2.10 Comparison of Q3 and ASL mtCOI haplotypes to similar populations

A Q3 population could not be included in our phylogenetic analyses because there was no mitogenome sequence available for the Q3 group from Burkina Faso described by Gueguen *et al.* (2010), nor the Q3 from Croatia (Chu *et al.*, 2012a). The genetic relatedness of these groups to the rest of the MED groups was, therefore, studied only at the level of the partial 3' mtCOI sequences available in GenBank. Five unique haplotypes of the two Q3 groups (Gueguen *et al.*, 2010; Chu *et al.*, 2012a) without identified errors (section 3.2.9) were compared to partial mtCOI sequences extracted from the HTS-derived mitogenomes generated in this study. The alignment and *p* distance matrix calculation were carried out in Geneious.

The 3' partial mtCOI sequences were also compared between the Uganda ASL from this study and published sequences of other MED-like populations from sub-Saharan Africa to see how our findings related to the previous literature. Our sample was compared to 27 haplotypes including "okra biotype" (Burban *et al.*, 1992; Omondi *et al.*, 2005), "Ug4" (Sseruwagi *et al.*, 2005) and "African silver-leafing" (Boykin *et al.*, 2007) samples. The sequence quality check, alignment and nucleotide distance calculation were carried out as above.

3.3 Results

3.3.1 Species and group assignment based on partial mtCOI sequence

The putative species identities of our colonies were checked by comparing their 3' partial mtCOI sequences to a set of reference consensus sequences (Dinsdale *et al.*, 2010). Our samples from Spain, Sudan and Israel clustered with the consensus sequence for Mediterranean species, while the population from Uganda was placed between Mediterranean and Middle East-Asia Minor 1 (Figure 3-2).

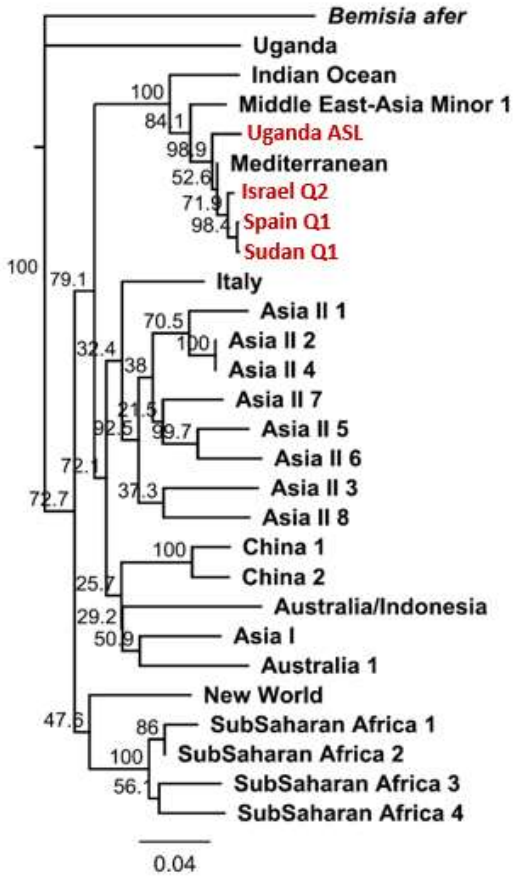


Figure 3-2: Rooted neighbor-joining tree based on 657 bp consensus sequences of partial 3' mtCOI sequences of *B. tabaci* putative species from Dinsdale *et al.* (2010) (in black) and populations in this study (in red), with *B. afer* sequence set as outgroup.

To place the whitefly populations within a framework of MED groups, the 3' partial mtCOI sequences were also compared to the dataset containing sequences from Q1, Q2, Q3 and ASL groups (Gueguen *et al.*, 2010). Our samples from Spain and Sudan clustered with the Q1 group, Israel with the Q2 group and Uganda with the ASL group (Figure 3-3). Israel and other Q2 samples from Israel, Cyprus and Turkey formed one cluster. Lastly, Uganda sample clustered more closely with ASL samples from Cameroon than those from Ivory Coast and Burkina Faso (Figure 3-3).

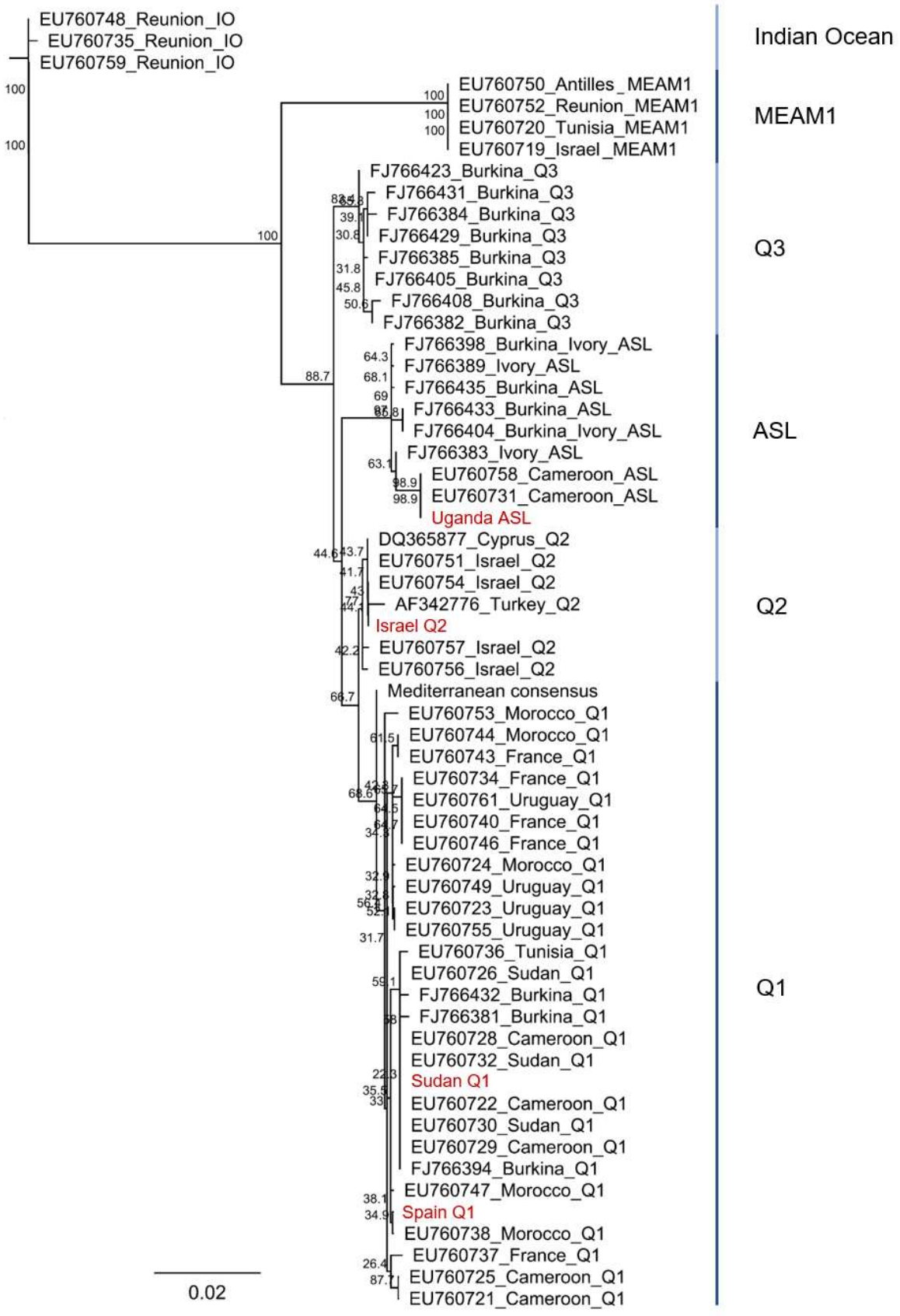


Figure 3-3: Rooted neighbor-joining tree based on 657 bp partial 3' mtCOI of MED, MEAM1 and IO sequences from Gueguen *et al.* (2010) (black) and our study (in red), with IO set as outgroup.

3.3.2 The assembly of mitogenomes

Draft mitogenomes were assembled from Illumina HiSeq reads of the total DNA from single adult males. The reads mapped to the reference mitogenome without any gaps. The assembly statistics are summarised in Table 3-3. After annotation and manual check of translations of the protein-coding genes, the mitogenomes were submitted to GenBank database (www.ncbi.nlm.nih.gov/genbank/; accession numbers in Table 3-2).

Table 3-3: Summary of the assembled mitogenomes.

Sample	Total reads	Mapped reads	Mean base coverage	Stadard deviation	Length (bp)
Spain Q1	78,657,159	513,612	4606.5	3502.6	15,632
Sudan Q1	80,600,051	238,833	2214.2	1494.0	15,634
Israel Q2	67,796,585	249,142	2231.5	1576.0	15,636
Uganda ASL	78,527,753	149,343	1351.0	666.0	15,637

All four mitogenomes generated in this study contained a standard number of genes for animal mtDNA (Boore, 1999): 13 protein-coding, 22 tRNA and 2 rDNA genes (Figure 3-4). The genes were in synteny with the reference MED mitogenome JQ906700 (Wang *et al.*, 2013).

There were small differences in the length of the low complexity AT-rich intergenic region between *cox3* and *tRNA-Ile*. The length in the reference MED mitogenome JQ906700 is 972 bp, however, in our samples the lengths were 971 bp (Uganda ASL), 973 bp (Spain Q1 and Israel Q2) and 975 bp (Sudan Q1).

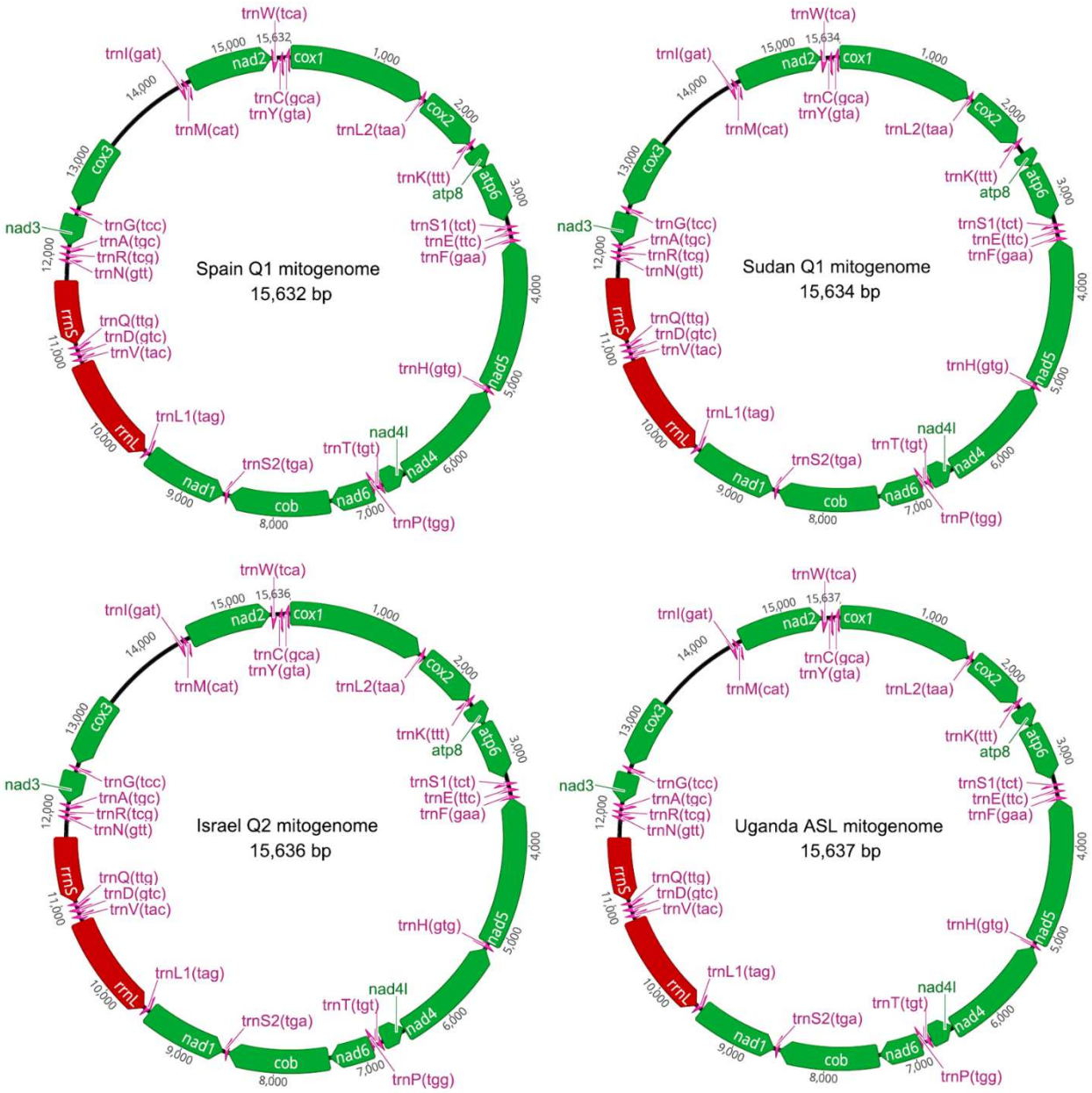


Figure 3-4: Assembled and annotated mitogenomes of the four populations in this study.

3.3.3 Mitogenome phylogeny of members of the *B. tabaci* complex

The genetic relationships among our samples and six other species of *B. tabaci* complex were reconstructed in a phylogenetic analysis based on the concatenations of 15 mitochondrial genes. Three different data partition schemes produced identical topologies with similar Ultrafast Bootstrap values (Figure 3-5 A–C).

Best-fit model according to the Bayesian information criterion (BIC) score was TVM+F+I+G4 for no data partitioning. The best substitution models identified in the second partition scheme were TVM+F+I+G4 (1st and 2nd codons), GTR+F+I+G4 (3rd codons) and TPM3u+F+G4 (rDNA genes). The substitution models used in the third partition scheme were TPM2u+F+G4 (atp6), K3Pu+F+I (atp8), HKY+F+G4 (cox1,cox2 and cytb), K3Pu+F+I+G4 (cox3), TPM3+F+I (nd1 and rrnL), TPM2u+F+I+G4 (nd2), TPM3+F+G4 (nd3), TIM3+F+I+G4 (nd4), TPM3u+F+I (nd4l), TIM3+F+I+G4 (nd5), TPM2+F+I (nd6), HKY+F+G4 (rrnS). Abbreviations and explanations of the models are listed in Appendix 1.

All branches had a high statistical support with Ultrafast Bootstrap value ranging from 70 to 100, which unlike the normal bootstrap corresponds roughly to 70–100% probability that the clades are true (Minh *et al.*, 2013). The partition scheme separating 15 individual genes had a slightly higher support than the other two schemes, *i.e.* 77% instead of 70% at the branch separating MEAM1 from ASL, Q1 and Q2 groups, and 96% instead of 94% at the separation between Q1 samples from Sudan and Burkina Faso (Figure 3-5 C). Uganda ASL was placed outside the Q1 and Q2 clade with 100% bootstrap support, between MED and MEAM1 species. IO had a basal position in the Africa/Middle East/Asia Minor clade.

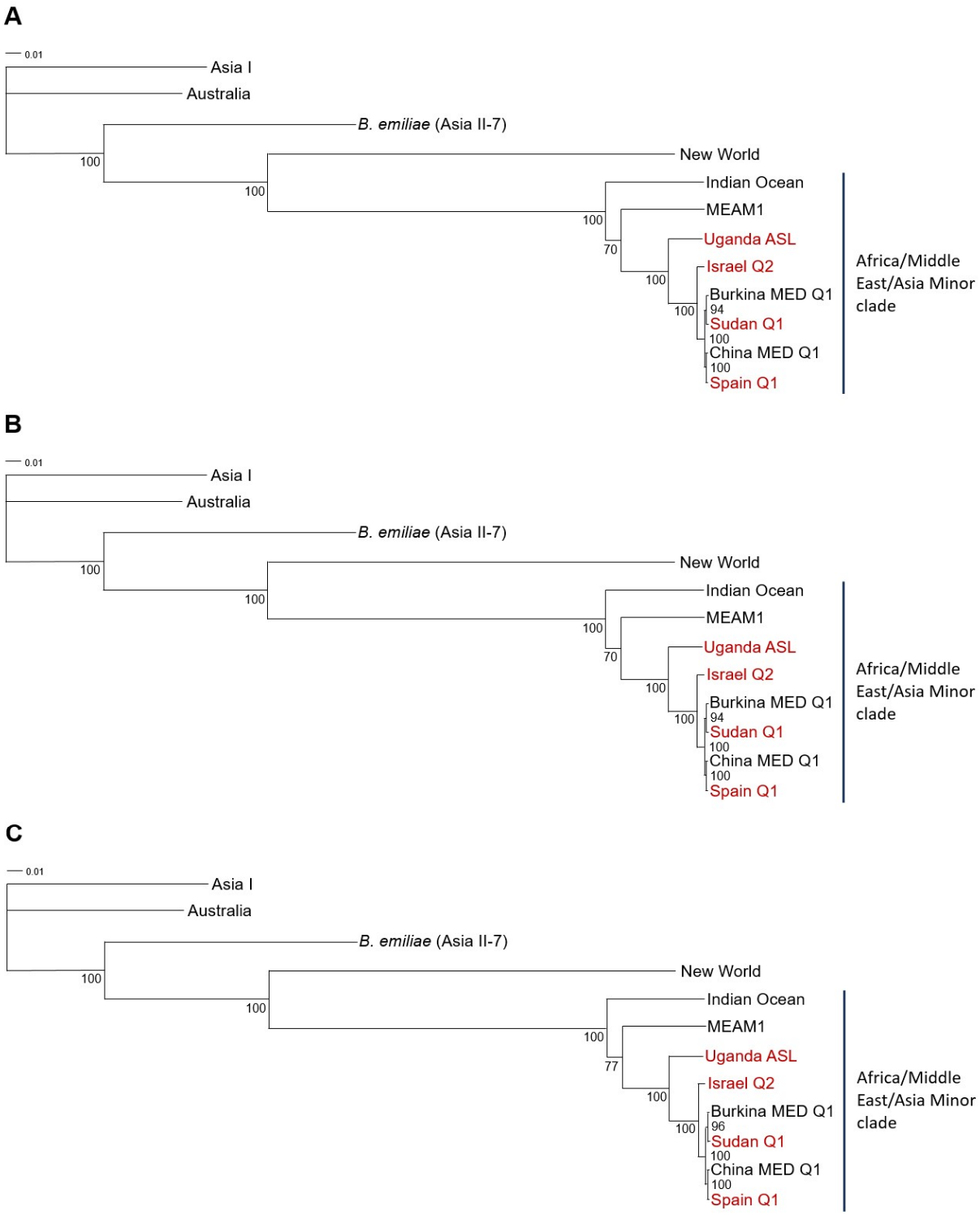


Figure 3-5: Unrooted maximum likelihood phylogenetic trees inferred from 12,595 bp concatenations of 15 mitochondrial genes from 12 *B. tabaci* populations using **A:** no partitioning of data, **B:** partitioning into 1st +2nd codons, 3rd codons and rDNA genes, **C:** partitioning into 15 individual genes. Values near nodes show statistical support of Ultrafast Bootstrap. Names in red mark samples generated in this study.

3.3.4 Nucleotide sequence distances within mitochondrial genes across *B. tabaci* populations

Pairwise nucleotide sequence distances were compared among 12 samples of the *B. tabaci* species complex for four regions of the mitogenome: (i) 3' mtCOI partial sequence of mtCOI, (ii) 5' partial sequence of mtCOI (iii) full mtCOI gene sequence and (iv) concatenation of 15 mitochondrial genes (Table 3-4).

The distances within the Africa/Middle East/Asia Minor clade ranged from 0 to 9.44%. Within the Q1, Q2 and ASL groups, the distance range was 0–4%. Across the Q1 and Q2 samples only, the distances remained low in all four regions (0–1.37%), however, the distance of Uganda ASL from these Q1 and Q2 groups ranged from 1.98 to 4% (Table 3-4).

Among all 12 populations, the 3' partial mtCOI sequence showed the lowest levels of nucleotide divergence on average (9.95%). The 5' partial mtCOI sequence and full mtCOI sequence had similar average divergences (10.22% and 10.10%, respectively). In contrast, the average distance among the 15 mitochondrial genes was 13.27%.

Table 3-4: mtDNA nucleotide sequence distances (%) among 12 populations of the *B. tabaci* complex in four regions of the mitogenome. The frame encompasses the Africa/Middle East/Asia Minor clade. Dark grey boxes highlight the divergences between Uganda ASL and MED populations, while divergences within Q1 and Q2 groups of MED are highlighted with light grey. Names with asterisks mark samples sequenced in this study.

3' mtCOI (657 bp)

	1	2	3	4	5	6	7	8	9	10	11
1 Asia I	-										
2 Australia I	10.81	-									
3 *B. emiliae*	12.33	13.85	-								
4 New World	14.16	16.59	16.13	-							
5 Indian Ocean	13.24	15.53	14.92	15.68	-						
6 MEAM1	14.16	14.31	15.22	14.46	7.00	-					
7 Uganda ASL *	14.61	15.83	14.31	15.22	6.54	4.72	-				
8 Israel Q2 *	14.00	15.98	14.76	14.61	6.09	4.41	1.98	-			
9 Burkina MED Q1	14.76	16.44	14.92	14.76	6.24	5.02	2.74	1.07	-		
10 China MED Q1	14.76	16.44	15.22	15.07	6.54	5.02	2.74	1.07	0.30	-	
11 Spain Q1 *	14.61	16.29	15.07	14.92	6.39	4.87	2.59	0.91	0.15	0.15	-
12 Sudan Q1 *	14.76	16.44	14.92	14.76	6.24	5.02	2.74	1.07	0.00	0.30	0.15

5' mtCOI (657 bp)

	1	2	3	4	5	6	7	8	9	10	11
1 Asia I	-										
2 Australia I	12.48	-									
3 *B. emiliae*	13.24	14.46	-								
4 New World	14.61	15.53	15.22	-							
5 Indian Ocean	15.07	14.46	15.53	14.76	-						
6 MEAM1	14.46	14.31	15.53	14.00	9.44	-					
7 Uganda ASL *	13.39	13.09	14.61	14.92	8.98	8.07	-				
8 Israel Q2 *	14.00	13.85	14.31	13.85	8.07	8.22	3.81	-			
9 Burkina MED Q1	14.16	14.00	15.07	13.85	7.76	7.91	3.81	1.22	-		
10 China MED Q1	14.00	14.00	15.22	13.7	8.22	8.07	3.96	1.37	0.46	-	
11 Spain Q1 *	14.16	14.00	15.07	13.55	7.76	7.61	3.81	1.22	0.30	0.46	-
12 Sudan Q1 *	14.00	13.85	14.92	13.70	7.91	7.76	3.65	1.07	0.15	0.30	0.15

Full mtCOI (1,537 bp)

	1	2	3	4	5	6	7	8	9	10	11
1 Asia I	-										
2 Australia I	11.65	-									
3 *B. emiliae*	13.01	14.18	-								
4 New World	14.38	15.81	15.94	-							
5 Indian Ocean	14.38	14.90	15.22	15.09	-						
6 MEAM1	14.38	14.05	15.55	14.25	7.81	-					
7 Uganda ASL *	13.92	14.44	14.64	14.83	7.81	6.57	-				
8 Israel Q2 *	14.18	14.77	14.77	13.92	7.03	6.44	3.25	-			
9 Burkina MED Q1	14.51	15.09	15.09	13.92	7.03	6.64	3.64	1.04	-		
10 China MED Q1	14.44	15.09	15.29	13.99	7.35	6.70	3.71	1.11	0.33	-	
11 Spain Q1 *	14.44	15.03	15.16	13.86	7.09	6.44	3.58	0.98	0.20	0.26	-
12 Sudan Q1 *	14.44	15.03	15.03	13.86	7.09	6.57	3.58	1.04	0.13	0.33	0.20

Concatenations (12,595 bp)

	1	2	3	4	5	6	7	8	9	10	11
1 Asia I	-										
2 Australia I	14.00	-									
3 *B. emiliae*	16.45	16.72	-								
4 New World	19.96	20.3	19.88	-							
5 Indian Ocean	19.83	20.37	19.75	19.67	-						
6 MEAM1	19.77	19.91	19.47	19.33	8.74	-					
7 Uganda ASL *	19.72	20.23	19.36	19.34	8.62	7.73	-				
8 Israel Q2 *	19.84	20.38	19.57	19.44	8.80	7.79	3.91	-			
9 Burkina MED Q1	19.98	20.43	19.66	19.34	8.62	7.88	3.91	1.09	-		
10 China MED Q1	20.02	20.44	19.72	19.40	8.71	7.95	3.94	1.07	0.37	-	
11 Spain Q1 *	19.97	20.39	19.65	19.36	8.66	7.91	3.93	1.06	0.35	0.15	-
12 Sudan Q1 *	20.02	20.48	19.69	19.45	8.71	7.93	4.00	1.17	0.37	0.47	0.45

The correlation between nucleotide distances in partial mitochondrial regions and 15 mitochondrial genes was high (>0.9) and the significance level was 0.001 in all comparisons (Table 3-5). The correlation between distances in 3' and 5' partial mtCOI sequences was the lowest, beneath the correlation between partial mtCOI regions and tRNA genes. The highest correlation occurred between distances in the full mtCOI gene and the 15 mitochondrial genes. The 5' partial mtCOI sequence best represented full mtCOI sequence, while the 3' partial mtCOI best corresponded to the 15 mitochondrial genes.

Table 3-5: Mantel test comparing nucleotide distance matrices among two partial mtCOI sequences (657 bp), full mtCOI (1,537 bp), concatenation of 11 tRNA genes (657 bp) and 15 mitochondrial genes (12,595 bp) from 12 *B. tabaci* populations.

	3' mtCOI	**5' mtCOI**	**full mtCOI**	**tRNAs**	**15 genes**
3' mtCOI	1				
5' mtCOI	0.9735	1			
full mtCOI	0.9923	0.9935	1		
tRNAs	0.9781	0.9766	0.9852	1	
15 genes	0.9948	0.9825	0.9951	0.9880	1

3.3.5 mtCOI nucleotide divergence within the Africa/Middle East/Asia Minor clade

Nucleotide sequence divergence within the mtCOI gene was studied in greater detail among members of the Africa/Middle East/Asia Minor clade (Dinsdale *et al.*, 2010). A sliding window analysis was performed in a 657 bp window across the full length of the gene (1,537 bp) and the distances for each of the seven pairwise comparisons were plotted.

The distribution of sequence divergence was not uniform across the gene length. The 5' half of the mtCOI gene showed higher level of nucleotide divergence among members of the clade compared with the 3' half (Figure 3-6). Maximum values detected across all seven comparisons occurred in windows positioned between 81–737 bp and 351–1007 bp, while minimum distances occurred in 526–1188 bp and 886–1542 bp (Appendix 2).

The nucleotide distance between ASL and Q1+Q2 groups ranged from 2.48 to 4.77% (3.56% on average), depending on the chosen region of mtCOI gene (Appendix 2). Thus, the divergence between the ASL population and other MEDs reached values 1.02%

below and 1.27% above the species delimitation threshold of 3.5% identified for the *B. tabaci* species complex (Dinsdale *et al.*, 2010). In contrast, the divergence between Q1 and Q2 populations showed a smaller variation in sequence divergence (0.61–1.38% depending on the window position; 1.01% on average).

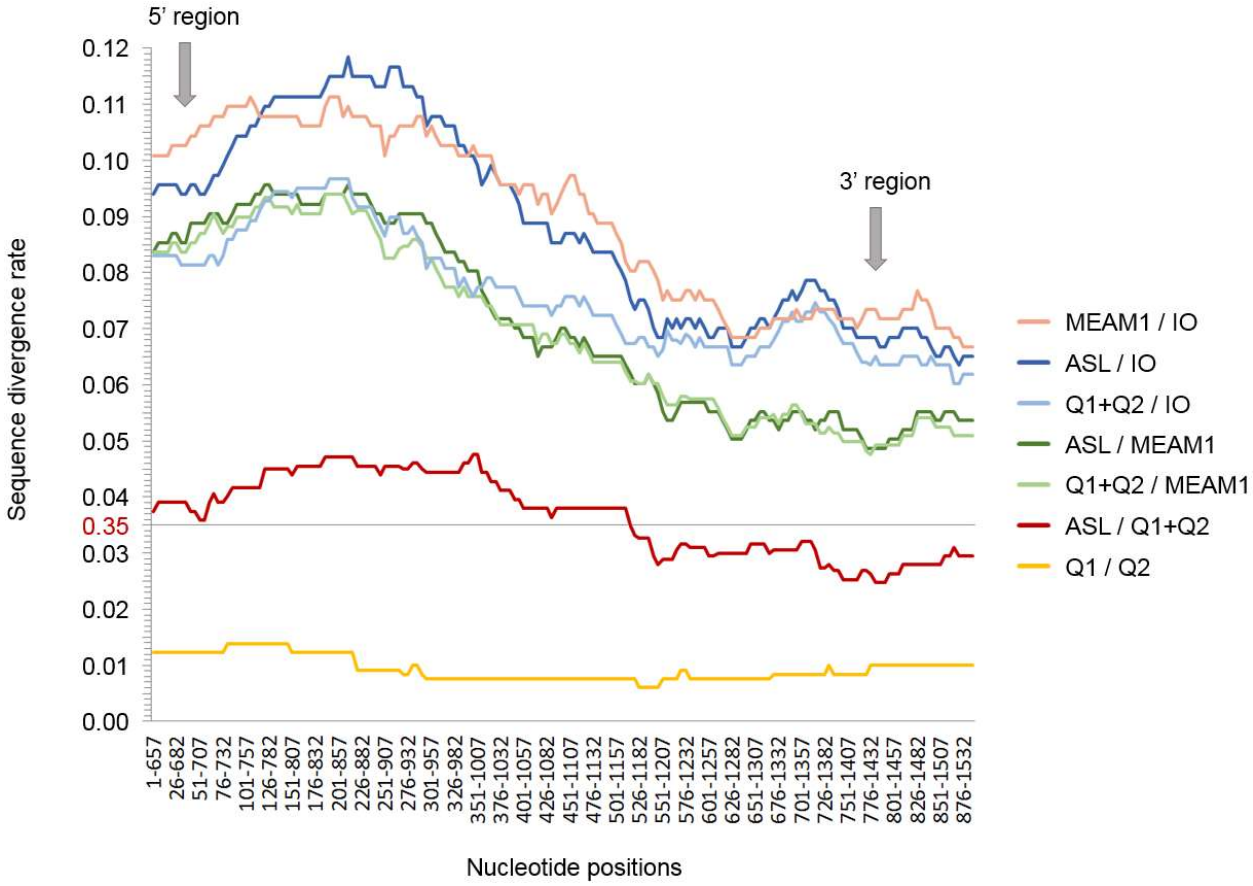


Figure 3-6: Pairwise sequence divergences across the full length of mtCOI gene from whiteflies of the Africa/Middle East/Asia Minor clade. The grey horizontal line marks the commonly applied 3.5% species threshold (Dinsdale *et al.*, 2010). The Q1 consists of two sequences (Spain Q1 and Burkina Faso Q1). Grey arrows indicate the position of 5′ and 3′ barcoding regions widely used to assist with species identification across animal taxa.

3.3.6 Pseudogene identification

Pooling the MED accessions from three publications and reducing them to unique haplotypes revealed the relationships between naming systems of different authors (Appendix 3). As expected, MedBasin1 and MedBasin2 (Ahmed *et al.*, 2009) corresponded respectively to Q1 and Q2 (Gueguen *et al.*, 2010; Chu *et al.*, 2012a). However, Q3 from Burkina Faso (Gueguen *et al.*, 2010) corresponded to Q5 (Chu *et al.*, 2012a), while Q3 from Croatia (Chu *et al.*, 2012a) was unique. Finally, sequences classified as ASL (Gueguen *et al.*, 2010) clustered with Q4 sequences (Chu *et al.*, 2012a).

Four criteria were applied in analysing published 3' partial mtCOI sequences of five MED groups in order to detect potential pseudogene sequences. The presence of INDELs (criterion i) and significant amino acid changes (criterion iv) occurred in sequences from all five MED groups. No premature stop codons (criterion ii) were detected, with the exception of clusters of non-synonymous mutations (criterion iii) found in Q1 and Q2 haplotypes. In total, 155/289 (53.6%) of the analysed haplotypes contained one or more errors (Table 3-6; Appendix 3).

Table 3-6: Summary of the analysis of 3' partial mtCOI MED sequences from GenBank against HTS-derived mitogenomes from this study. Numbers of haplotypes failing each criterion are shown in columns. Criterion i = INDELs, ii = STOP codons, iii = anomalous polymorphisms, iv = non-synonymous mutations.

MED group	Haplotypes	i	ii	iii	iv[a]	Potential NUMTs	
						Total	%
Q1[b,c]	210	67	0	23	47	114	54.3%
Q2[b,c]	36	9	0	3	10	19	52.8%
Q3[c]	2	0	0	0	1	1	50.0%
Q3[b] (=Q5[c])	8	1	0	0	3	4	50.0%
ASL[b] (=Q4[c])	33	7	0	0	10	17	51.5%

[a] Only sequences without INDELs were analysed at the amino acid level. Only amino acid changes in sites conserved across all 9 reference *Bemisia* sequences were considered.
[b] Naming system of Gueguen *et al.* (2010).
[c] Naming system of Chu *et al.* (2012a).

Within the Q1 dataset, 67 (31.6%) of the 210 unique haplotypes failed criterion i, out of which 12 overlapped with criterion iii. A particularly common INDEL was the deletion of ''A'' at position 750 or 751 that occurred in 17 and 25 haplotypes, respectively. Criterion iii excluded 23 Q1 haplotypes, out of which 11 contained 9–116 bp long stretches of

non-synonymous mutations at the beginning or the end of the sequence (Appendix 4) and another 12 were identical to accessions from a different *Bemisia* species or chloroplast genome sequences in BLASTn search (Appendix 3). A total of 47 (22.2%) Q1 haplotypes failed criterion iv based on the alignment with a reference set of mtCOI protein sequences, which revealed amino acid changes in positions that were fully conserved across the *B. tabaci* complex and the *Bemisia* "JpL" species.

The Q2 dataset also included haplotypes with INDELs (9/37, 24.3%), failing criterion i. Three Q2 haplotypes were detected by criterion iii, which contained clusters with high density of non-synonymous mutations (20–21 bp on 5' end and 87–89 bp at 3' end; Appendix 5), including multiple INDELs. In addition, ten haplotypes (27.0%) showed non-synonymous substitutions resulting in amino acid changes in positions that were conserved across the reference set of protein sequences, failing criterion iv.

From the two haplotypes from Croatia, called Q3 in Chu *et al.* (2012a), one contained a non-synonymous mutation in a conserved site. The group of haplotypes called Q3 (Gueguen *et al.*, 2010) or Q5 (Chu *et al.*, 2012a), all from Burkina Faso, contained one haplotype with an INDEL and three haplotypes with one significant amino acid change. Finally, within the ASL (Gueguen *et al.*, 2010) or Q4 (Chu *et al.*, 2012a) group, 7/33 (21.2%) haplotypes contained INDELs and additional 10/33 (30.3%) haplotypes showed amino acid changes in conserved sites.

3.3.7 Comparison of Q3 and ASL mtCOI haplotypes to similar populations

Comparison of nucleotide distances in the 3' partial mtCOI region between four Q3 haplotypes from Burkina Faso (Gueguen *et al.*, 2010) to our mitogenome haplotypes revealed that they were most closely related to Israel Q2 (0.46–1.07%) (Table 3-7). The Q3 haplotype from Croatia (Chu *et al.*, 2012a) showed equally low nucleotide distance of 0.46% from both Israel Q2 and Spain Q1 sequences. All five Q3 haplotypes were the most distant from the Uganda ASL sequence (1.83–2.13%). The Croatia Q3 haplotype was 1.22–1.52% divergent from Q3 haplotypes from Burkina Faso.

Table 3-7: Nucleotide distances (%) in the 3' partial mtCOI sequence (657 bp) among haplotypes of Q3 populations from Burkina Faso (Gueguen *et al.*, 2010) and Croatia (Chu *et al.*, 2012a) and the four populations from this study. Only haplotypes without identified errors were used in this comparison. The inner frame marks comparisons between our samples and the Q3 haplotypes. Grey cells highlight the lowest distance for each Q3 haplotype within that frame.

	1	2	3	4	5	6	7	8	9
1 Burkina Q3 FJ766385	-								
2 Burkina Q3 FJ766400	0.15	-							
3 Burkina Q3 FJ766405	0.30	0.15	-						
4 Burkina Q3 FJ766382	0.46	0.3	0.15	-					
5 Croatia Q3 GU086335	1.37	1.22	1.37	1.52	-				
6 Uganda ASL	1.98	1.83	1.98	2.13	2.13	-			
7 Spain Q1	1.83	1.67	1.83	1.98	0.46	2.59	-		
8 Sudan Q1	1.98	1.83	1.98	2.13	0.61	2.74	0.15	-	
9 Israel Q2	0.91	0.76	0.91	1.07	0.46	1.98	0.91	1.07	-

The Uganda ASL partial mtCOI sequence was compared to 27 published sequences of related sub-Saharan African populations; however, after checking the sequence quality against the four criteria for identifying pseudogenes and errors (section 3.2.9), 14 haplotypes were not considered further. The remaining 13 haplotypes were 0–1.13% distant from Uganda ASL (Table 3-8). These 13 haplotypes were obtained from samples in Uganda (Sseruwagi *et al.*, 2005), but also in West Africa in Ghana (Omondi *et al.*, 2005; De la Rúa *et al.*, 2006).

Table 3-8: Sequence divergence in the 3' partial mtCOI sequence (708 bp) between Uganda ASL population and published sequences of "African-silverleafing", "Ug4" or "okra" biotypes. The last column contains INDELs (orange) orr amino acid changes (purple) detected in comparison with reference mtCOI from Uganda ASL mitogenome. The number refers to the nucleotide position of the full mtCOI gene.

Sequence	Reference	Distance from Uganda ASL	INDELs and amino acid substitutions
Nigeria AY827606	De la Rúa *et al.* (2006), Boykin *et al.* (2007)	0.14%	775 E->K[a]
Ghana AY827582		2.13%	764 L->C; 775 E->K; 791 E->A; 817 Y->M; 1415 L->S; 1487 N->I; 1491 K->N
Ghana AY827579		1.13%	
Ghana AY827580		1.28%	1250 F->C; 1407 F->L; 1476 E->D
Ghana AY827581		1.42%	1250 F->C
Cameroon AF344258	Berry *et al.* (2004), Boykin *et al.* (2007)	1.56%	366 del T; 1357 del T; 1396 del T
Zimbabwe AF344285		0.43%	366 del T; 1387 del T
Zimbabwe AF344286		0.99%	366 del T; 1226 del T; 1465 del A; 1509 ins C
Ghana AY827588	Omondi *et al.* (2005)	1.13%	
Ghana AY827590		1.13%	
Ghana AY827587		0.71%	
Ghana AY827589		0.99%	
Uganda AY903573	Sseruwagi *et al.* (2005)	0.57%	1003 P->A; 1361 S->F
Uganda AY903549		0.42%	950 ins T
Uganda AY903551		0.14%	1395 I->M; 1532 F->S
Uganda AY903531		0.14%	
Uganda AY903532		0%	
Uganda AY903545		0%	
Uganda AY903546		0%	
Uganda AY903571		0.28%	952 del A; 1158 del T
Uganda AY903550		0.14%	943 ins A
Uganda AY903533		0%	
Uganda AY903549		0%	
Uganda AY903572		0%	
Uganda AY903574		0%	
Uganda AY903534		0.99%	1542 del T
Uganda AY903535		1.27%	1083 ins C

Ins = insertion, del = deletion.
[a] Indication of a non-synonymous mutation at the position shown, resulting in amino acid change (IUPAC one-letter codes)

3.4 Discussion

Findings presented in this chapter have contributed to our understanding of the evolutionary relationships within a subset of the *B. tabaci* species complex. A reconstruction of the phylogenetic tree and measuring the nucleotide distances among 12 members of the complex answered our first research question about the evolutionary relationships of the MED groups. These molecular analyses revealed that the ASL population is genetically distinct from MED groups Q1 and Q2. The second research question was addressed by comparing the nucleotide divergence among several regions of the mitogenome and quality checking of the published MED 3' partial mtCOI sequences, which revealed inconsistencies in the barcoding results and inaccuracies in the publicly available sequences.

Results from our study provided examples of all four problematic issues with DNA barcoding, as outlined in the introduction. Firstly, there was an overlap between the intraspecific and, what was proposed to be, interspecific nucleotide distance. This could be observed in the comparison of ASL and Q1+Q2 groups, distance of which fluctuated above and below the 3.5% species threshold. Secondly, the arbitrary choice between using the 3' or 5' partial mtCOI sequence for species assignment would result in two different classifications. Nucleotide divergence in the 5' region would place the ASL population outside MED species, however, choosing the 3' region resulted in its formal inclusion within MED (Dinsdale *et al.*, 2010). Thirdly, the scale of erroneous and misidentified 3' partial mtCOI sequences published by previous studies, in which the intraspecific phylogenetic relationships of MED were inferred, became evident. Over 53% of the 289 unique haplotypes contained errors, or were misidentified as MED despite originating from a different species or organism. Lastly, a considerable portion of these identified erroneous sequences probably originated from nuclear pseudogene sequences instead of mtDNA (Lopez *et al.*, 1994).

The Q3 population from Croatia (Chu *et al.*, 2012a) might actually be part of Q1 or Q2, based on the low nucleotide distance (0.46%) of its 3' partial mtCOI sequence from Spain Q1 and Israel Q2 sequences. Moreover, as the identification of this group was based on only two haplotypes (Chu *et al.*, 2012a), out of which one contained errors (Table 3-6), it is possible that this group is an artefact resulting from the intra-species genetic variability of MED and/or the inclusion of NUMT sequences. In contrast, the Q3 haplotypes of samples from Burkina Faso were the closest to Israel Q2 haplotype (up to 1.07%) (Table 3-7), which is about the same level of divergence as between Israel Q2 and Spain Q1

(Table 3-4). More molecular data and a mating study would be needed to elucidate the relationships between this Q3 and the other MED groups.

The genetic distinctness of the ASL population suggests that it might be a separate species, as it was placed outside the MED cluster with 100% statistical support. The tree also supported the consensus that Indian Ocean occupies a basal phylogenetic position within the Africa/Middle East/Asia Minor clade (Dinsdale *et al.*, 2010; De Barro *et al.*, 2011; Lee *et al.*, 2013). The species status of ASL should, however, be further confirmed in crossing trials (Chapter 4) and biological experiments (Chapter 5). The low nucleotide distance of 0–1.13% (Table 3-8) in 3' partial mtCOI sequences among samples of "African silver-leafing" (Boykin *et al.*, 2007), "okra biotype" (Omondi *et al.*, 2005) and "Ug4" (Sseruwagi *et al.*, 2005) indicates that these groups might belong to the same species, but it remains to be confirmed. Specifically, reproductive compatibility of ASL populations from East and West Africa should be assessed.

The ''Sub-Saharan Africa silver-leafing" clade was originally placed separately from the "Mediterranean invasive" clade in the first global Bayesian phylogenetic study (Boykin *et al.*, 2007), but these two clades were later merged to form one "Mediterranean'' putative species (Dinsdale *et al.*, 2010). This placement was in accordance with the 3.5% species-level threshold in the 3' barcoding region of mtCOI. We show here, however, that the distance values varied considerably across the full length of mtCOI gene (Figure 3-5). Results from the sliding window analysis showed that the sequence divergence between ASL and Q1+Q2 populations in a 657 bp long window ranged from 2.48 to 4.77% (a difference of 2.29%), depending on the region chosen for barcoding. Such variability was not apparent between Q1 and Q2, divergence of which remained low across the full gene length (0.61–1.38%, difference of 0.77%). A similar pattern was observed in the comparison at the mitogenome scale (Table 3-4).

The correlation between nucleotide distances measured in partial mitochondrial sequences and across 15 genes was high, however, there were differences in how well the commonly used 3' and 5' partial mtCOI regions represent the larger-scale mitogenome distances. Results from our analysis using 12 *Bemisia* populations suggested that the 5' partial mtCOI sequence represented well the full mtCOI sequence, while the 3' partial mtCOI better corresponded to the 15 mitochondrial genes. The distances in 3' and 5' partial regions had the lowest correlation coefficient, even lower than their comparison to tRNA sequences, which can lead to discordant results. It is, therefore, useful to study multiple genome regions to make analyses more robust to such discrepancies.

The importance of understanding the cryptic species diversity in the *B. tabaci* complex is three-fold. In practical terms, it is paramount for devising efficient pest management strategies and biosecurity measures. Secondly, it brings more consistency and clarity to communication within the scientific community, as well as with the wider public. Lastly, our findings emphasize the simplistic and inaccurate approach to species delimitation that is currently in place. DNA barcoding is an invaluable method for fast classification of field samples, however, the large (and sometimes exclusive) role it plays in species delimitation has been criticised (Collins and Cruickshank, 2013), especially for closely related species (Meyer and Paulay, 2005).

The issue of over-reliance on DNA barcoding sequences is further exacerbated by the volume of faulty sequences submitted to public repositories. Our analysis of 289 unique mtCOI haplotypes in GenBank assigned to MED subclades Q1–Q5 and ASL (Chu *et al.*, 2008; Ahmed *et al.*, 2009; Gueguen *et al.*, 2010; Chu *et al.*, 2012a), revealed the extent of NUMTs or erroneous sequences included in previous molecular studies. In all five subclades, a half or more (50–54%) of the unique haplotypes contained mutations that are unlikely to occur in a functional protein-coding sequence. It is important to note that the FaBox tool assigned identical sequences with different lengths as unique haplotypes. Thus, the percentages in Table 3-5 probably do not reflect the frequency with which pseudogenes are picked up by the primers.

The inclusion of NUMT sequences in phylogenetic analyses can overestimate the species richness (Song *et al.*, 2008) and confound phylogenetic analyses (van der Kuyl *et al.*, 1995; Thalmann *et al.*, 2004), as well as create confusion and inconclusive results in the research community. Pseudogene sequences have recently been shown to cause an overestimation of species richness within the Africa/Middle East/Asia Minor clade (the case of MEAM2; Tay *et al.*, 2017a). This indicates that unrecognised NUMT sequences might commonly occur across multiple clades of the *B. tabaci* species complex and interfere with efforts to resolve its systematics. It is therefore important to recognise this issue and to analyse critically the quality of putative mitochondrial sequences amplified with universal primers. The pseudogene sequences are not immediately obvious if they do not contain stop codons or significant amino-acid changes within the analysed partial sequence. We emphasise the utility of HTS and acquiring mitochondrial sequences by assembly from a large number of reads. Such sequences can then be used as a reference for all the Sanger-sequenced marker sequences (e.g. from a field collection), to identify sequencing errors and potential pseudogenes and exclude them from downstream analyses and from submitting to public databases.

A limitation of our study, and of the literature in general, is that different methods for calculating the nucleotide distances have been used. The 3.5% threshold (Dinsdale *et al.*, 2010) was identified based on distances corrected with the Hasegawa-Kishino-Yano 85 (HKY85) substitution model, which does not assume equal DNA base frequencies and accounts for the difference between transitions (purine to purine or pyrimidine to pyrimidine base) and transversions (purine to pyrimidine base or *vice versa*), with each site changing independently of others (Hasegawa *et al.*, 1985). Some of the subsequent studies identifying new putative species followed this method (e.g. Parrella *et al.*, 2012). Other commonly used model is the Kimura's two parameter (K2P) that also weights transitions and transversions differently, but assumes equal base frequencies and equal rates across sites (Kimura, 1980). This model was used in Hu *et al.* (2011), Esterhuizen *et al.* (2013), and Hu *et al.* (2014; 2017). As different models provide different corrections, the resulting values of nucleotide distances can vary (Collins *et al.*, 2012). This is evident in our study as well, as Table 3-4 presents uncorrected *p* distances, but the output of the sliding window analysis were distances corrected with the JC model that assumes an equality of substitution rates among sites, equal nucleotide frequencies and the same rate of transitions and transversions. This explains the slightly different values resulting from these two methods after comparing identical sequences. Both *p* distance and the JC model were, however, shown to be more accurate measures of nucleotide distances in DNA barcoding than the K2P model (Collins *et al.*, 2012; Srivathsan and Meier, 2012).

Another limitation of our results stems from using very short reads (150 bp) for sequencing total whitefly DNA. The difficulties of assembling repetitive regions with sequences generated in a shotgun sequencing approach is well known (Treangen and Salzberg, 2011). It is, therefore, not possible at this point to determine whether the differences in length of the low complexity region between *cox3* and *tRNA-Ile* are genuine or whether they are an artefact generated through the assembly from short reads. A long PCR with primers specific to regions inside *cox3* and *tRNA-Ile* genes would be needed to verify the sequence in this region of the mitogenome. On the other hand, the range of length differences was only 1–4 bp and any intergenic sequences were not included in our analysis, so the short reads assembly should not have an effect on our phylogeny and nucleotide distances. The mapping of our reads without gaps, the correct translation of protein-coding sequences and the good alignment of our sequence with other published mitogenomes indicates a good overall quality of our mitogenomes.

By using HTS-derived mitogenomes, we were able to study the genetic relationships among the MED populations and several other *Bemisia* species using true mitochondrial

sequences (*i.e.* no pseudogenes) and multiple genes (*i.e.* not only mtCOI). This approach contrasts previous studies inferring phylogenetic relationships and sequence divergence among MED populations from 3' partial mtCOI sequences, out of which about 54% contain errors and potentially are NUMTs. The scientific community would benefit from clearing sequence databases of errors and by applying higher standards of sequence quality control prior to sequence submission.

The research in the future should also move towards using multiple genes, and ideally genome-wide data, in the phylogenetic analyses. A greater integration of nuclear and mitochondrial genomic studies is necessary to further our understanding of insect genome evolution (Cameron, 2014). The short mtDNA marker is a useful proxy, but mtDNA is known to often not represent the full phylogenetic history (Collins and Cruickshank, 2013) and so the genetic divergences in mtDNA do not necessarily track species boundaries (Avise, 2004). With the increasing ease with which genome-wide SNPs can be obtained, combining nucleotide polymorphisms from both mitochondrial and nuclear genomes (e.g. Elfekih *et al.*, 2018) will offer greater power for inference of *B. tabaci* cryptic species phylogenetic relationships.

4. Reproductive compatibility among MED populations and their bacterial endosymbionts[2]

4.1 Introduction

Data presented in Chapter 3 showed that the ASL population was genetically more distant from MED populations Q1 and Q2 than were Q1 and Q2 from each other. This finding raised a question as to whether ASL really belongs to the same biological species as the Q1 and Q2 groups.

Species within the *B. tabaci* complex are cryptic and their morphology is plastic depending on the host-plant that they develop on (Russell, 1957; Mound, 1963), so the traditional morphology-driven taxonomy and species designations are not possible. DNA barcoding is a useful substitute method for specimen classification (demonstrated in section 3.3.1), however, it is not always sufficient for species delimitation (Collins and Cruickshank, 2013). A more comprehensive study is needed to assess the species status of the populations of interest. This is particularly the case for pests, because the adequate species identification and description of pest species are crucial for biosecurity and pest management.

Is has been suggested that a combined approach of molecular phylogenetics and reproductive compatibility can provide an adequate resolution of the *B. tabaci* species complex (Liu *et al.*, 2012b). In the present chapter, therefore, the species status of the ASL, Q1 and Q2 groups was investigated by applying the biological species concept and assessing the reproductive compatibility among them by conducting reciprocal crossing experiments. An additional reason to study the gene flow among various whitefly populations is that such knowledge has implications for global biosecurity (De Barro and Hart, 2000; Boykin *et al.*, 2012), as interbreeding could potentially lead to an exchange of genes conferring favourable phenotypes, such as insecticide resistance.

The reproductive relationships within MED have been studied to a limited extent by microsatellite analyses of sympatric field populations. Three such studies carried out across five Mediterranean countries and La Réunion suggested that Q1 and Q2 were reproductively compatible (Gauthier *et al.*, 2014; Terraz *et al.*, 2014; Thierry *et al.*, 2015). In contrast, no evidence of gene flow was detected between sympatric field populations of Q1 and ASL in Burkina Faso (Mouton *et al.*, 2015). However, the model-based

[2] This chapter formed a part of publication Vyskočilová *et al*. (2018); page 231.

Bayesian methods for detecting hybrids implemented in these microsatellite field studies have a limitation because they are biased towards the detection of F_1 hybrids (Vähä and Primmer, 2006). To ascertain the reproductive compatibility among populations, it is important to study the fertility of F_1 hybrids by testing their ability to produce F_2 offspring (Xu *et al.*, 2010).

Another limitation in the published literature on mating compatibility among members of the *B. tabaci* complex has been the lack of rigour in the verification of hybrid progeny. A lack of hybrid confirmation can lead to false positive (type I) errors. In some studies, the identity of F_1 females was only studied by sequencing their partial mtCOI sequence (e.g. Qin *et al.*, 2016), which can indicate exclusively the maternal parentage, or the genetic identity was not checked at all. In the remaining cases, the genetic make-up of the F_1 females was verified by RAPD (De Barro and Hart, 2000; Omondi *et al.*, 2005; Xu *et al.*, 2010; Sun *et al.*, 2011; Wang *et al.*, 2011a). This method does have the capacity to detect both maternal and paternal genetic material, however, it is not always efficient due to partial failures in the PCR amplification, co-migration of different fragments with similar length and low reproducibility of the results (Caterino *et al.*, 2000; Xu *et al.*, 2010).

The aim of this chapter was to fill these gaps by investigating the reproductive compatibility among MED populations in laboratory crossing experiments coupled with a specific molecular marker for hybrid confirmation, as well as to formulate hypotheses about the potential effect of endosymbiotic bacteria on these reproductive relationships. The research objectives of this chapter were to: (1) study the mating compatibility among the MED populations in reciprocal crossing experiments, (2) verify the fertility of hybrids, (3) develop a nuclear marker and confirm the parental origin of the F_1 hybrids, and (4) identify the bacterial endosymbionts of each experimental population.

4.2 Material and methods

4.2.1 Growing plants and rearing insects

Aubergine plants *S. melongena* 'Black Beauty' were grown and *B. tabaci* colonies were reared as detailed in section 3.2.1.

4.2.2 Reciprocal crossing experiments

All crossing experiments were conducted on aubergine plants at the three to seven true-leaf stage, rooted in soil and enclosed in Lock&Lock cages. At least 24 h before introducing the adults, all leaves, except a fully expanded one, were removed to facilitate the contact between mating partners.

The virginity of females in all crossing experiments was ensured by their emergence in isolation. One to two days prior to the experiment, leaves carrying whitefly nymphs were taken from the core colony. The 4th instar nymphs with red eyes (Figure 2-5 B) were cut from the leaf together with the piece of supporting leaf and placed individually in glass tubes (0.5 cm in diameter, 6.0 cm long) secured with a piece of cotton wool. If the position of two nymphs was too close to be separated, they were placed in a tube together. Adults emerged from such samples were only used if they were of the same sex. The sex was determined by examining the shape of their abdomen under a stereomicroscope (Figure 4-1).

Figure 4-1: Male (left) and female (right) of *B. tabaci* (Uganda ASL population).

Each replicate consisted of one virgin female and three virgin males. Survival of these adults was monitored periodically and deceased parental males were replaced by new males from the respective core colony. In the case of female death, the replicate was discarded.

In the control crosses the female and males originated from the same population. Reciprocal crosses were performed with females and males from two different populations. The parental adults were collected after a seven-day mating period and stored in 90% ethanol at -20^{0}C. Adults of the F_1 generation were collected until all emerged. Their sex was determined and any F_1 females were stored in 90% ethanol at -20^{0}C for subsequent molecular analysis.

4.2.3 Back-crossing F_1 hybrids to test their fertility

A separate set of crossing experiments was performed to test the fertility of F_1 hybrids. To obtain virgin hybrid F_1 females, a group mating cage was set up with 20 virgin parental females and 20 males from each respective colony in both directions for the period of 14 days. F_1 nymphs were then harvested for their emergence in isolation as described in section 4.2.2. Virgin F_1 females from both directions of the group mating were then back-crossed with males of either parental type, resulting in four types of back-crosses. In each replicate of the back-crosses, one female and three males were used and the rest of the experiment performed as described in 4.2.2.

4.2.4 Statistical analyses

Statistical analyses were performed using R (www.R-project.org; R Core Team, 2013). Counts of offspring generated in crossing experiments were analysed by a generalised linear model with negative binomial error distribution and a log link using the MASS library (Venables and Ripley, 2002). For the proportion of female progeny, a generalized linear model with quasibinomial error distribution and logit link of the proportional data was used. Multiple comparisons of offspring counts and female proportions were performed by Tukey's test (Tukey, 1949) using the multcomp package (Hothorn *et al.*, 2008) and significant differences were demonstrated by compact letter display. Reciprocal crosses that produced only male offspring were omitted from the multiple comparison. A separate analysis was also carried out with results from the F_1 and F_2 generations relevant only to the reproductively compatible populations.

4.2.5 Selection of a diagnostic nuclear marker

A molecular marker was developed for verifying the hybrid origin of F_1 females and thus ruling out false positive results. It was necessary to identify a marker from nuclear DNA, because mitochondrial DNA has a maternal mode of inheritance that renders it ineffective for discerning the paternal origin of F_1 females.

A marker suitable for distinguishing between the closely related MED populations needed to contain sufficient intraspecies variability. Therefore, we targeted nuclear genes encoding mitochondrial proteins, because of their potentially faster mutation rate (Gershoni *et al.*, 2009). Candidate genes (Appendix 6) were extracted from transcriptome sequences of Spain Q1 and Israel Q2 adults (Patel *et al.*, unpublished data) based on *Drosophila melanogaster* homologues (www.flybase.org). The gene sequences were aligned and regions containing polymorphisms were identified. Four domains in three genes (NADH-Q-oxidoreductase 75kDa subunit, Succinate dehydrogenase, subunit D and Glutamate carrier 2) with the highest density of polymorphisms were identified (Appendices 7–10), amplified with primers designed in this study (Table 4-1) and sequenced by Sanger method detailed in section 3.2.2. In addition, three nuclear genes (shaw, RNApyII and prp8) previously studied in *B. tabaci* (Hsieh *et al.*, 2014) were amplified and sequenced.

Table 4-1: Primers and PCR conditions used for amplification of mtCOI and nuclear genes. Ann ^{0}C = annealing temperature.

Target	Primer name	Sequence (5' – 3')	Amplicon	Ann ^{0}C	PCR program					Cycles
shaw [a]	ShawF ShawR	ATCTGCATCTCGATCCTCTC AGTTGAAGTCGTTGTGCGG	510 bp	60^{0}C						
RNApylI [a]	RNApylIF RNApylIR	CACAAAATGAGTATGATGGGTC CTTCACGACCTCCCATAGC	961 bp	60^{0}C	94^{0}C 5 min	94^{0}C 1 min	Ann^{0}C 1 min	72^{0}C 2 min	72^{0}C 10 min	35x
prp8 [a]	Prp8F Prp8R Prp8seqMF	GCCTTGGGAGGTGTTGAAG GGCTTGCATCCAGGGTACC CTGGAGTTCTCATTGCGATC	1060 bp	56^{0}C						
NADH-Q-oxidoreductase, 75kDa subunit	g17_FWD_2206 g17_REV_3177	GTCCAAGCTGTGTTGAAGCA GCAAGAAAGGAACAATGCACAG	972 bp	60^{0}C						
Succinate dehydrogenase, subunit D	g2_FWD g2_REV	TGGGACAGATGGTAAAATTCTGA TGGGCGTGTAAAATGTTCAAG	508 bp	59^{0}C						
Glutamate carrier 2, region 1 (GC1 marker [b])	g3r1_FWD g3r1_REV2	AGCAACCAGATAAGGGCTAGT TGTCATCAAAAGTCAGTCGCC	662 bp	55^{0}C	95^{0}C 3 min	94^{0}C 30 s	Ann^{0}C 1 min	72^{0}C 1 min	72^{0}C 10 min	35x
	g3r1_FWD3 [c] g3r1_REV3 [c]	TGTTTTGTATTTTGATCTATTCA CGAGGAGGAAATGTAAACAA	770 bp	52^{0}C						
Glutamate carrier 2, region 2	g3r2_FWD g3r2_REV	ACCTTGAAGTGTGCCTCGAA CGAGGAGGAAATGTAAACAA	350 bp	57^{0}C						

[a] Primers from Hsieh *et al.* (2014).

[b] This region was later identified as a part of the predicted Glutamate carrier 1 gene in the MEAM1 genome (see section 4.2.7).

[c] Primers designed to generate a different amplicon length for RFLP test.

4.2.6 Diagnostic RFLP test development

Each of the seven targeted nuclear sequences (Table 4.1) from each population was analysed by NEBcutter 2.0 (http://nc2.neb.com/NEBcutter2/; Vincze *et al.*, 2003) in order to find specific restriction sites that would enable populations of interest to be distinguished by their restriction patterns. In addition, new primers (g3r1_FWD3 and g3r1_REV3) were designed to generate a longer amplicon needed for some of the pairwise combinations. The final design consisted of a single region (Glutamate carrier 2, region1) amplified with two sets of primers to produce 662 bp and 770 bp amplicons, digested with enzymes *Bsp*1286I and *Bfu*CI, respectively. Since the time of our analysis, the *Bfu*CI in the NEB database was replaced by its isoschizomer *Sau*3AI, which has the same cut site.

After amplification by PCR, 0.5 µg of the purified product was digested by the restriction enzyme in CutSmart buffer (NEB) at 37°C for 3 hours (*Bsp*1286I) or overnight (*Bfu*CI). The total reaction volume (25 µl) was then loaded onto 2.5% (w/v) agarose gel in 0.5 x TBE with RedSafe staining solution (iNtRON Biotechnology, US) and ran for 1 h and 30 min at 4.5 V/cm. The 1 Kb+ DNA Ladder (Thermo Scientific) was used as a molecular size standard. The DNA was visualised under UV light (302 nm) and pictures were processed in GeneSnap (Syngene, UK). Efficiency of the method was tested by using simulated F_1 female hybrid DNA as PCR template prepared by 1:1 mixture of total DNA extracted from two females originating from two populations of interest.

4.2.7 Molecular cloning of the nuclear marker

The marker chosen for distinguishing between Spain Q1 and Israel Q2 populations was the 770 bp amplicon of Glutamate carrier 2, region 1. After a closer analysis it corresponded to a segment of 3' untranslated region of the predicted Glutamate Carrier 1 (GC1) gene (Appendix 11), annotated as Bta11593 in the genome of MEAM1 (Chen *et al.*, 2016). The alignment of our marker and the Bta11593 gene in Geneious 10.0.8 (Kearse *et al.*, 2012) revealed their 97–98% identity in the nucleotide sequence.

The GC1 amplicons from representative F_1 females originating from control and reciprocal crosses between Spain Q1 and Israel Q2 were cloned into pGEM®-T Vector System I (Promega) with a 2:1 insert:vector ratio. Colonies of transformant *Escherichia coli* cells JM109 were selected randomly and individually lysed in 200 µl sterile water by brief vortexing, then used as a template for PCR with vector-specific T7 and SP6 primers. Selected amplified inserts were sequenced at the ACRF Biomolecular Resource Facility at the Australian National University in Canberra. Trace files were analysed using

the Pregap and Gap4 programs within the Staden molecular analysis software (Bonfield *et al.*, 1995). The insert sequences were reduced to unique haplotypes in FaBox DNA Collapser version 1.41 (Villesen, 2007) and mapped to reference sequences from control crosses.

4.2.8 Screening for bacterial endosymbionts by PCR test

The presence of endosymbiotic bacteria was tested by PCR using genus-specific primers for the primary endosymbiont *Portiera* and five secondary endosymbionts (Table 4-2). A positive control was included in the screening for each secondary endosymbiont. DNA extracted from 15 whiteflies from a SSA1-SG3 colony was used as the positive control for *Arsenophonus*, DNA from one Asia II-1 individual for *Cardinium*, DNA from six Israel MEAM1 adults for both *Hamiltonella* and *Rickettsia* and DNA from six Asia I adults for *Wolbachia.* No positive control was included in the *Portiera*-specific reaction as the presence of the endosymbiont in the whitefly is obligatory (Thao and Baumann, 2004). The PCR products were visualised by loading 4 µl of the reaction volume onto 1% (w/v) agarose in 0.5 x TBE gel with RedSafe staining solution and running for about 1 h at 4.5 V/cm. The 1 Kb+ DNA Ladder (Thermo Scientific, UK) was used as a molecular size standard. The PCR products were visualised under UV light and pictures were processed in GeneSnap (Syngene, UK). Positive results were confirmed by sequencing the PCR products from two samples per population. Sanger sequencing was performed as described in section 3.2.2 and the sequences were compared to the NCBI database using BLASTn program (Altschul *et al.*, 1990).

Table 4-2: Primers and PCR conditions used for amplification of genus-specific amplicons of six bacterial endosymbionts.

Target	Primer name	Sequence (5' – 3')	Amplicon	Annealing ^{0}C	PCR program	Reference
Portiera 16S rDNA	28F 1098R	TGCAAGTCGAGCGGCATCAT AAAGTTCCCGCCTTATGCGT	1050 bp	58^{0}C		Zchori-Fein and Brown (2002)
Arsenophonus 23S rDNA	Ars23S-1 Ars23S-2	CGTTTGATGAATTCATAGTCAAA GGTCCTCCAGTTAGTGTTACCCAAC	750 bp	58^{0}C	94^{0}C, 5 min 94^{0}C, 30 s Ann ^{0}C, 45 s 72 ^{0}C, 1 min 30 s 72^{0}C, 10 min 35 cycles	Chiel *et al.* (2007)
Cardinium 16S rDNA	Card-F Card-R	TAGACACACACGAAAGTTCATGT GCATGCAATCTACTTTACACTGG	650 bp	57^{0}C		Ghosh *et al.* (2015)
Hamiltonella 16S rDNA	Hb-F Hb-R	TGAGTAAAGTCTGGGAATCTGG AGTTCAAGACCGCAACCTC	730 bp	58^{0}C		Gueguen *et al.* (2010)
Rickettsia 16S rDNA	Rb-F Rb-R	GCTCAGAACGAACGCTATC GAAGGAAAGCATCTCTGC	960 bp	58^{0}C		Gottlieb *et al.* (2006)
Wolbachia 16S rDNA	Wol16S-F Wol16S-R	CGGGGGAAAAATTTATTGCT CCCCATCCCTTCGAATAGGTAT	730 bp	58^{0}C		Ghosh *et al.* (2015), Heddi *et al.* (1999)

*Ann ^{0}C = annealing temperature.

4.2.9 High-throughput sequencing from single females

Total DNA from single female whiteflies, representing each laboratory colony, was isolated using the silica spin-column method (Thangaraj *et al.*, 2016) and quantified using Qubit 3.0 Fluorimeter (Thermo Scientific, UK). Purified total DNA was sheared into 400–500 bp fragments by an ultrasonicator (Covaris, US), and T4 DNA polymerase was applied to generate blunt ends. After adding an 'A' base to the 3' end of the blunt phosphorylated DNA fragments, adapters were ligated to the ends of the DNA fragments. After an enriching step via PCR with adapter-specific primers, the fragments were purified through agarose gel electrophoresis. The index tag was introduced into the adapter at the PCR stage, followed by a library quality test. The multiplexed library was then sequenced on a single lane of Illumina HiSeq 2000 platform (Novogene Bioinformatics Institute, Beijing, China) and 150 bp paired-end reads were generated. Raw sequencing data were generated by the base calling software CASAVA 1.8.2 (Illumina, US) according to its corresponding guidance.

4.2.10 Endosymbiont genome assembly and read mapping

Raw reads from two Illumina HiSeq datasets (single males in section 3.2.4 and single females in section 4.2.9) were mapped to bacterial genomes in the NCBI database using BLASTn (Altschul *et al.*, 1990). Relative abundances of each genus of bacterial endosymbiont were expressed as percentages of the mapped reads to a particular endosymbiont from all reads mapped to bacterial genomes. Tentative thresholds for positive infections were identified by plotting the percentages in a histogram using MS Excel.

4.3 Results

4.3.1 Reciprocal crossing experiments and hybrid fertility

All four control crosses produced both male and female progeny (Table 4-3 A). In the reciprocal F_1 crosses, however, no females were found in 10 out of 12 crosses (Table 4-3 B). The only type of reciprocal cross from which F_1 females emerged was between Spain Q1 and Israel Q2.

The highest number of F_1 offspring emerged in ASL control cross, which was significantly ($P<0.001$) different from the lowest number produced in the ♀Spain Q1 x ♂Sudan Q1 cross. The percentage of females did not differ significantly among the successful crosses (Table 4-3 A and 4-3 B). P values from the multiple comparisons among control and F_1 crosses can be found in Appendix 12.

Fertility of the F_1 females resulting from crosses between Spain Q1 and Israel Q2 was investigated by back-crossing them with males of both parental types. In all four combinations, female offspring was produced (Table 4-3 C). The F_2 females were present in 33 out of 35 replicates of all four combinations; the only two replicates with purely male offspring occurred in the ♀(♀Q2 x ♂SpQ1) x ♂Q2 cross.

There were, however, significant differences in the F_2 offspring counts, depending on the direction of the cross (Figure 4-2, Appendix 13). The offspring of (♀Q2 x ♂SpQ1) females back-crossed with Q2 males was 6–6.2 times smaller than the one produced by (♀SpQ1 x ♂Q2) females back-crossed to males of either parental type, which was a highly significant difference ($P<0.001$). The offspring of (♀Q2 x ♂SpQ1) females crossed with SpQ1 males was 3.5–3.6 times smaller than that of (♀SpQ1 x ♂Q2) females back-crossed to males of either parental type, which was also significantly different ($P<0.01$). The F_2 offspring of (♀Q2 x ♂SpQ1) females back-crossed with Q2 males was also significantly ($P<0.01$) smaller than the F_1 progeny produced in reciprocal crosses between SpQ1 and Q2 (4.8–5.4-fold).

Table 4-3: Means and standard errors from reciprocal crossing experiments among the four populations in this study. **A:** Control crosses between females and males from the same population. **B:** Reciprocal crosses between females and males from different populations, showing results from the F_1 generation. **C:** The F_2 progeny from back-crossing of hybrid F_1 females with males of either parental type. Different superscript letters indicate statistically significant differences ($P<0.05$) between crosses (Tukey's test). Multiple comparison including the F_2 offspring was done in a separate analysis (Figure 4-2).

Cross (1♀ x 3♂)	n	Mean no. of progeny	Mean no. of females	Mean % females
A: Controls				
♀Spain Q1 x ♂Spain Q1	9	21.8 ± 4.6 ^{bc}	9.1 ± 2.5	41.8 ± 8.4 ^{a}
♀Sudan Q1 x ♂Sudan Q1	10	28.2 ± 5.6 ^{bc}	11.5 ± 2.9	40.8 ± 6.9 ^{a}
♀Israel Q2 x ♂IsraelQ2	7	22.6 ± 5.4 ^{bc}	8.1 ± 2.5	36.1 ± 9.1 ^{a}
♀Uganda ASL x ♂Uganda ASL	10	42.7 ±8.4 ^{c}	13.4 ± 3.4	31.4 ± 5.4 ^{a}
B: Reciprocal crosses (F_1)				
♀Spain Q1 x ♂Uganda ASL	12	25.9 ± 4.7 ^{bc}	0.0 ± 0.0	0.0 ± 0.0
♀Uganda ASL x ♂Spain Q1	5	22.6 ± 6.4 ^{bc}	0.0 ± 0.0	0.0 ± 0.0
♀Israel Q2 x ♂Uganda ASL	13	17.4 ± 4.6 ^{ab}	0.0 ± 0.0	0.0 ± 0.0
♀Uganda ASL x ♂Israel Q2	7	20.1 ± 4.9 ^{bc}	0.0 ± 0.0	0.0 ± 0.0
♀Sudan Q1 x ♂Spain Q1	5	10.8 ± 3.3 ^{ab}	0.0 ± 0.0	0.0 ± 0.0
♀Spain Q1 x ♂Sudan Q1	9	11.0 ± 2.5 ^{b}	0.0 ± 0.0	0.0 ± 0.0
♀Sudan Q1 x ♂Israel Q2	8	27.9 ± 6.2 ^{bc}	0.0 ± 0.0	0.0 ± 0.0
♀Israel Q2 x ♂Sudan Q1	5	13.4 ± 4.0 ^{ab}	0.0 ± 0.0	0.0 ± 0.0
♀Sudan Q1 x ♂Uganda ASL	7	13.0 ± 3.3 ^{ab}	0.0 ± 0.0	0.0 ± 0.0
♀Uganda ASL x ♂Sudan Q1	11	31.3 ± 5.9 ^{ac}	0.0 ± 0.0	0.0 ± 0.0
♀Spain Q1 x ♂Israel Q2	7	36.6 ± 8.6 ^{ac}	13.4 ± 4.0	36.7 ± 7.2 ^{a}
♀Israel Q2 x ♂Spain Q1	6	32.8 ± 8.4 ^{bc}	14.8 ± 4.8	45.2 ± 8.5 ^{a}
C: Back-crosses (F_2)				
♀(♀Spain Q1 x ♂Israel Q2) x ♂Spain Q1	12	42.2 ± 7.6	25.1 ± 5.6	59.5 ± 5.2
♀(♀Spain Q1 x ♂Israel Q2) x ♂Israel Q2	11	41.1 ± 7.7	26.8 ± 6.2	65.3 ± 5.3
♀(♀Israel Q2 x ♂Spain Q1) x ♂Spain Q1	7	11.7 ± 3.0	9.6 ± 3.0	81.7 ± 10.2
♀(♀Israel Q2 x ♂Spain Q1) x ♂Israel Q2	5	6.8 ± 2.2	6.0 ± 2.3	88.2 ± 13.2

The proportion of females also differed between F_1 and F_2 generations and between the directions of the cross (Figure 4-2). In controls and F_1 reciprocal crosses of Spain Q1 and Q2, the female percentage ranged from 36.1% to 45.2%. However, back-crossing the hybrid F_1 females resulted in female-biased progeny with 59.5–93.8% of F_2 females. The bias was stronger (albeit non-significant) in F_1 hybrid females with Q2 maternal origin (83.8% and 93.8% of F_2 females) compared with their counterparts with Q1 maternal background (59.5% and 65.3%).

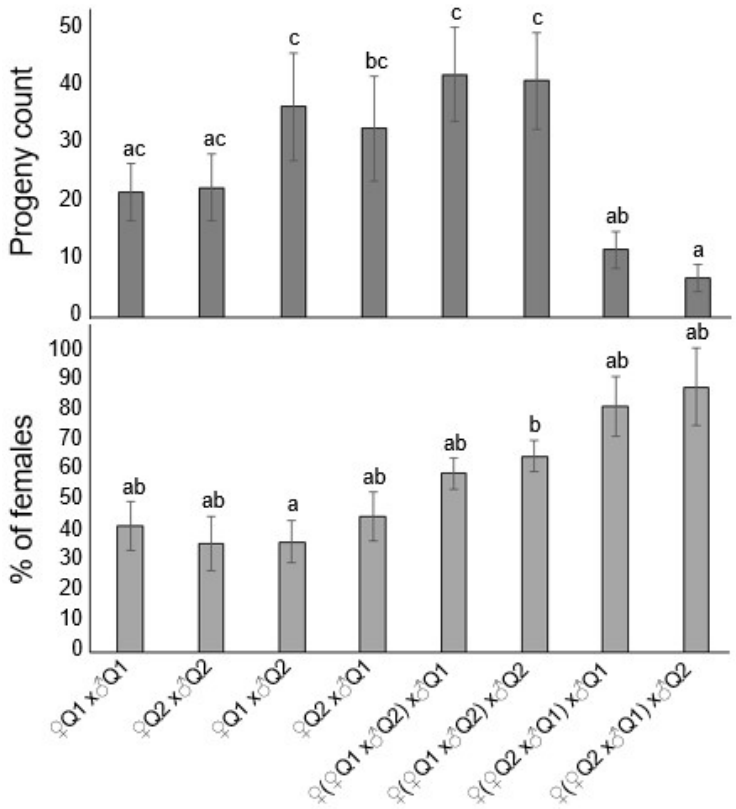


Figure 4-2: Results from F_1 and F_2 crossing experiments between Spain Q1 and Israel Q2.

4.3.2 Diagnostic RFLP test of the nuclear marker

Region 1 of the Glutamate carrier 2 candidate gene, later identified as the 3' untranslated region of the predicted GC1 gene ("GC1 marker" hereafter), was chosen as a suitable marker for distinguishing among our Q1, Q2 and ASL populations. The 662 bp amplicon of the GC1 marker contained two restriction sites recognised by *Bsp*1286I, however, while the site in position 467 bp was shared among all four populations, the site in position 321 bp was absent from sequence from Uganda ASL (Figure 4-3). The 770 bp amplicon of the GC1 marker contained three restriction sites for *Bfu*CI, but the one in position 526 bp was specific only for Spain Q1 and Uganda ASL (Figure 4-3).

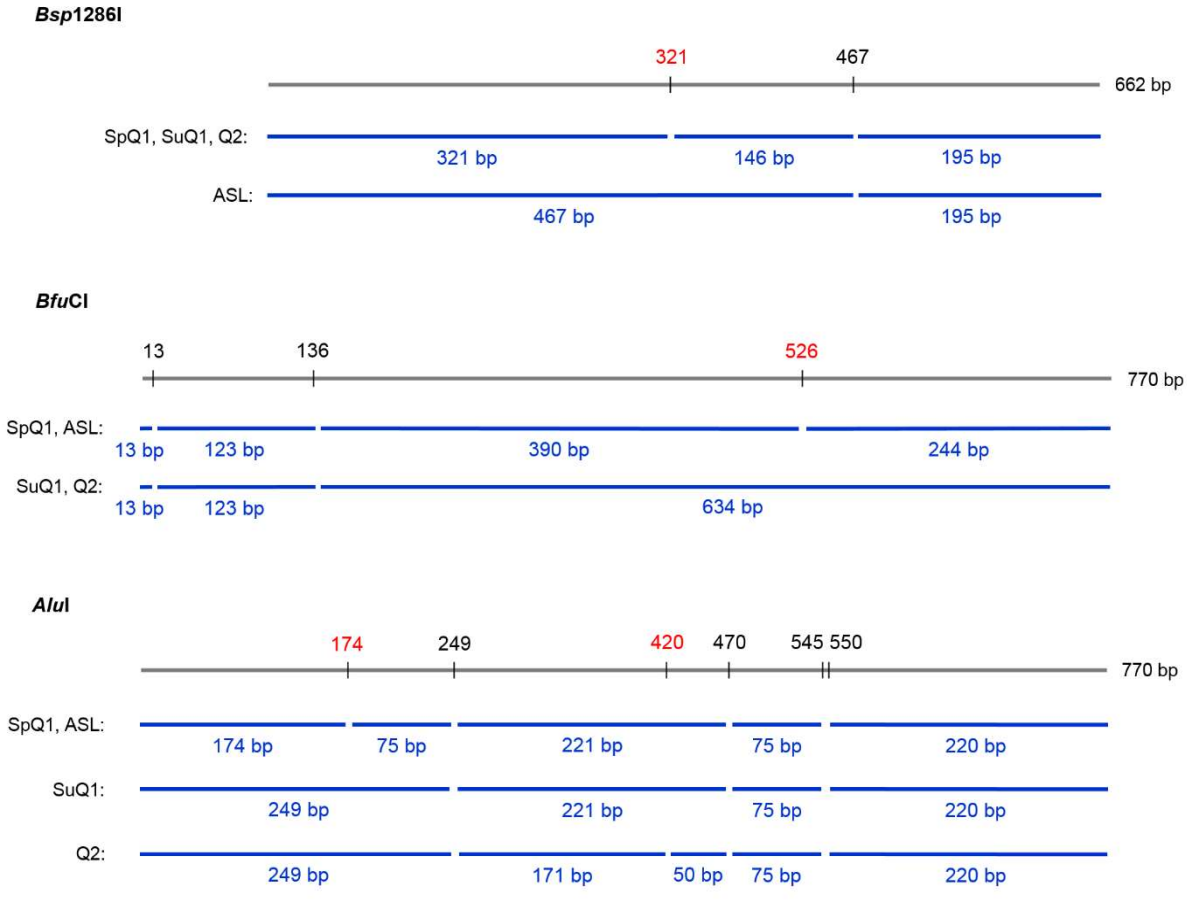


Figure 4-3: Scheme representing cut sites of *Bsp*1286I, *Bfu*CI and *Alu*I in the GC1 nuclear marker sequence (grey) and the resulting fragments (blue) for each population. Positions highlighted in red mark diagnostic sites.

These intraspecific polymorphisms allowed the development of an RFLP diagnostic test to study the parental origins of F_1 females potentially emerging from reciprocal crosses of each pairwise combination among the four populations (Table 4-4). The only exception was the cross between Sudan Q1 and Israel Q2, for which no specific differences were found in the restriction sites for *Bsp*1286I or *Bfu*CI. Only one enzyme, *Alu*I, was identified as potentially useful in distinguishing between Sudan Q1 and Israel Q2 (Figure 4-3). The digestion with *Alu*I, however, was not tested in the laboratory.

Table 4-4: Design of the RFLP diagnostic test (marker, amplicon size, restriction enzyme) for each pairwise combination among the four populations in this study.

	Spain Q1	**Sudan Q1**	**Israel Q2**
Sudan Q1	GC1 (770 bp) *Bfu*CI		
Israel Q2	GC1 (770 bp) *Bfu*CI	GC1 (770 bp) *Alu*I [a]	
Uganda ASL	GC1 (662 bp) *Bsp*1286I	GC1 (662 bp) *Bsp*1286I	GC1 (662 bp) *Bsp*1286I

[a] Not tested.

The output of an enzymatic digestion and a subsequent gel electrophoresis is a restriction profile that can be compared among samples. After cutting the 662 bp amplicon with *Bsp*1286I, a sample from Uganda ASL could be distinguished based on two fragments sized 195 and 467 bp, while the Q1 and Q2 populations had three fragments (146, 195 and 321 bp) (Figure 4-4 A and 4-4 B). An exception was one male from Sudan Q1 that showed a pattern typical for Uganda ASL (male a in Figure 4-4 C). Partial 3' mtCOI sequence obtained from this specimen, however, was identical to Sudan Q1.

Digestion of the 770 bp amplicon with *Bfu*CI allowed the distinction between Spain Q1 (123, 244 and 390 bp) and Sudan Q1 or Israel Q2 (123 and 634 bp) (Figure 4-5). All populations also shared a 13 bp fragment, however, molecules below 50 bp were not visible in the agarose gel after electrophoresis. The sample from Spain Q1 was repeatedly not fully digested, resulting in five visible fragments instead of three (Figure 4-5).

The hypothetical test involving *Alu*I would generate three bands for Sudan Q1 (fragments 75, 220 + 221 and 249 bp), and five in Israel Q2 (50, 75, 171, 220 and 249 bp) (Figure 4-3).

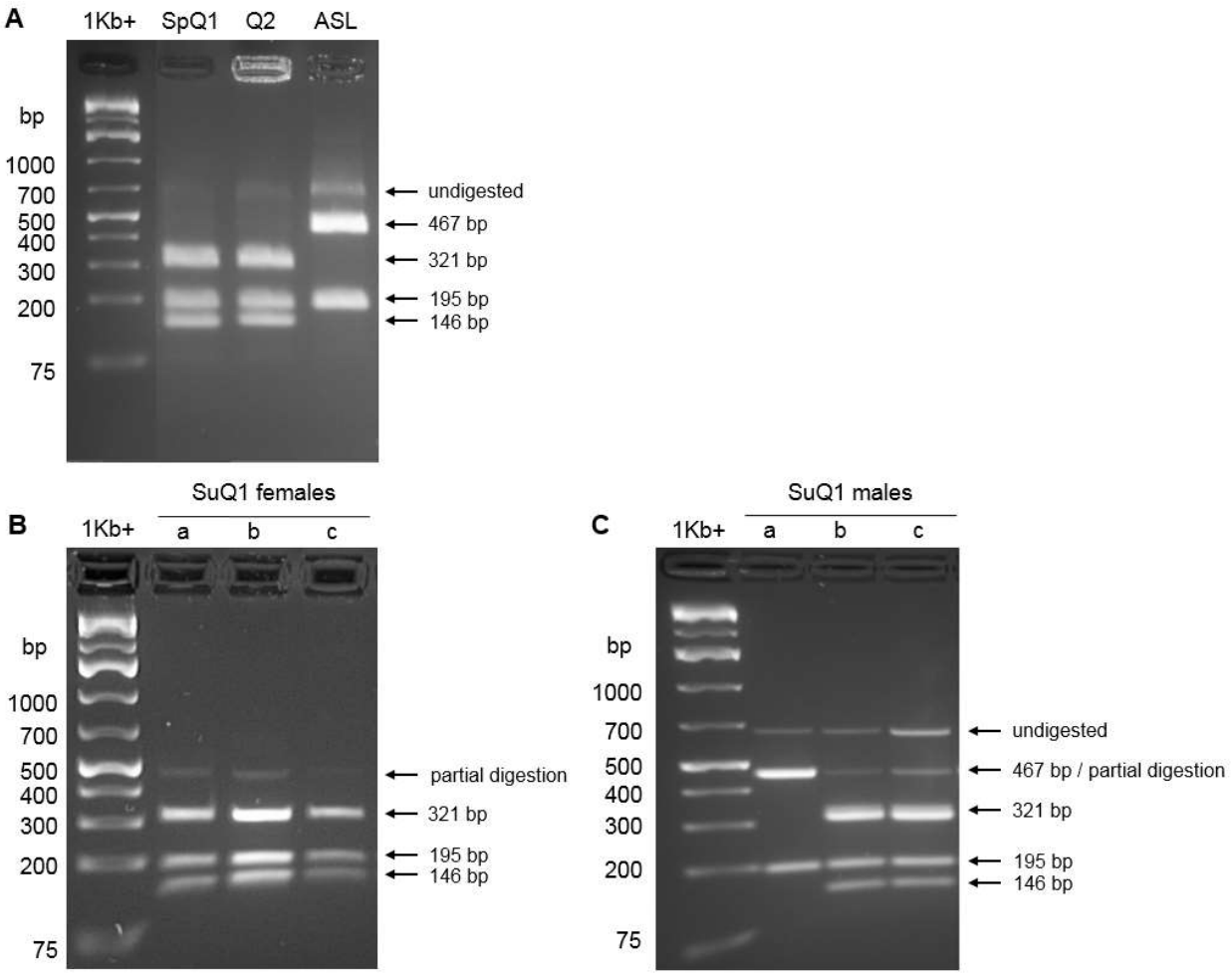


Figure 4-4: Restriction patterns by *Bsp*1286I visualised in 2.5% agarose gel with 1Kb+ DNA ladder (Thermo Scientific) used as a molecular standard. **A:** Patterns of Spain Q1, Israel Q2 and Uganda ASL. **B:** Three samples of Sudan Q1 females showing the same pattern as Spain Q1 and Israel Q2. **C:** Three samples of Sudan Q1 males. Male a shows pattern identical to the one of Uganda ASL, while males b and c share the same pattern with all other tested Q1 and Q2 individuals.

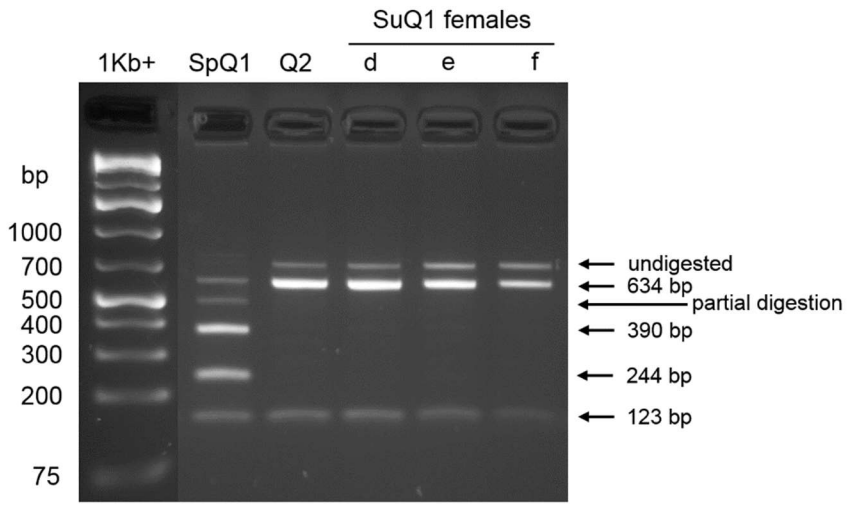


Figure 4-5: Restriction patterns by *Bfu*CI visualised in 2.5% agarose gel with the 1Kb+ DNA ladder (Thermo Scientific) used as a molecular standard. Spain Q1 has a distinctive pattern from Israel Q2 and Sudan Q1.

The trial digestion using simulated hybrid female DNA showed that both DNA templates in the 1:1 mix amplified successfully, as restriction patterns resulting from the mixed samples contained fragments specific for both parental populations (Figure 4-6).

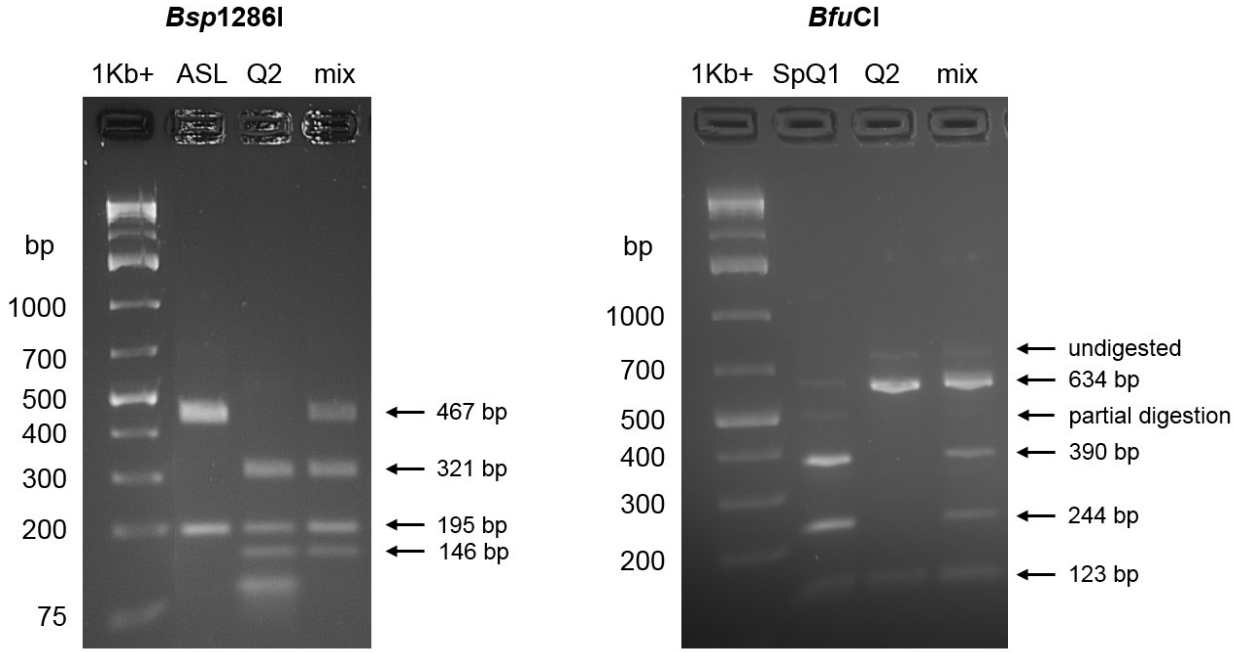


Figure 4-6: RFLP design tested on simulated hybrid female DNA (mix). Left: 662 bp amplicon of the GC1 marker digested with *Bsp*1286I with Uganda ASL and Israel Q2 as an example. Right: 770 bp amplicon of the GC1 marker digested with *Bfu*CI to distinguish between Spain Q1 and Israel Q2. The 1 Kb+ DNA ladder (Thermo Scientific) was used as a molecular standard.

The RFLP test developed was used to check the parental origin of F_1 females that emerged from the reciprocal crosses between Spain Q1 and Israel Q2. All tested F_1 females showed a combined pattern from both parental populations (Figure 4-7), similar to the one seen in simulated SpQ1 x Q2 hybrid (Figure 4-6). However, the expected 634 bp fragment appeared as two clearly separated bands after visualization in the agarose gel (Figure 4-7), which was not apparent in the previous test.

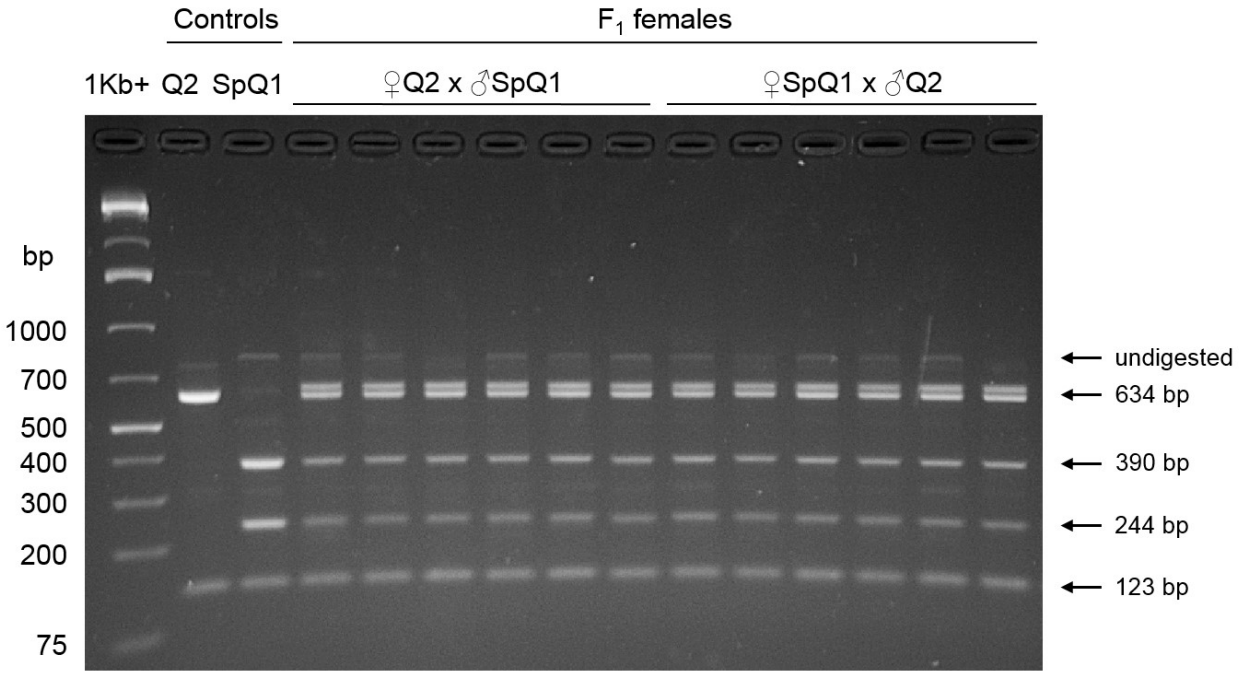


Figure 4-7: RFLP screening of 12 F_1 females from reciprocal crosses between Spain Q1 and Israel Q2 and parental adults as controls. The 1 Kb+ DNA ladder (Thermo Scientific) was used as a molecular standard.

4.3.3 Molecular cloning of the nuclear marker

For an unequivocal confirmation of the hybrid status of the F_1 females from Spain Q1 x Israel Q2 reciprocal crosses, their parental origin was analysed in addition by Sanger sequencing of the GC1 nuclear DNA marker. The marker sequence enabled the distinction between these two populations based on a 4 bp insertion/deletion (INDEL) and two single nucleotide polymorphisms (SNPs). GC1 PCR products were initially sequenced directly from two parental controls and four F_1 females: f1 and f2 originating from the ♀SpQ1 x ♂Q2 cross; f3 and f4 from ♀Q2 x ♂SpQ1. The trace files from direct Sanger sequencing revealed heterozygous sequences in the F_1 females, compared with the homozygous ones from control individuals. The heterozygosity was demonstrated as mixed traces downstream or upstream from the INDEL site, depending on whether FWD3 or REV3 was used as the sequencing primer (Figure 4-8).

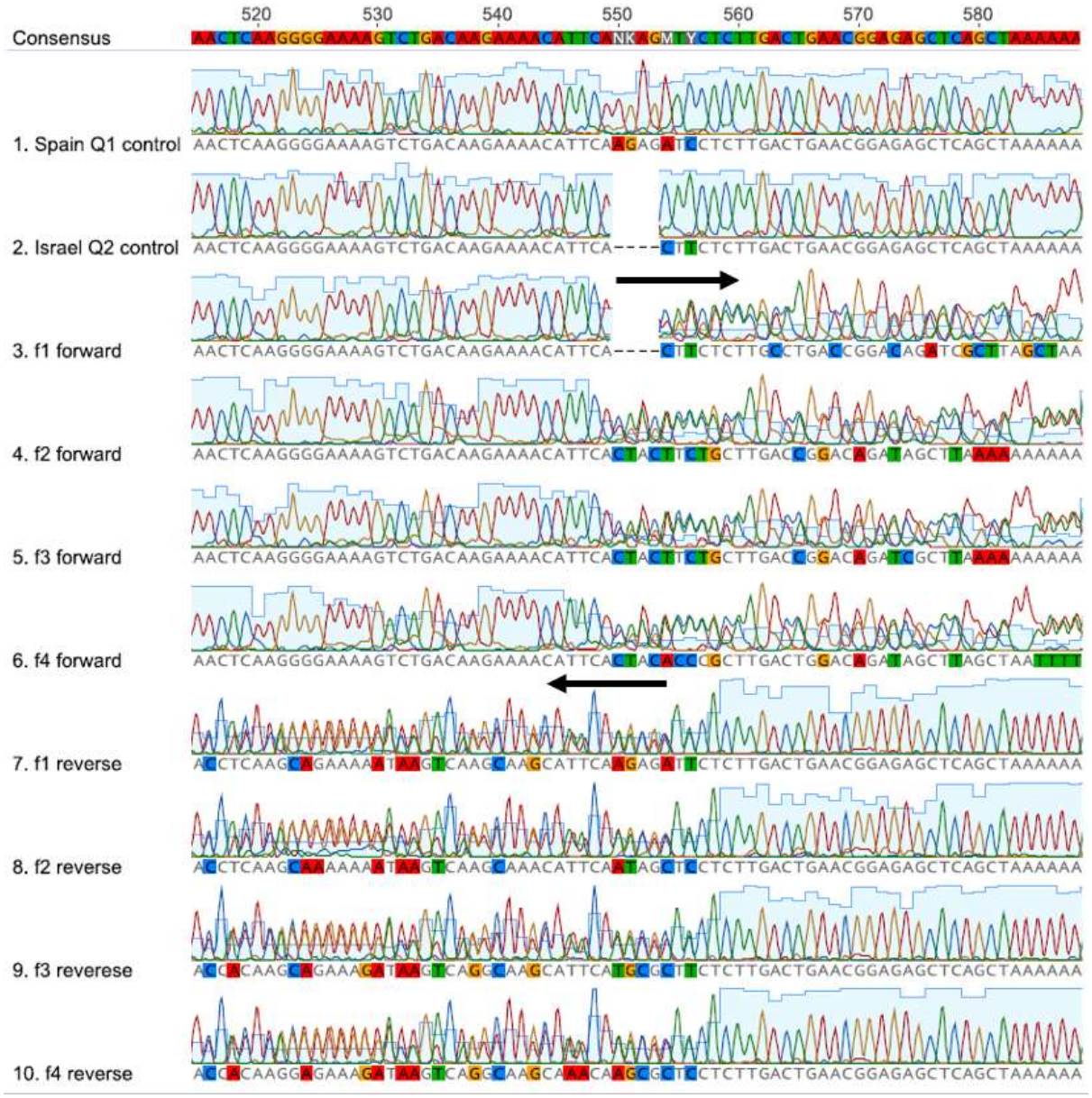


Figure 4-8: Sequence alignment and chromatograms of the directly sequenced PCR products of the GC1 marker from Spain Q1 control, Israel Q2 control and four hybrid F_1 females sequenced with both forward and reverse primers. Black arrows highlight the direction of sequencing with forward (sequences 3–6) and reverse (sequences 7–10) primers and therefore the location of mixed traces downstream and upstream of the INDEL site, respectively.

The mixed sequences of GC1 marker from females f1–f4 were separated by molecular cloning in *E. coli*. Sixteen marker sequences (four from each female) were reduced to 12 unique haplotypes and aligned. Aligning the haplotypes with the reference sequences from controls revealed that haplotypes Hap1–Hap6 were characteristic for Spain Q1, while Hap7–Hap12 were of the Israel Q2 type. Each of the four F_1 females contained sequences of both parental types (Figure 4-9).

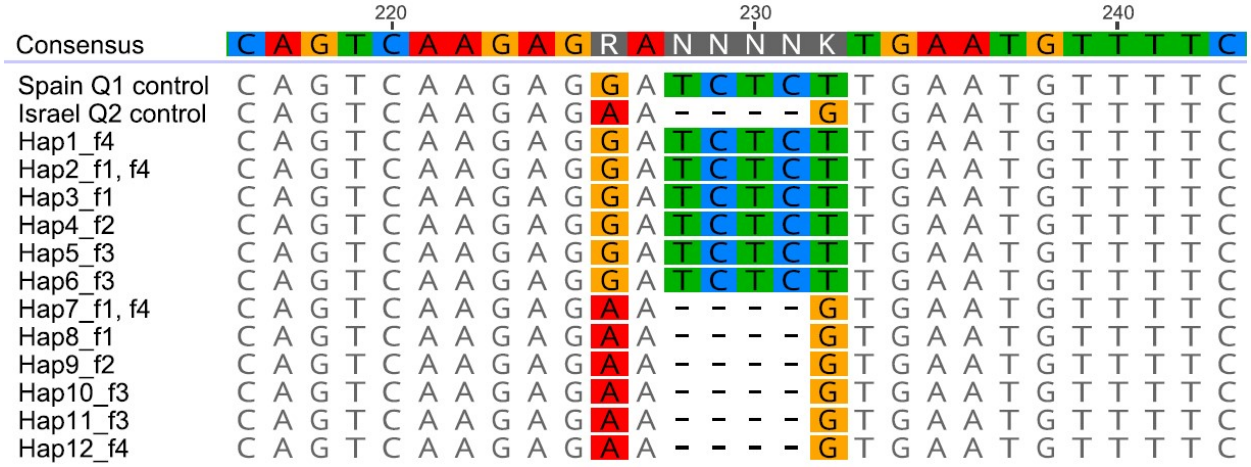


Figure 4-9: Molecular proof of parental origin of hybrid F_1 females. Spain Q1 control and Israel Q2 control show population-specific polymorphisms: 4bp 'GAGA' INDEL and SNPs at positions 509 (SpQ1: 'A', Q2: 'C') and 515 (SpQ1: 'C', Q2: 'T'). Unique haplotypes of the GC1 marker (Hap1–Hap12) from four F_1 females produced in reciprocal crosses between ♀SpQ1 x ♂Q2 (f1, f2) and ♀Q2 x ♂SpQ1 (f3, f4) were aligned against GC1 sequences from females produced in control crosses (Spain Q1 control, Israel Q2 control). The presence of both Spain Q1- and Israel Q2-specific sequences in all four females confirmed that the F_1 females were hybrids between these two MED populations.

4.3.4 Infection status by bacterial endosymbionts

The infection status of the Spain Q1, Sudan Q1, Israel Q2 and Uganda ASL core colonies was investigated by PCR test using primers specific for the primary endosymbiont *Portiera* and five genera of common secondary endosymbiotic bacteria (*Arsenophonus*, *Cardinium*, *Hamiltonella*, *Rickettsia* and *Wolbachia*). Ten females sampled from each colony were used for the screening.

The primary endosymbiont *Portiera aleyrodidarum* was present in all individuals from all populations (Figure 4-10 A) and served as a quality control for the DNA extracted from the 40 females used in the screening. None of the four populations was infected with *Arsenophonus*. *Cardinium*-specific products appeared only in two Sudan Q1 samples and in low quantity (Figure 4-10 A). *Hamiltonella* was present in all Spain Q1 and most Sudan Q1 females, while it was absent from Israel Q2 and Uganda ASL (Figure 4-10 B). *Rickettsia* was present in all populations except for Spain Q1, and *Wolbachia* only occurred in Sudan Q1 and Uganda ASL (Figure 4-10 B). The results of this PCR test are summarised in Table 4-5. PCR products from positive samples were sequenced and compared to the NCBI database using BLASTn. All tested sequences were 99–100% identical to samples from their respective bacteria infecting *B. tabaci*.

The intensity of PCR products sometimes differed among positive results. For example, strong signals were visible in *Portiera* across all populations, *Hamiltonella* in both Spain Q1 and Sudan Q1 and *Wolbachia* in Uganda ASL. In the case of *Rickettsia*, however, the PCR products were produced in much lower quantity in Uganda ASL compared with Israel Q2 and the quantity of product in Sudan Q1 was intermediate.

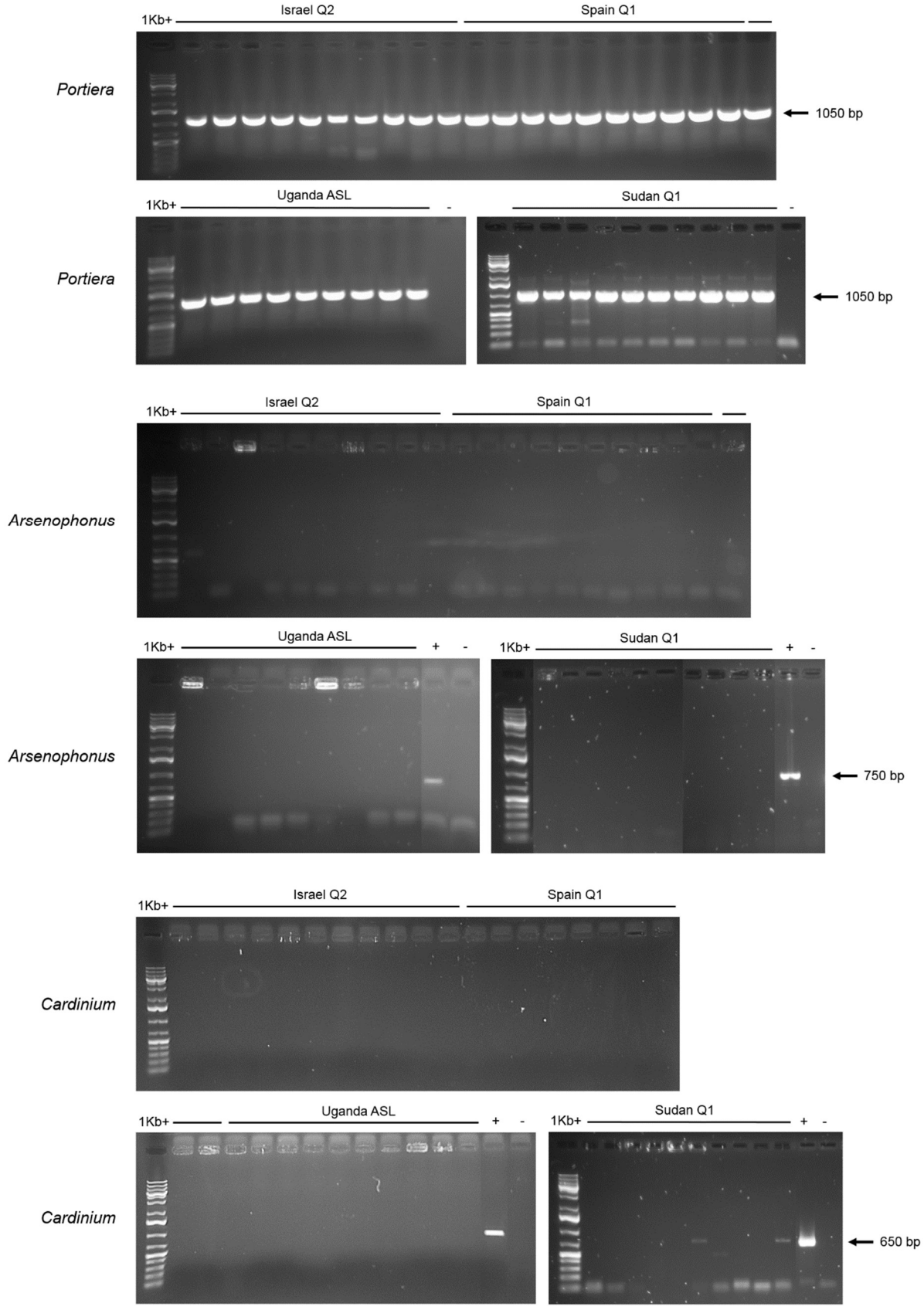


Figure 4-10 A: PCR screening for six genera of bacterial endosymbionts across four laboratory populations of MED putative species. The 1 Kb+ DNA ladder (Thermo Scientific) was used as a molecular size standard. The plus and minus signs mark positive and negative controls, respectively. The expected molecular size of PCR product specific for each bacterium (Ghosh *et al.*, 2015) is indicated with an arrow.

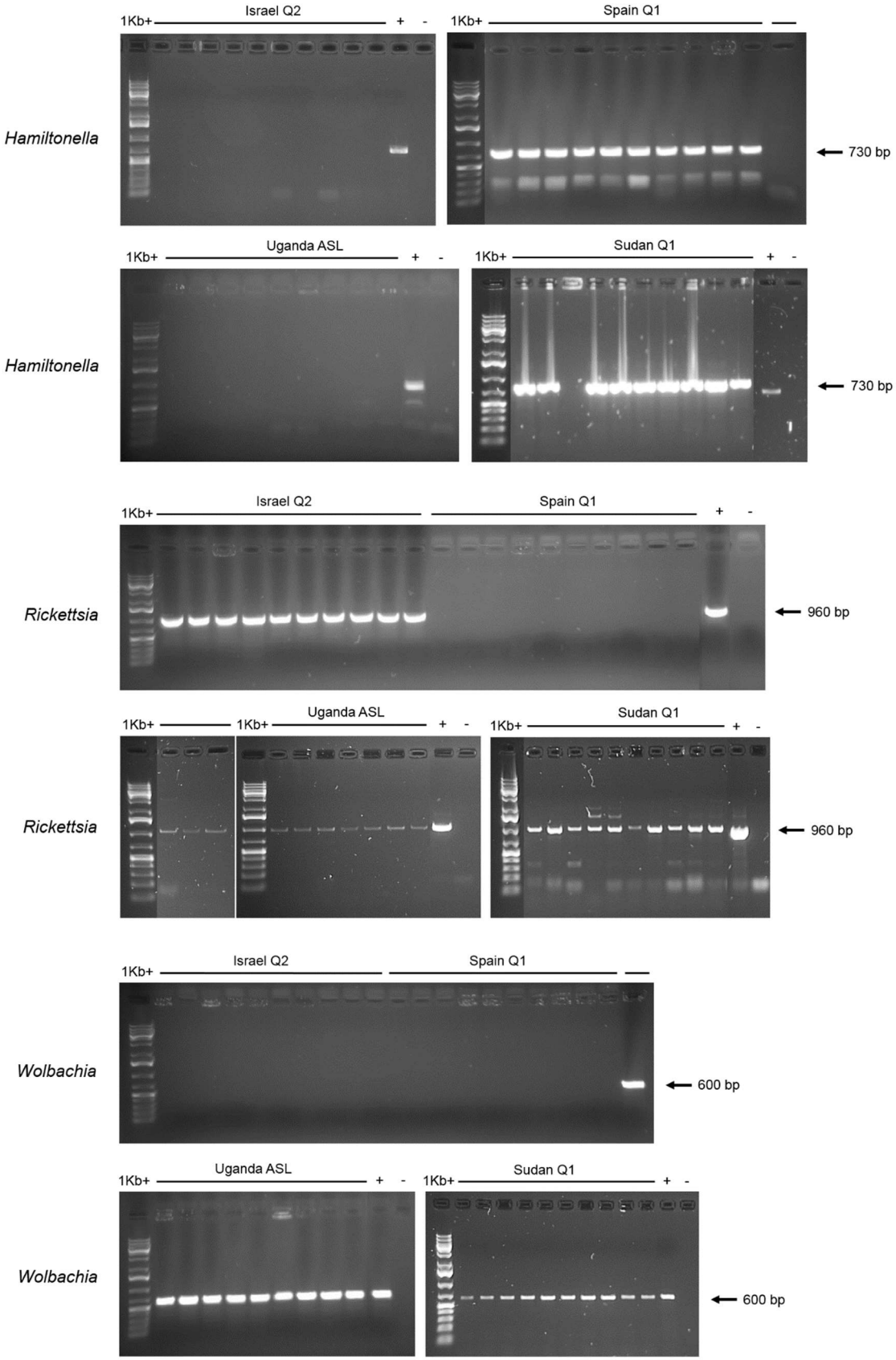


Figure 4-10 B: PCR screening for six genera of bacterial endosymbionts across four laboratory populations of MED putative species. The 1Kb+ DNA ladder (Thermo Scientific) was used as a molecular size standard. The plus and minus signs mark positive and negative controls, respectively. The expected molecular size of PCR product specific for each bacterium (Ghosh *et al.*, 2015) is indicated with an arrow.

Table 4-5: Infection status of whitefly colonies used in this study by primary (*P* = *Portiera*) and secondary endosymbiotic bacteria (*A* = *Arsenophonus*, *C* = *Cardinium*, *H* = *Hamiltonella*, *R* = *Rickettsia*, *W* = *Wolbachia*). The numbers of positive samples out of the 10 females tested are indicated, assessed by the presence of the PCR products after amplification with genus-specific primers. Dark grey highlights samples with a majority of females that tested positive, light grey with a minority.

	P	*A*	*C*	*H*	*R*	*W*
Israel Q2	10	0	0	0	10	0
Spain Q1	10	0	0	10	0	0
Sudan Q1	10	0	2	9	10	10
Uganda ASL	10	0	0	0	10	10

4.3.5 Relative abundance of bacterial endosymbionts

The availability of a genomic resource for our populations allowed for estimating the relative abundance of bacterial endosymbionts in their whitefly hosts. The frequency histogram plotting the percentages of mapped reads revealed two tentative cut-off values for a positive infection; one at 3% and the other at 5% (Figure 4-11).

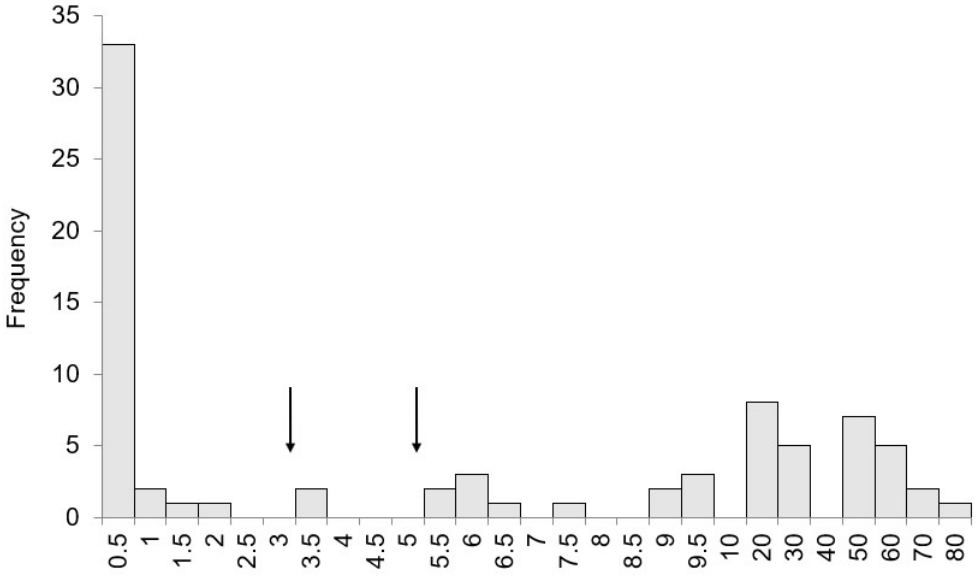


Figure 4-11: Frequency histogram showing the distribution of percentages of mapped reads to endosymbiont genomes. Arrows indicate the 3% and 5% tentative cut-offs.

The relative abundances of reads mapped to endosymbiont genomes were congruous between male and female datasets in most cases (Table 4-6). In four instances, however, there was a single outlier in an otherwise consistent result, which can be seen for *Portiera* in Sudan Q1, *Cardinium* in Spain Q1, *Cardinium* in Sudan Q1 and *Wolbachia* in Spain Q1 (Table 4-6). In addition, there was a drastic difference in the abundance of *Portiera* between male and female datasets of Spain Q1 (37- to 58-fold) and Uganda ASL (118-fold). The female samples were unequivocally positive (43.8–71.5%), while the male samples appeared negative (0.4–1.6%).

The relative abundances of reads that mapped to *Rickettsia* in Israel Q2 and *Wolbachia* in Uganda ASL were the highest among all secondary endosymbionts in all populations (Figure 4-12). In Sudan Q1 and Uganda ASL *Rickettsia* reads were nearly absent, unlike the results from the PCR test. Similarly, the presence of reads that mapped to *Cardinium* in Uganda ASL and *Wolbachia* in Spain Q1 (Figure 4-12) were not picked up in the PCR test (Table 4-5).

Table 4-6: Percentages of reads mapped to genomes of bacterial endosymbionts (*P* = *Portiera*, *C* = *Cardinium*, *H* = *Hamiltonella*, *R* = *Rickettsia*, *W* = *Wolbachia*). Results are shown for each dataset (M = male, F = female) and each read file (R1, R2). *Arsenophonus* is not shown due to the lack of *Arsenophonus*-like sequences in all datasets. Light grey cells highlight values between 3–5%, dark grey over 5%.

Whitefly population	Sample code	Sex	Read file	% of mapped reads				
				P	*C*	*H*	*R*	*W*
Israel Q2	31F	F	R1	40.22	0.00	0.00	56.16	0.00
Israel Q2	31F	F	R2	40.12	0.00	0.00	55.94	0.00
Israel Q2	SV8	M	R1	27.80	0.01	0.02	60.34	0.03
Israel Q2	SV8	M	R2	29.54	0.01	0.02	60.24	0.02
Spain Q1	29F	F	R1	57.30	3.32	19.57	0.10	14.35
Spain Q1	29F	F	R2	71.52	0.00	19.81	0.07	0.01
Spain Q1	SV10	M	R1	1.38	0.03	6.48	0.81	0.19
Spain Q1	SV10	M	R2	1.57	0.04	7.26	0.87	0.20
Sudan Q1	32F	F	R1	5.13	9.23	14.52	0.28	27.45
Sudan Q1	32F	F	R2	57.01	3.31	19.46	0.10	14.26
Sudan Q1	SV2	M	R1	5.13	9.23	14.52	0.28	27.45
Sudan Q1	SV2	M	R2	5.70	9.27	15.29	0.29	28.67
Uganda ASL	36F	F	R1	43.93	5.84	0.00	0.11	45.39
Uganda ASL	36F	F	R2	43.75	5.79	0.00	0.10	45.22
Uganda ASL	SV5	M	R1	0.37	8.88	0.04	0.28	49.46
Uganda ASL	SV5	M	R2	0.37	8.87	0.04	0.29	51.13

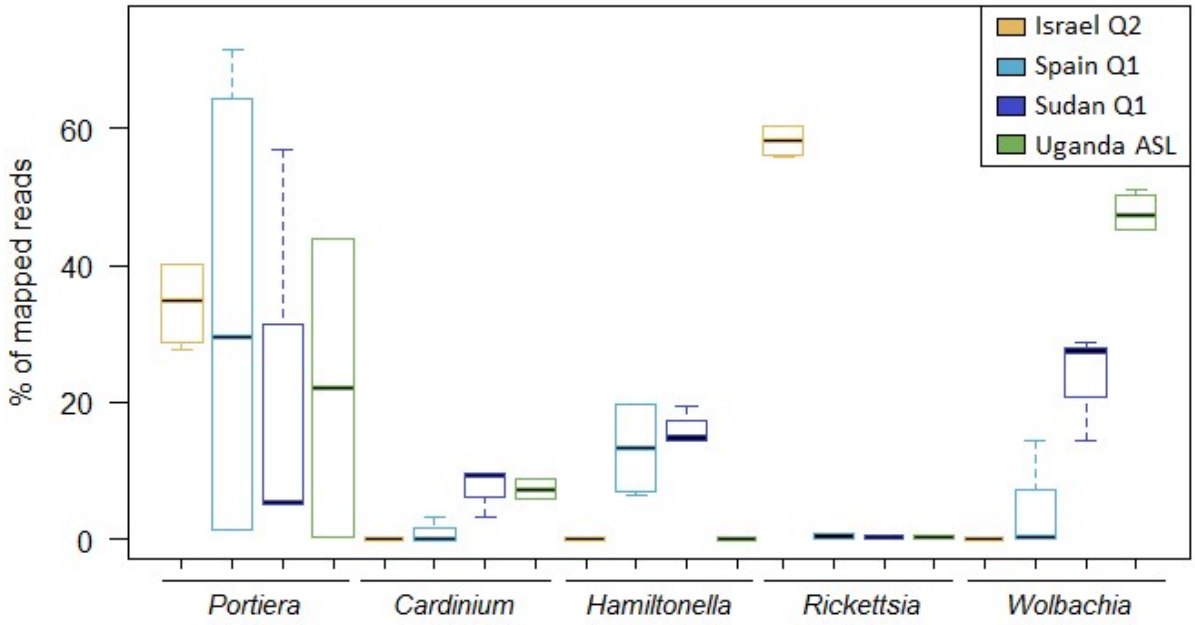


Figure 4-12: Percentages of reads combined from male and female HTS datasets mapped to genomes of bacterial endosymbionts. *Arsenophonus* is not shown due to the lack of *Arsenophonus*-like sequences in all datasets. Box lengths represent the interquartile range, whiskers express the full range and black bars correspond to the medians.

4.4 Discussion

This chapter expanded on the knowledge about reproductive relationships among MED groups, answering our third research question. We addressed this question by conducting laboratory crossing experiments in which progeny counts and sex ratios, as well as the viability and fertility of the hybrid progeny, was observed and compared. This information was missing from the microsatellite field studies of Gauthier *et al.* (2014), Terraz *et al.* (2014), Mouton *et al.* (2015) and Thierry *et al.* (2015). In addition, the endosymbiotic component of our whitefly population was investigated, which enabled the formulation of hypotheses for future research and addressing the fourth research question about the potential involvement of endosymbionts in the reproductive relationships within the MED putative species.

Based on previously published research, our hypotheses for the reproductive relationships among the MED populations were: (i) ASL is reproductively isolated from Q1 and Q2, and (ii) the Q1 and Q2 populations are reproductively compatible. The first hypothesis was supported by the purely male progeny in all reciprocal crosses involving the Uganda ASL population, indicating a presence of reproductive barriers preventing copulation, egg fertilisation or viability of any hybrid progeny between ASL and Q1 or Q2 populations. The second hypothesis was supported only partially. On one hand, the Mediterranean populations of Q1 and Q2 from Spain and Israel mated successfully and produced fertile hybrids. On the other hand, however, the sub-Saharan Q1 population from Sudan did not interbreed with any other population, including Israel Q2 and Spain Q1. This was surprising, considering the close genetic relationships among Q1 and Q2 populations observed at the mtDNA level in Chapter 3.

Our case of inconsistency between the level of genetic relatedness and the expected level of reproductive compatibility is not an isolated one. Evidence of reproductive isolation between very closely related populations was reported within the putative species SSA1, which has been divided into four subgroups (SG1, SG2, SG3 and SG4) (Legg *et al.*, 2014). Mugerwa (2018) reported that while SG1 and SG2 were compatible (3' partial mtCOI nucleotide distance 1.7%), the SG3 was incompatible with both SG1 and SG2 (distance 1.3% and 1.8%, respectively). Our MED populations showed an even closer mtCOI nucleotide identities than the SSA1 subgroups. The 3' partial mtCOI nucleotide distance was 1.07% between Sudan Q1 and Israel Q2, and only 0.15% between Sudan Q1 and Spain Q1 (Table 3-4). Conversely, cases of successful hybridisation between genetically distant populations (4.5–14.3%) have also been reported, namely between Asia II-3 x Asia II-9, MEAM1 x IO, MED x IO and

MEAM1 x Australia (see section 2.2.2) (De Barro and Hart, 2000; Delatte *et al.*, 2006; Qin *et al.*, 2016; Elfekih *et al.*, 2018). This indicates that the proposed 3.5% boundary in partial mtCOI nucleotide distance is not fully consistent with the presence of reproductive barriers among *B. tabaci* populations.

It could be argued that the mating behaviour of insects in a laboratory is not always a true representation of their behaviour in the field. However, such concerns usually relate to cases in which the organisms interbreed in captivity, but not in the wild (Coyne and Orr, 1989; Benirschke and Kumamoto, 1991; Wang, 2007). The case demonstrated in this chapter, *i.e.* the failure to interbreed under laboratory conditions, while the reproduction in control crosses is normal, is usually indicative of reproductive isolation in the field as well (Diehl and Bush, 1984). In addition, our results are congruent with the previously published evidence of ASL being a reproductive isolated entity from Q1 (Mouton *et al.*, 2015), and of the compatibility between Q1 and Q2 groups (Gauthier *et al.*, 2014; Terraz *et al.*, 2014; Thierry *et al.*, 2015; Hadjistylli *et al.*, 2016), although the latter seems to only apply if the Q1 population originates from the Mediterranean Basin and not sub-Saharan Africa.

The question of fertility of Spain Q1 x Israel Q2 hybrids was addressed by conducting back-crossing experiments with F_1 females and parental type males. In all four types of back-crosses, the F_1 hybrids were fertile, regardless of which population was the maternal and paternal genetic source. There were substantial differences, however, in the number and sex ratio of the F_2 progeny produced by F_1 females from different directions of the reciprocal crosses. Hybrid females with Q1 maternal background produced larger number of offspring ($P<0.01$) and a relatively normal percentage of females (59.5–65.3%), while the offspring of their counterparts with Q2 maternal background were few and female-biased (81.7–88.2%, non-significant). This result suggests that there are additional factors playing a role in the reproductive relationship between these two populations and making the interaction asymmetric.

The asymmetry of gene flow between the Mediterranean Q1 and Q2 populations has previously been suggested, because the nuclear background associated with Q1 individuals introgressed into, or completely replaced, the nuclear background of individuals with Q2 mitochondrial type, but not *vice versa* (Gauthier *et al.*, 2014). However, a different population genetic study based on microsatellites showed evidence for both directions of hybridisation (Hadjistylli *et al.*, 2016). Samples from France and Canary Islands had “Western Mediterranean” (Q1) nuclear background and Q2 mitochondrial type, while a sample from Greece had Q1-type mitochondria and “Eastern

Mediterranean" (Q2) nuclear DNA introgressed into its genome (Hadjistylli *et al.*, 2016). The interaction also did not appear asymmetric in Terraz *et al.* (2014), as the sympatric populations in France and Spain had homogeneous nuclear backgrounds regardless of the mitochondrial haplotype. In our study, no asymmetry was observed between the two directions of the cross in the F_1 generation. In the F_2 generation, however, the higher fecundity ($P<0.01$) of F_1 females with Q1 mitochondria compared with their Q2 counterparts could lead to a higher prevalence of hybrids with Q1 mitochondrial type in the field.

The validity of our results from the crossing experiments was ensured by rigorous testing of parental origin of the F_1 females to confirm that they were hybrids and discard any false positives that might have occurred due to a contamination. A nuclear molecular marker was identified to enable the identification of both paternal and maternal origin. The 3' untranslated region of the predicted GC1 gene proved to be a useful marker with sufficient capacity to distinguish between closely related populations within the MED putative species. F_1 females potentially emerging from five out of six pairwise combinations of crosses could be distinguished by using a simple PCR and two relatively cheap restriction enzymes *Bsp*1286I and *Bfu*CI. The sixth combination (Israel Q2 x Sudan Q1) could theoretically be resolved by using the third enzyme *Alu*I, however, this method has not been tested because the results from crossing experiments were already available at the time and it was no longer necessary to design the test as no F_1 female emerged from this cross.

Two limitations of the RFLP method were encountered in this work. Firstly, it can produce ambiguous results by incomplete digestion. This issue can be overcome by adjusting the reaction conditions and the amount and purity of the DNA template. Secondly, despite all effort, optimal results were not obtained for GC1 product from Spain Q1 cut with the *Bfu*CI enzyme. Distinguishing between Spain Q1 and Israel Q2 was crucial, as they were the only two populations producing F_1 females in reciprocal crosses. Thus, this problem was circumvented by molecular cloning of the GC1 marker to produce unambiguous evidence of the parental origin of the hybrids.

The number of unique GC1 haplotypes present in three out of the four F_1 females tested (f1, f3 and f4) exceeded the theoretically possible number of alleles per individual. For example, two Q1-specific (Hap2 and Hap3) and two Q2-specific haplotypes (Hap7 and Hap8) were detected in the hybrid female f1 (Figure 4-9), while in reality she would have received only one allele from each parent. The 12 GC1 haplotypes differed from one another only by one to four SNPs that occurred outside the diagnostic region shown

in Figure 4-9. These SNPs were probably errors introduced during the PCR amplification step by DreamTaq polymerase. The error rate of *Taq* DNA polymerases is 1–20 × 10^{-5} (Tindall and Kunkel, 1988; Keohavong and Thilly, 1989; Cline *et al.*, 1996), which is 6- to 50-fold higher than more precise DNA polymerases such as *Pfu* or Phusion Hot Start (McInerney *et al.*, 2014). Despite these introduced errors, the cloned sequences showed unequivocally that each hybrid female contained a copy from both parental types as seen in the INDEL region. This method of checking for genuine hybrid female progeny could be improved by using a high-fidelity DNA polymerase and a reduced number of PCR cycles for amplification of the marker sequence prior to cloning. Alternatively, the hybrids could be detected using microsatellite markers (McKenzie *et al.*, 2012; Gauthier *et al.*, 2014; Terraz *et al.*, 2014; Thierry *et al.*, 2015) or genome-wide SNPs (Elfekih *et al.*, 2018).

Analysis of the composition of bacterial endosymbionts in each population was done by combining two methods. The test consisting of conventional PCR with genus-specific primers has often been used before (e.g. Gueguen *et al.*, 2010; Thierry *et al.*, 2011; Ghosh *et al.*, 2015; Mugerwa, 2018; Tang *et al.*, 2018). The genomic approach of assembling genomes and mapping reads, making the use of available HTS data, has not been used for *B. tabaci* before.

Both methods have their limitations. The PCR method can suffer from false negative results stemming from a mismatch between the bacterial DNA sequence and the genus-specific primer. It could also produce false positives in the form of non-specific PCR products; however, these errors were eliminated in the process of sequencing of the products and comparing them to the NCBI database. An advantage of the PCR method is that it is more cost-effective and it is easier to scale up by testing many individuals separately. Given the relatively high cost of HTS, scaling up could be done by pooling many individuals into one sample, however, the relative frequencies of reads might not accurately reflect the frequency of positive infections in the pooled sample. In our study, only two individuals (one male and one female) per population were studied by the HTS method, which increases the probability of obtaining skewed results due to small sample size. In addition, coverage may also present an issue, which became apparent after mapping *Portiera* reads from samples of Spain Q1 and Uganda ASL males. The relative abundance was very low compared with other samples, which seems unlikely as the presence of *Portiera* is obligatory. Furthermore, the unequal coverage also means that the relative abundance of reads mapped to each endosymbiont genome is not strictly indicative of the actual abundance of the bacteria. A quantitative PCR would be needed to verify the abundance of each endosymbiont in each sample, as was

previously done in *B. tabaci* (Pan *et al.*, 2013; Su *et al.*, 2013b; Ghosh, 2016) and other hemipterans, such as aphids and psyllids (Enders and Miller, 2016; Morrow *et al.*, 2017).

The discrepancies between the PCR and HTS methods in our study can be explained by random sampling and/or primer mismatch, coverage and contamination. Firstly, only two Sudan Q1 individuals were *Cardinium*-positive in the PCR test, but both male and female appeared positive in the HTS screening, which could have been a result of chance. In contrast, 10 out of 10 Uganda ASL females were *Cardinium*-negative in the PCR test, while both male and the female picked for the HTS appeared positive. It seems unlikely that two individuals with a very rare and undetected infection would be selected from the same laboratory colony. It is more likely that there was an issue with the specificity of the primers targeting *Cardinium* 16S rDNA. To test this hypothesis, we assembled a contig by mapping the HTS reads from Uganda ASL and Sudan Q1 to a reference sequence of Cardinium 16S from MED-Q1 (AB981353) and then mapped the Card-F and Card-R primer sequences to the contig. Card-R matched the sequence with 100% identity, however, the best fit for Card-F primer contained two INDELs and two mismatches and would theoretically result in 729 bp amplicon (Appendix 14). We hypothesise that this mismatch caused the negative results in the PCR screening. In addition, we speculate that the two light bands observed for Sudan Q1 (Figure 4-10 A, *Cardinium*) were probably genuine products of the primers, because the mismatches occured in the middle of the Card-F sequence, while the first 10 bases at 5' end and 7 bases at 3' end match the template sequence well. Secondly, the very low amount of *Rickettsia*-related reads in Sudan Q1 and Uganda ASL was surprising and was potentially caused by the same coverage issue that caused an under-representation of *Portiera* sequences in Spain Q1 and Uganda ASL males. Lastly, the relatively large abundance of *Wolbachia*-related reads in only one read file from the Spain Q1 female might indicate a contamination issue in the HTS run.

Some of our results from the endosymbiont screening were well aligned with the literature. The two most consistent findings were the positive infections by *Hamiltonella* in Q1 (Gueguen *et al.*, 2010; Gnankiné *et al.*, 2013; Parrella *et al.*, 2014; Terraz *et al.*, 2014; Thierry *et al.*, 2015), and *Rickettsia* in Q2 (Parrella *et al.*, 2014; Terraz *et al.*, 2014; Thierry *et al.*, 2015). The previously reported co-infection by *Hamiltonella* and *Wolbachia* in Q1 (Terraz *et al.*, 2014; Thierry *et al.*, 2015) only occurred in our Sudan Q1 population and not in Spain Q1. The absence of *Arsenophonus* and *Hamiltonella* in Q1 and Q2, respectively (Gueguen *et al.*, 2010), as well as the absence of *Cardinium* in Israeli populations (Chiel *et al.*, 2007; Gueguen *et al.*, 2010) were congruent with our data.

In contrast, an infection by *Arsenophonus* was reported repeatedly in Q2 and ASL groups (Gueguen *et al.*, 2010; Gnankiné *et al.*, 2013; Parrella *et al.*, 2014; Terraz *et al.*, 2014; Thierry *et al.*, 2015), while none of our populations was infected by this endosymbiont.

We hypothesise that some of the mating interactions observed in this chapter could potentially be influenced by secondary endosymbiotic bacteria. Some endosymbionts residing within insects, notably *Wolbachia* and *Cardinium*, are known for manipulating the host's reproduction (Werren *et al.*, 2008; Mann *et al.*, 2017), and some endosymbionts were correlated with increased or decreased fitness of their hosts (Brumin *et al.*, 2011; Himler *et al.*, 2011; Xue *et al.*, 2012; Su *et al.*, 2013a; Raina *et al.*, 2015). The presence of both *Wolbachia* and *Cardinium* in Sudan Q1 and Uganda ASL suggests their potential role in the reproductive isolation of these two populations from all other studied populations, most notably the incompatibility between Sudan Q1 and Spain Q1. The potential manipulation of reproduction among members of the *B. tabaci* complex by *Cardinium* has already been proposed (Thierry *et al.*, 2011), although no evidence of *Cardinium*-induced incompatibility was found in intraspecies crosses within MED (Fang *et al.*, 2014). The involvement of *Wolbachia* in reproductive incompatibility within the *B. tabaci* complex has also been suggested (De Barro and Hart, 2000). Crossing trials using endosymbiont-positive and endosymbiont-negative individuals would help to elucidate the role of *Cardinium* and *Wolbachia* in the reproductive compatibility among groups of the MED putative species.

Another case of potential endosymbiont influence on mating interactions occurred in the back-crossing of the Spain Q1 x Israel Q2 hybrids. None of these two populations contained *Wolbachia* or *Cardinium*, however, Spain Q1 was infected by *Hamiltonella*, while Israel Q2 harboured abundant *Rickettsia*. A correlation between *Rickettsia* infection and a female-biased progeny has been previously reported (Himler *et al.*, 2011; Asiimwe *et al.*, 2014), which could also explain the higher female proportion (non-significant) in F_1 and F_2 offspring originating from Israel Q2 mothers passing on the *Rickettsia* infection. Such offspring, however, was less numerous than the one with Spain Q1 maternal origin, which is not in agreement with studies reporting a fitness advantage conferred by the *Rickettsia* infection (Himler *et al.*, 2011; Asiimwe *et al.*, 2014). In contrast, our results are on par with the report of higher fitness of *Hamiltonella*-positive whiteflies (Su *et al.*, 2013a), because in both F_1 and F_2 generation the offspring with Spain Q1 maternal origin, and hence with inherited *Hamiltonella*, was more numerous (non-significant in F_1; $P<0.01$ in F_2).

This chapter has shown that, in addition to genetic differences, there are also reproductive barriers between populations from the Mediterranean Basin and sub-Saharan Africa, as well as between the sub-Saharan populations, suggesting that there is more than one biological species within the MED putative species. In addition, we showed that all four populations contain a different combination of endosymbiotic bacteria, which can affect the reproduction and life history traits of their hosts. The next chapter will deal with the question whether these population share the same affinity for various host plants or whether they have evolved different biological adaptations.

5. Host-plant range and performance differences among MED populations[3]

5.1 Introduction

It is important to understand the ecology of a pest organism to design efficient management strategies to control the pest and, if it is also a vector, the viruses it transmits. In the field, many interactions among whiteflies, their predators, symbionts and pathogens can be affected by the plant host (Inbar and Gerling, 2008). Knowledge about the range of plants suitable for pest reproduction and the effect of the plants on pest's life history traits helps make predictions about its population dynamics in the field. It is, therefore, particularly important for invasive pests, such as MEAM1 and MED putative species within the *B. tabaci* complex. However, specific and accurate information about host range of individual species within the *B. tabaci* complex is limited.

Until the early 1990's*, B. tabaci* (*sensu* Russell, 1957) was considered to be a single polyphagous species with 506 plant species listed as hosts (Mound and Halsey, 1978; Greathead, 1986). As early as in the 1950s, however, different host-associated races within *B. tabaci* were distinguished. Examples include the Sida and Jatropha races in Puerto Rico (Bird, 1957) and okra and cassava biotypes in sub-Saharan Africa (Burban *et al.*, 1992; Omondi *et al.*, 2005). Following the recognition that *B. tabaci* is a cryptic species complex and assignment of *B. tabaci* populations into different putative species, knowledge on the host range of individual species remains to be reassessed.

MEAM1 has received the most attention as it was the first reported globally invasive species from the *B. tabaci* complex (Toscano *et al.*, 1998). Since the beginning of the 2000's the host use of MEAM1 has often been compared with MED, because MED was the second invasive species reported from the complex (Chu *et al.*, 2006; Dennehy *et al.*, 2010). Subsequently, in the light of new findings about the genetic and biological heterogeneity among different groups within MED, a gap in our knowledge about their host-plant relations became evident.

There have been some reports indicating host-plant range differences among the groups of the putative MED species. Results from a field survey in Uganda suggested that sympatric ASL and Q1 populations exhibited different preference for host-plants, because Q1 was mostly found on *Nicotiana tabacum*, while the majority of ASL occurred on *Cucurbita moschata*, *Sida acuta*, *Ipomoea batatas*, *Pavonia urens* and *S. melongena*

[3] This chapter formed a part of manuscript Vyskočilová *et al*. accepted for publication in 2019; page 244.

(Mugerwa, 2018). A similar trend was observed by Sseruwagi *et al.* (2005), where Q1 (referred to as “Ug5”) was only collected from *N. tabacum*, while ASL (referred to as “Ug4”) occurred on *Cucumis sativus*, *Cucurbita pepo*, *S. melongena*, *Leonotis nepetifolia* and *P. urens*. Different host preferences were also reported for MED populations Q1 and Q2 in Italy; individuals from Q2 were more predominant on solanaceous hosts, while Q1 adults were collected from members of Convolvulaceae, Malvaceae and Asteraceae (Parrella *et al.*, 2014).

There are, however, limitations in assessing the host range of an herbivore insect species exclusively based on field collections data. If only adult whiteflies are collected and analysed, the results do not provide certainty that the plant they were collected from is a good reproductive host for that whitefly species. This problem can be circumvented by collecting immature stages developing on the plant, which gives a clearer indication of the plant’s suitability for the pest’s reproduction. However, a controlled experiment is needed to determine whether the pest can complete a full life cycle on the host and to quantify and compare its life history traits.

The aim of this chapter was to assess the performance of the MED populations on a variety of host plants and compare their host-plant ranges. The research objectives of this chapter were to: (1) compare life history traits (oviposition rate, adult survival time, number and sex ratio of the progeny) among the MED populations on a variety of plants, (2) classify the plants based on their suitability as reproductive hosts and compare the host ranges among the MED populations, and (3) carry out preliminary work for a transcriptomic analysis to study the molecular mechanisms underpinning differential host-plant utilisation among the MED populations.

5.2 Material and methods

5.2.1 Growing plants and rearing insects

Plant species used in the host-plant performance bioassays are listed in Table 5-1. Plants were grown in a whitefly-free room at 28 ± 2^0C, 50–60% relative humidity and a 14:10 light:dark photoperiod. After plantlets rooted in loam-based compost (J. Arthur Bower's John Innes No. 2, UK) reached 10–15 cm above soil level, they were individually enclosed in Lock&Lock whitefly-proof cages (Wang *et al.*, 2011a) with two additional side openings in the upper container covered by 160 µm nylon mesh (Figure 3-1 B–D). The plants were watered twice a week. Colonies of *B. tabaci* (Table 3-1) were reared as described in section 3.2.1.

Table 5-1: Plant hosts used in the study with the information on the taxonomic families and propagation methods.

Host	Species name	Variety	Family	Prop.
Aubergine	*Solanum melongena*	Black Beauty	Solanaceae	S
Bean	*Phaseolus vulgaris*	Tendergreen	Fabaceae	S
Borecole	*Brassica oleracea*	Dwarf Green Curled	Brassicaceae	S
Cassava	*Manihot esculenta*	Colombian	Euphorbiaceae	TC
Chard	*Beta vulgaris*	Rhubarb Red	Amaranthaceae	S
Cotton	*Gossypium hirsutum*	-	Malvaceae	S
Mint	*Mentha piperita*	Peppermint	Lamiaceae	C
Okra	*Abelmoschus esculentus*	Clemson Spineless	Malvaceae	S
Pepper	*Capsicum annuum*	Californian Wonder	Solanaceae	S
Squash	*Cucurbita pepo*	All Green Bush	Cucurbitaceae	S
Sweet potato	*Ipomoea batatas*	Naspot11, Beauregard	Convolvulaceae	C
Tobacco	*Nicotiana tabacum*	Izmir, Basma	Solanaceae	S
Tomato	*Solanum lycopersicum*	Moneymaker	Solanaceae	S

Prop. = Propagation method
S = seeds, C = cuttings, TC = tissue culture.

5.2.2 Photographs of abaxial leaf surface

For capturing leaf surface photographs a separate set of plants were grown as described in 5.2.1. Detached leaves in different growth stages were placed under the stereomicroscope (Nikon SMZ18) with an attached DSLR camera (Nikon D5300) and illuminated with a LED light source (Photonic Optics F3000). Photographs were taken with automatic camera settings at 20x and 50x magnification.

5.2.3 Oviposition and leaf preference assay

All plants from Table 5-1, except mint, were used for the oviposition assay. Three to six experimental replicates were used for each combination of plant and whitefly population. The leaves were removed, leaving only three leaves in different stages: (i) “mature”, the first true leaf, (ii) “young”, the last fully expanded leaf, and (iii) “immature”, the youngest, not fully expanded leaf. The exceptions were bean and chard in which only two leaves of similar age were used and hence were excluded from the leaf age preference comparison.

Whitefly adults used in the experiments were collected from a synchronized colony. To obtain such colony, aubergine plants carrying large numbers of nymphs were taken from each core colony, devoid of all adults and put in a separate cage. After 24–48 h, adults that have emerged on these plants (“young adults” hereafter) were collected into glass tubes (0.5 cm in diameter, 6.0 cm long) and their sex was determined using a stereomicroscope. Thirty young adults (15 males and 15 females) were released onto each experimental plant, except for cotton, okra and pepper, onto which 50 adults (25 males and 25 females) were released. The parental adults were removed from the plant 48 h after release. The eggs laid on the abaxial leaf side were counted using a stereomicroscope and a photograph of the leaf was taken with a grid paper background to allow size calibration in surface area calculation. The oviposition rate was expressed as the number of eggs laid per female in 48 h.

To determine the oviposition preference for leaves of different age, numbers of eggs/female/48 h were further divided by the leaf surface area. The areas were calculated based on the leaf photographs using the online tool SketchAndCalc (Dobbs, 2011), www.sketchandcalc.com) (Appendix 15). The eggs/female/48 h/cm^2 values were multiplied by an appropriate number assigned to the leaf stage (mature = 1, young = 2, immature = 3) and summed. Leaf preference scores were calculated for each plant by dividing the sum of eggs/female/48 h/cm^2 values by the sum of eggs/female/48 h/cm^2 x leaf stage number.

5.2.4 Survival, fecundity and progeny sex ratio assay

Data on adult survival, progeny counts and progeny sex ratio were recorded from a separate set of plants from the oviposition assay. All plants from Table 5-1 were used, with three to six replicates for each combination of plant and whitefly population. Thirty young adults (15 females and 15 males) from a synchronized colony (detailed in section 5.2.3) were released onto each experimental plant. The numbers of surviving adults were

monitored during the period of 16–19 days, after which the parental generation was collected to avoid interference with the count of the F_1 generation. Adults of the F_1 generation from all replicates were collected at three time-points of 21, 28 and 35 days post-release. The collected F_1 adults were frozen at -20°C in order to be counted and their sex determined later. The total cumulative count of F_1 adults from each replicate was used for the analysis of progeny counts and sex ratio.

5.2.5 Classification of hosts

We classified the 13 plants into three categories based on their suitability as a reproductive host for each whitefly population. The categories were delineated based on gaps in distribution of progeny counts, identified by plotting the mean progeny counts in a histogram using MS Excel. Differences between the ranges of favourable plants among whitefly populations were visualised by a Venn diagram.

5.2.6 Statistical analyses

The statistical analyses and plots were carried out in R (R Core Team, 2013). Whitefly performance on different plant hosts was analysed at four levels: (i) female oviposition rate, (ii) adult whitefly survival, (iii) F_1 progeny counts and (iv) proportion of females in the F_1 progeny.

The counts of eggs and F_1 adults were analysed by a generalised linear model with a negative binomial error distribution and a log link using the MASS library (Venables and Ripley, 2002). The estimated mean survival times were computed by survival analysis with censoring. The data file contained information on time of death in days for each whitefly (denoted by “1”) in each replicate from day 1 to day 16–19; the whiteflies surviving beyond that period were censored (denoted by “0”). The survival regression model was fitted by survreg function assuming a Weibull distribution. The survival curves were plotted using the raw data and survfit function. For the proportion of F_1 female progeny, a generalized linear model with binomial error distribution and logit link of the proportional data was used.

Multiple comparisons were performed by Tukey test (Tukey, 1949) using the multcomp package (Hothorn *et al.*, 2008) and significant differences were demonstrated by compact letter display. Comparisons of egg counts, progeny counts, survival times and female proportions between whitefly populations were conducted for each host individually, except for oviposition preference for different leaf stages, in which all plants were included in one comparison.

5.2.7 Silver-leafing bioassay

The capacity of the Uganda ASL population to induce the squash silver-leafing symptoms was tested by releasing 30–60 adults on squash plants *Cucurbita pepo* 'All Green Bush', individually enclosed in Lock&Lock cages in five replicates. As a positive control, two replicates with 60 MEAM1 adults from Peru were set up using the same protocol. After two weeks the adults were collected and plants were visually assessed for leaf silvering symptoms, for up to five weeks post-infestation.

5.2.8 Sample collection for transcriptomics

Samples were collected for a future comparative transcriptomic study aimed at identifying differentially expressed genes on favourable and challenging hosts. Populations chosen for this study were Spain Q1 and Uganda ASL on three plants: one favourable for both, a second more favourable for Spain Q1, and a third more favourable for Uganda ASL.

Whiteflies in five different life stages were collected (Figure 5-1). To generate the sample material, 700–3000 adults from a core colony were released onto four or five large plants in a BugDorm cage. These adults were collected after 24–48 h. Egg-bearing plants were transferred into a clean cage and the development of the immature stages was monitored periodically. The eggs and 2nd instars were collected from detached leaves, but the 3rd and 4th instars were collected from leaves attached to the plant to avoid wastage of the whitefly material. The leaves were viewed under a stereomicroscope, specimens were picked up with 0.30 x 60 mm needles (Seirin, US) while minimising the damage to the leaf and the specimens, and transferred into a 1.5ml Safe-Lock microtube (Eppendorf, UK) in a dry ice bath (-78.5^{0}C). One sample consisted of about 200 eggs, 100 nymphs or 100 adults with unspecified sex ratio, each contained within one or more tubes. As soon as the collection was finished, the tubes were frozen in liquid nitrogen (-195.79°C) for 30–60 s, immediately transferred into a chilled box and stored at -80°C.

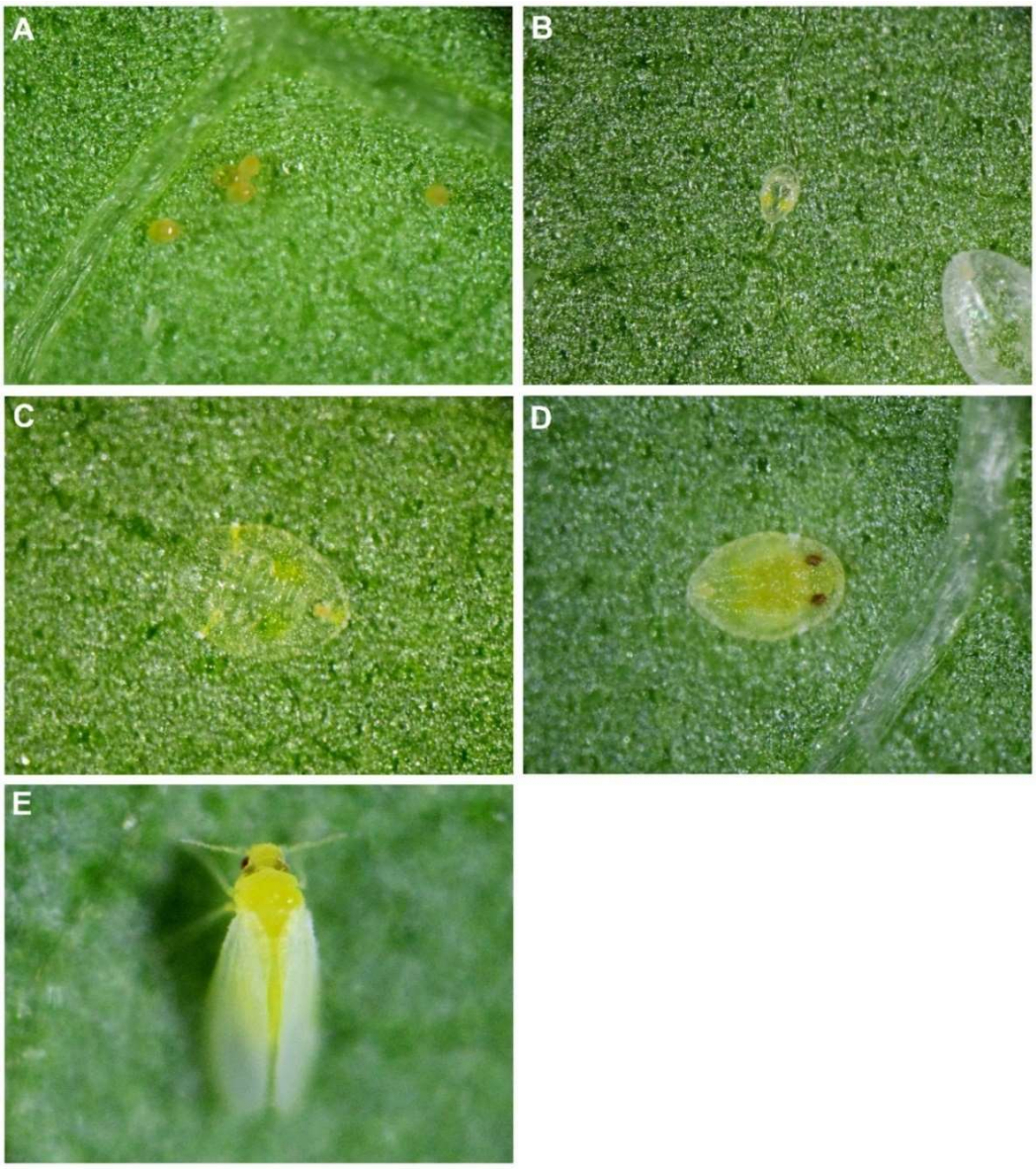


Figure 5-1: Whitefly life stages collected for RNA extraction. **A**: eggs, **B**: 2nd instar (with empty 4th instar exuviae in the bottom right corner), **C**: 3rd instar, **D**: 4th instar, **E**: adult. Magnification 80x. Photographs taken with Uganda ASL whiteflies on okra leaves.

5.2.9 RNA extraction and quality assessment

A trial extraction of total RNA was performed to determine the efficiency of the collection method and evaluate the amount of material needed to get a sufficient amount of good quality RNA (2–5 µg) for sequencing. Six samples from section 5.2.8 were selected to represent each of the five developmental stages and each of the three host plants (Table 5-2). The kit used for RNA extraction was Isolate II RNA Mini Kit (Bioline, UK). The frozen whitefly samples were transferred from -80°C into a dewar with liquid nitrogen and homogenised with a plastic pestle while immersed in a dry ice bath. After 15–30 s of grinding, 350 µl of Buffer RLY (from Isolate II RNA Mini Kit) and 3.5 µl 2-mercaptoethanol was added and sample continued to be ground and mixed for 30–60 s. The mixture was

filtered and RNA extracted using the Isolate II RNA Mini Kit spin columns, as per the manufacturer's instructions. The RNA bound on the silica membrane was eluted in 2 x 50 µl, which were then combined. The extracted RNA in 100 µl elution was purified by the RNA Clean & Concentrator Kit (Zymo Research, UK) as per the manufacturer's instructions. Final elution was done in 2 x 30 µl, combined into 60 µl of purified RNA.

Table 5-2: Summary of the samples used for RNA extraction. Each sample was contained in one individual tube.

Sample	Life stage	Population	Host plant	Specimens in tube
1	Eggs	Spain Q1	Pepper	200
2	Eggs	Spain Q1	Cotton	200
3	2. instar	Spain Q1	Pepper	98
4	3. instar	Uganda ASL	Okra	100
5	4. instar	Uganda ASL	Okra	60
6	Adults	Spain Q1	Pepper	100

The yield and integrity of RNA was assessed by employing three methods, (i) agarose gel electrophoresis, (ii) spectrophotometry and (iii) fluorometrics. The gel electrophoresis was performed in 2% agarose gel with 10% RedSafe staining solution. After denaturation of the RNA in a 70°C water bath for 2 min, 10 µl of eluate with 2 µl of orange loading dye was loaded from each sample. The electrophoresis ran for 1h 15min at 8 V/cm. The spectrophotometric quantification was done using NanoDrop 2000 (Thermo Scientific, UK). Denatured RNA eluate (1 µl) was used to obtain a reading, each sample was repeated twice and the values averaged. Molecular grade water (Sigma-Aldrich, UK) was used as a blank. The fluorometric sizing and quantitation were done on the Bioanalyzer 2100 platform (Agilent Technologies, US). Samples were run through the Agilent chip electrophoresis instruments as per manufacturer's instructions. Plant RNA assay was selected in the settings due to running together with plant samples. The outcomes of the analysis were electropherogram plots showing fragment size distribution and concentration, but an accurate RNA Integrity Number (RIN) could not be calculated. The yield was calculated as concentration (ng/µl) x 60 µl divided by 1000 to get the total estimated amount of extracted RNA in µg.

5.3 Results

6 5.3.1 Comparison of abaxial leaf surfaces

The morphology of abaxial side of leaves was observed under a stereomicroscope. The full set of photographs taken at 50x magnification can be found in Appendix 16. We observed that in some plants leaves of all ages had a smooth surface, while in others the hairiness varied with leaf stage.

Mint was a specific case because its leaves were not hairy, but carried glandular trichomes that secrete essential oils (Maffei *et al.*, 1989) (Figure 5-2). Density of these glands was higher on younger leaves (Figure 5-2 B, C).

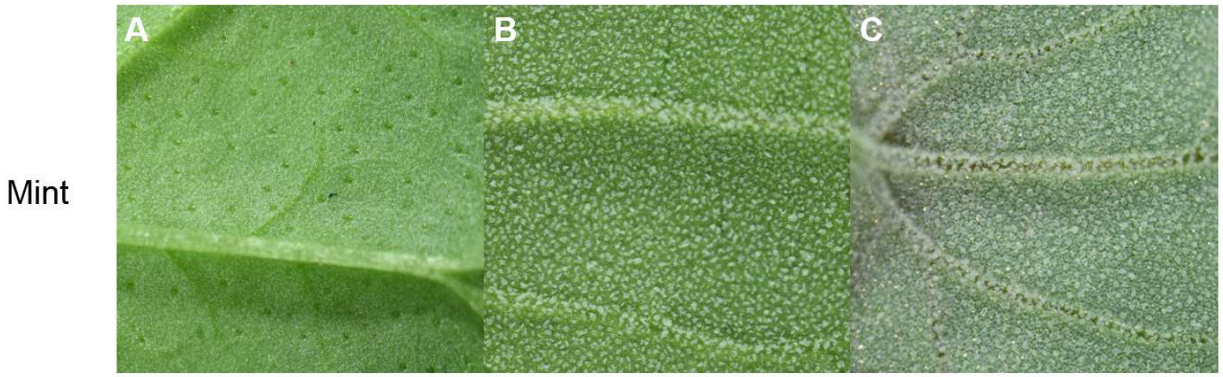


Figure 5-2: Abaxial surface of mint leaf magnified 20x. **A** = old leaf, **B** = young leaf, **C** = immature leaf.

Plants with glabrous leaves included borecole, cassava, chard, cotton, pepper and sweet potato (Figure 5-3). The remaining six plants (aubergine, bean, okra, squash, tobacco and tomato) shared a common pattern, in which the old leaf was the smoothest, and the young and immature leaves were hairier (Figure 5-4).

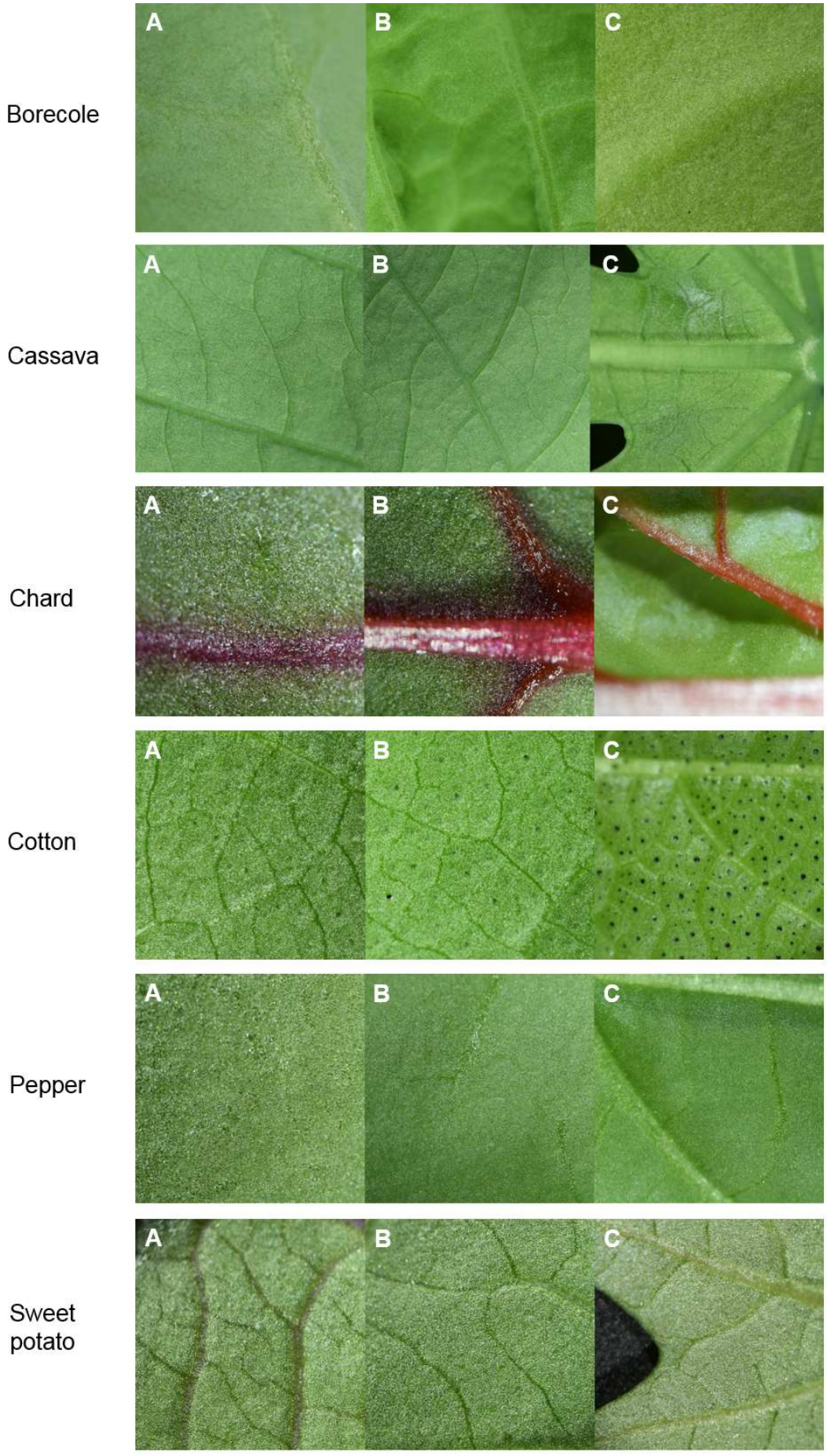


Figure 5-3: Abaxial surface of glabrous leaves magnified 20x. **A** = old leaf, **B** = young leaf, **C** = immature leaf.

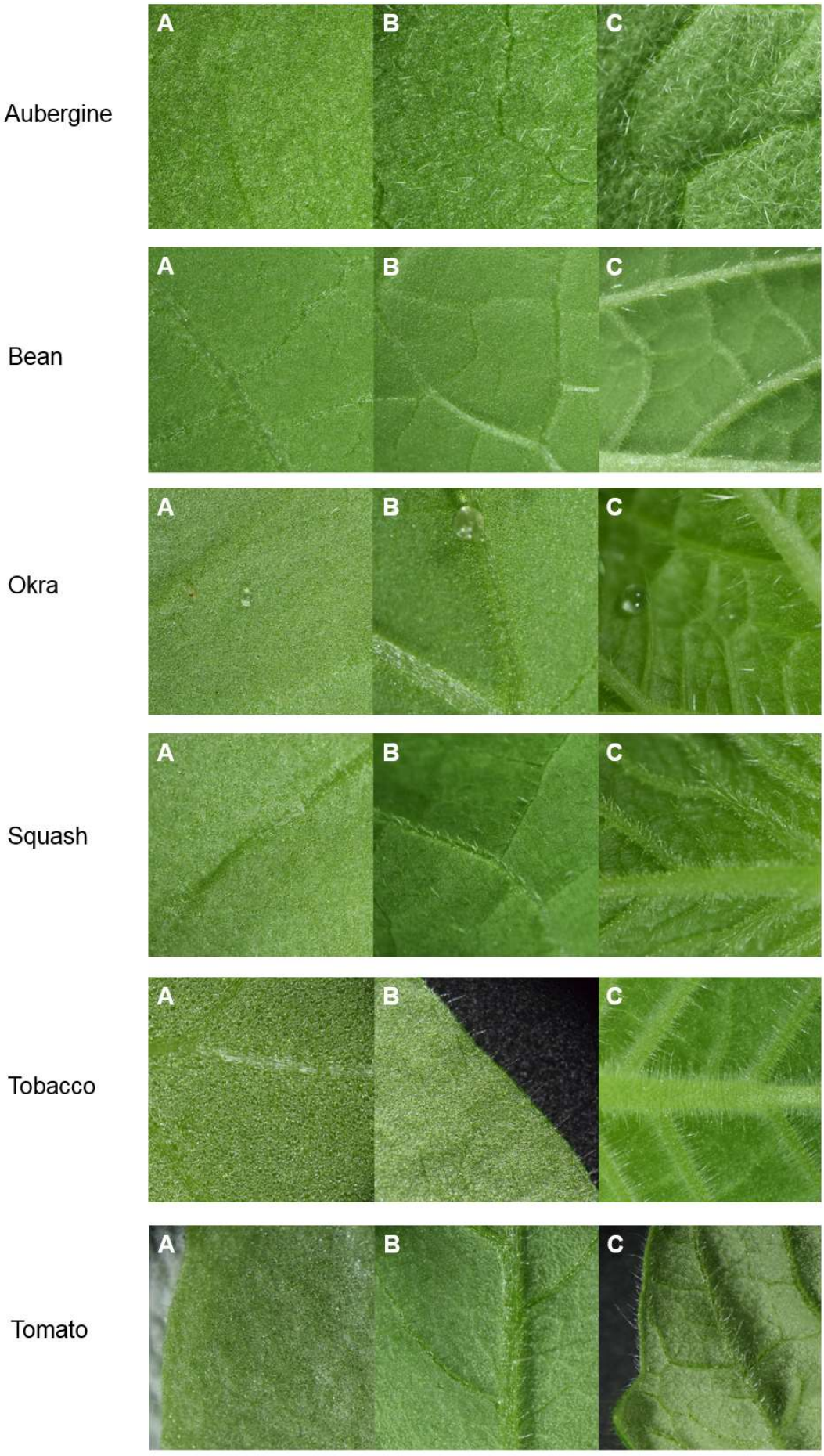


Figure 5-4: Abaxial surface of hairy leaves magnified 20x. **A** = old leaf, **B** = young leaf, **C** = immature leaf.

5.3.2 Oviposition rate

The oviposition rate was expressed as number of eggs laid by one female in 48 h. Significant differences (P<0.05) between populations occurred on 6 out of 12 plants: bean, cassava, cotton, okra, pepper and tomato (Figure 5-5 and Appendix 17).

The biggest differences among populations occurred on cassava, bean, okra and cotton. Israel Q2 females laid 20 times more eggs than Uganda ASL females on cassava (P<0.01). Spain Q1 and Uganda ASL females laid respectively 3.3 (P<0.05) and 3.9 (P<0.01) times more eggs than Israel Q2 females on bean. On okra, Uganda ASL females laid 2.6–2.8 times more eggs (P<0.01) compared with females from the three MED populations. Another marked difference occurred on cotton where Spain Q1 females laid 2.6 times more eggs than Israel Q2 (P<0.05).

Collectively for all four populations, the highest numbers of eggs were laid on sweet potato, aubergine and bean (Table 5-3). The lowest numbers of eggs occurred on chard, cassava and tobacco.

The mean oviposition rate across all plants revealed that Uganda ASL and Spain Q1 females were more fertile than Sudan Q1 and Israel Q2 females. The highest mean fertility was achieved by Spain Q1 on aubergine and bean, and Uganda ASL on bean and okra (Table 5-3).

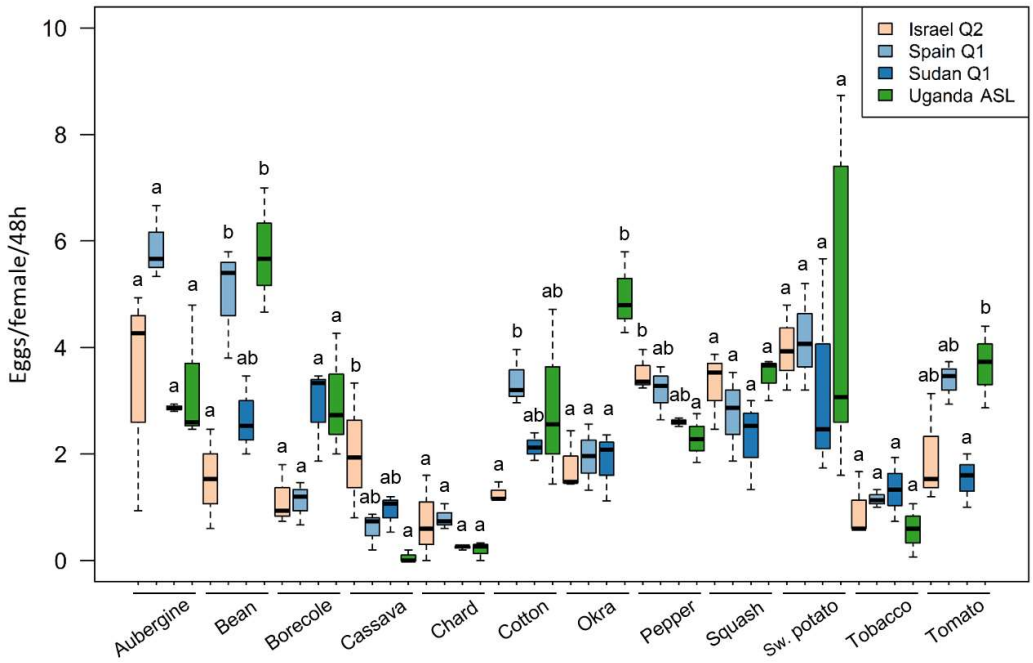


Figure 5-5: Oviposition rates of females from four whitefly populations on 12 host plants. Box length represents the interquartile range, whiskers express the full range and black bar corresponds to the median. Different letters above the boxes indicate significant differences between whitefly populations on each plant individually (Tukey's test, P<0.05).

Table 5-3: Mean oviposition rates (eggs/female/48 h) ± standard errors of four whitefly populations on 12 host plants.

Host	Israel Q2	Spain Q1	Sudan Q1	Uganda ASL
Aubergine	3.4 ± 1.1	5.9 ± 1.4	2.9 ± 1.0	3.3 ± 1.0
Bean	1.5 ± 0.7	5.0 ± 1.3	2.7 ± 0.9	5.8 ± 1.4
Borecole	1.2 ± 0.6	1.1 ± 0.6	2.9 ± 1.0	3.0 ± 1.0
Cassava	2.0 ± 0.8	0.6 ± 0.4	0.9 ± 0.6	0.1 ± 0.1
Chard	0.7 ± 0.5	0.8 ± 0.5	0.2 ± 0.3	0.2 ± 0.3
Cotton	1.3 ± 0.6	3.4 ± 1.1	2.1 ± 0.8	2.9 ± 1.0
Okra	1.8 ± 0.8	1.9 ± 0.8	1.9 ± 0.8	5.0 ± 1.3
Pepper	3.5 ± 1.1	3.2 ± 1.0	2.6 ± 0.9	2.3 ± 0.9
Squash	3.3 ± 1.0	2.8 ± 1.0	2.3 ± 0.9	3.5 ± 1.1
Sweet potato	4.0 ± 1.2	4.2 ± 1.2	3.3 ± 1.0	4.4 ± 0.9
Tobacco	1.0 ± 0.6	1.2 ± 0.6	1.3 ± 0.7	0.6 ± 0.4
Tomato	2.0 ± 0.8	3.4 ± 1.1	1.5 ± 0.7	3.7 ± 1.1
Mean	2.1 ± 0.3	2.8 ± 0.3	2.1 ± 0.2	3.0 ± 0.3

The comparison of egg distribution between mature, young and immature leaves revealed differences in female oviposition preferences between hosts (Figure 5-6). On tobacco and tomato the preference to oviposit on older leaves was significantly higher ($P<0.05$) than on all other plants. In contrast, young leaves were strongly preferred on sweet potato and pepper. The leaf preference did not differ significantly among the whitefly populations ($P>0.5$).

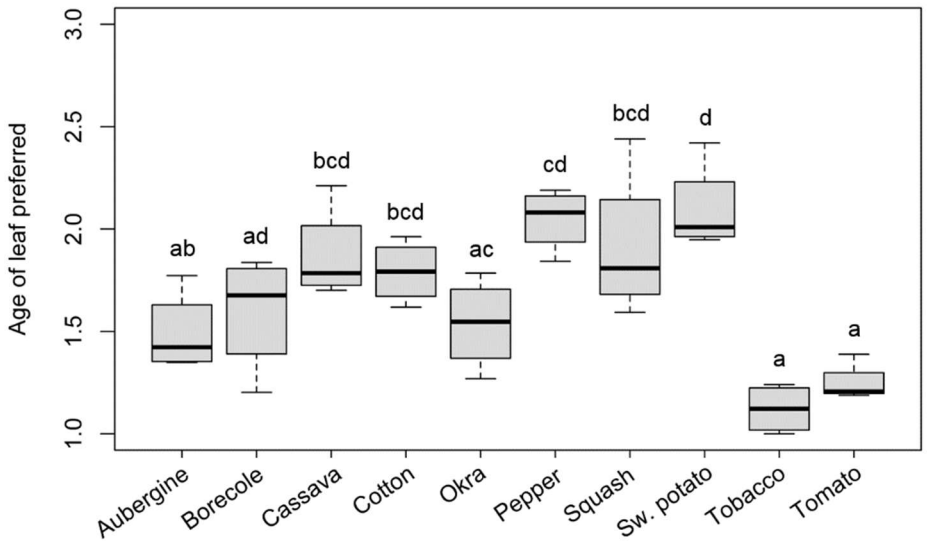


Figure 5-6: Plant scores representing female oviposition preference for leaves of different age (mature = 1, young = 2, immature = 3). Box length represents the interquartile range, whiskers express the full range and black bar corresponds to the median. Different letters above the boxes indicate significant differences between plants (Tukey's test, $P<0.05$).

5.3.3 Adult survival times and curves

The average survival times were estimated in the survival analysis as the number of days from release on the plant to death of an individual. Significant differences (P<0.05) among populations occurred on all 13 plants (Figure 5-7 and Appendix 18).

For all populations, the highest adult survival occurred on sweet potato, cotton and aubergine. The shortest adult survival times were observed on cassava, chard and tomato. Spain Q1 and Israel Q2 lived longer on average across all plants, compared with Sudan Q1 and Uganda ASL (Table 5-4). Despite the low mean survival times across plants, Uganda ASL adults feeding on sweet potato reached an exceptionally high estimated life span of 47.4 ± 8.6 days. Relatively long survival times (about 32 days) on sweet potato were also observed for Spain Q1 and Sudan Q1. A similar life span (28–30 days) was observed for Israel Q2 on cotton and Spain Q1 on aubergine and cotton.

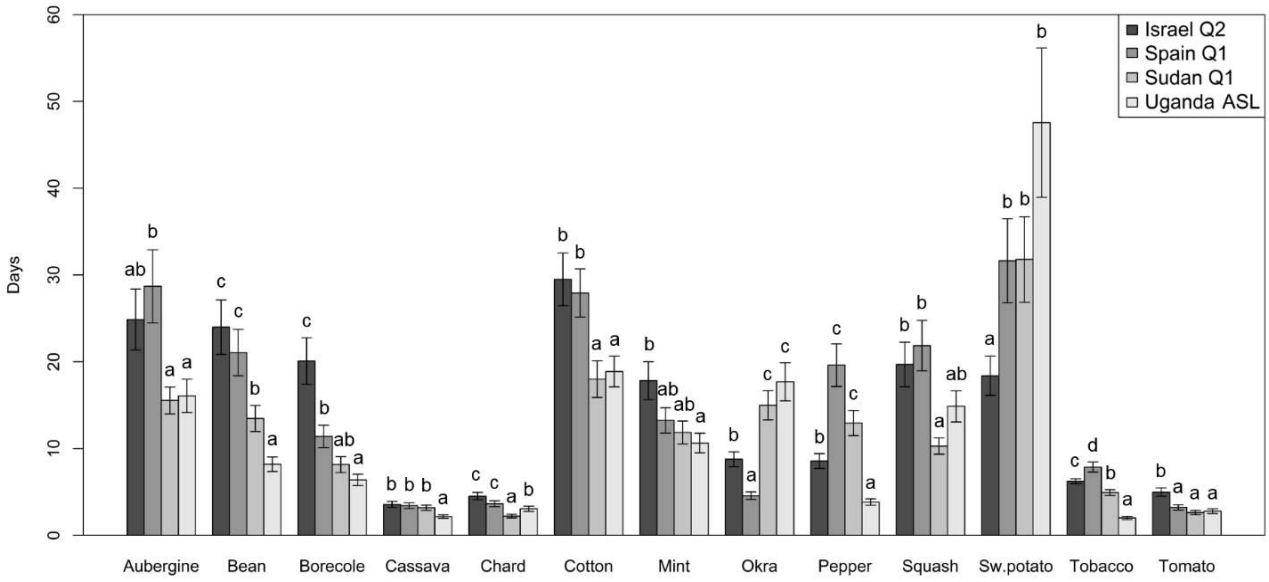


Figure 5-7: Mean adult survival times of the four whitefly populations on 13 host plants. Different letters above the error bars indicate significant differences between whitefly populations on each plant individually (Tukey's test, P<0.05).

Table 5-4: Mean adult survival times (days) ± standard errors for four whitefly populations on 13 host plants.

Host	Israel Q2	Spain Q1	Sudan Q1	Uganda ASL
Aubergine	24.8 ± 3.5	28.7 ± 4.2	15.6 ± 1.6	16.1 ± 1.9
Bean	24.0 ± 3.5	21.0 ± 2.7	13.5 ± 1.5	8.2 ± 0.8
Borecole	20.1 ± 2.7	11.4 ± 1.3	8.2 ± 0.9	6.4 ± 0.7
Cassava	3.6 ± 0.3	3.4 ± 0.3	3.2 ± 0.3	2.2 ± 0.2
Chard	3.6 ± 0.4	3.6 ± 0.3	2.2 ± 0.2	3.1 ± 0.3
Cotton	29.5 ± 3.0	27.9 ± 2.8	18.0 ± 2.1	18.9 ± 1.8
Mint	17.8 ± 2.2	13.3 ± 1.5	11.8 ± 1.3	10.6 ± 1.1
Okra	8.8 ± 0.9	4.6 ± 0.4	15.0 ± 1.7	17.7 ± 2.2
Pepper	8.6 ± 0.9	19.6 ± 2.5	12.9 ± 1.4	3.8 ± 0.4
Squash	19.7 ± 2.6	21.8 ± 2.9	10.3 ± 0.9	14.9 ± 1.8
Sweet potato	18.4 ± 2.3	31.6 ± 4.8	31.7 ± 4.9	47.4 ± 8.6
Tobacco	6.2 ± 0.3	7.9 ± 0.6	4.9 ± 0.3	2.0 ± 0.2
Tomato	5.0 ± 0.5	3.2 ± 0.3	2.6 ± 0.2	2.8 ± 0.3
Mean	11.3 ± 0.4	12.9 ± 0.5	9.6 ± 0.3	9.2 ± 0.3

The dynamics of population decline on each plant species was demonstrated by survival curves (Figure 5-8). The initial death rate was the fastest on tomato and cassava, on which population sizes dropped to 50% and below within 2.5 and 3.3 days on average, respectively. On tobacco and chard, the population decline to 50% occurred after 4 days on average. The continued death rate was the fastest on cassava, chard and tomato, on which the populations declined to or below 10% in 4.3, 5.5 and 8.6 days, respectively.

The patterns of adult survival differed between Uganda ASL and other three populations. For example, the ASL population on tobacco dropped to 10% within two days, while for other populations the same decline took 13 or more days. Similarly, on pepper the ASL population declined by two thirds within two days, while the same decline took 11 or more days for Q1 and Q2 populations. Bean was favourable for the survival of all Q1 and Q2 populations, but ASL population dropped below 50% within three days. In contrast, about 40% of ASL population survived for 18 days on okra, but in the other three populations only 0–10% survived the same conditions. Plants favourable for adult survival of all four populations were cotton, aubergine and sweet potato.

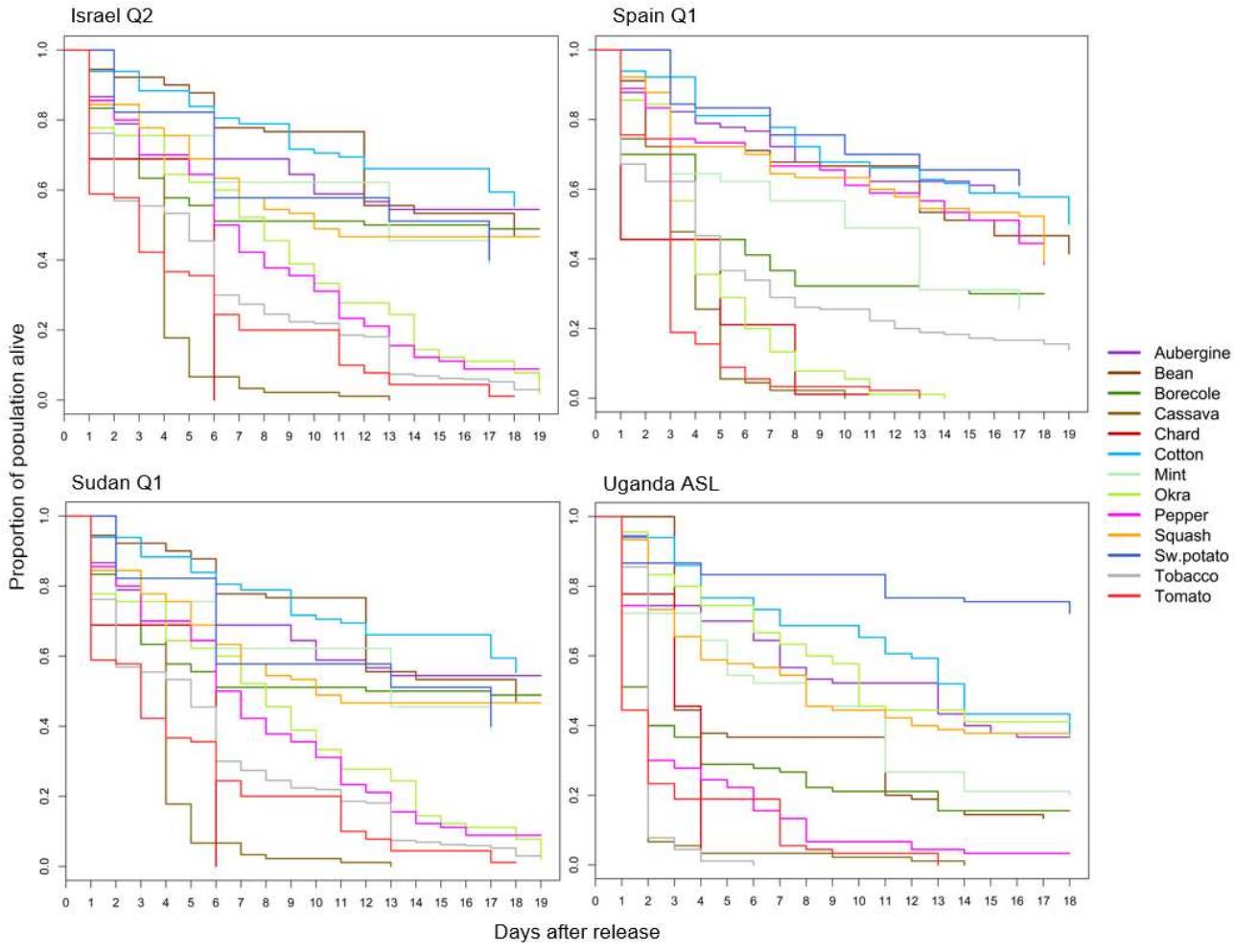


Figure 5-8: Survival curves showing whitefly population survival dynamics on 13 host plants over a period of 16–19 days post-release.

5.3.4 F_1 progeny counts

The progeny counts were expressed as cumulative sums of adults emerged up to 35 days after the parental release in each replicate. Significant differences ($P<0.05$) among populations occurred on 10 out of 13 plants (Figure 5-9 and Appendix 19). The three exceptions were cassava, chard and tomato, on which either no, or only a single, F_1 adult emerged for all populations (Table 5-5).

The most significant differences among populations occurred on okra, bean, pepper and squash. Okra was an exceptionally suitable host for Uganda ASL, as the number of progeny produced was significantly larger than that of Israel Q2 (29.9-fold, $P<0.001$) and Sudan Q1 (3.7 fold, $P<0.05$). Spain Q1 failed to produce any offspring on okra, but was the most successful population on bean, producing 56.3 times more progeny than Uganda ASL ($P<0.001$) and 2.9 times more than Sudan Q1 (non-significant). Israel Q2 failed to develop on bean. Pepper was only suitable for the Q1 populations. Squash was very favourable for Uganda ASL; the numbers of progeny produced were 65.4 and 13.6 times greater than for Israel Q2 and Sudan Q1, respectively ($P<0.001$ and $P<0.05$, respectively). The hosts suitable for the reproduction of all four populations were sweet potato, cotton, and aubergine.

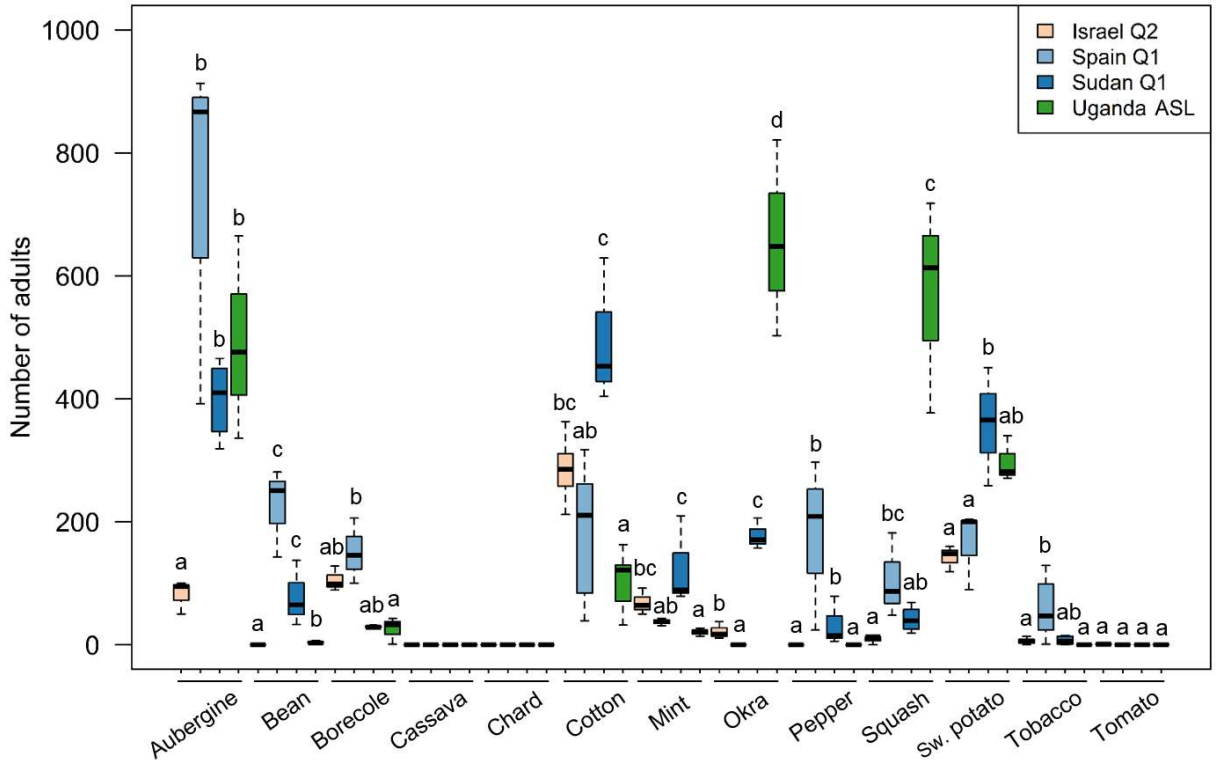


Figure 5-9: Numbers of F_1 adults in progeny produced by four whitefly populations on 13 host plants. Box length represents the interquartile range, whiskers express the full range and black bar corresponds to the median. Different letters above the boxes indicate significant differences between whitefly populations on each plant individually (Tukey's test, $P<0.05$).

The largest numbers of progeny across populations were produced on aubergine, cotton, sweet potato and okra (Table 5-5). In addition to cassava and chard, on which no adults were produced, the lowest number of F_1 adults emerged on tomato, tobacco and pepper.

Uganda ASL was the most fecund population on average (Table 5-5), but also displayed the most dramatic differences between plants. It failed to produce offspring on five host plants (cassava, chard, pepper, tobacco, and tomato), and on further three plants (bean, mint and borecole) produced only a few offspring. Spain Q1 and Sudan Q1 had more consistent and moderately high fecundity rates. Israel Q2 was appreciably less fecund in comparison to the other three populations (2.5-fold less than Uganda ASL and about 2-fold less than Spain and Sudan Q1).

Table 5-5: Mean F_1 progeny counts (produced by 15 pairs per replicate) ± standard errors for four whitefly populations on 13 host plants.

Host	Israel Q2	Spain Q1	Sudan Q1	Uganda ASL
Aubergine	81.7 ± 23.7	724.0 ± 205.3	398.4 ± 87.7	492.3 ± 139.8
Bean	0.0 ± 0.0	225.0 ± 64.2	78.3 ± 22.7	4.0 ± 1.6
Borecole	105.3 ± 30.4	150.7 ± 43.2	29.7 ± 9.0	25.7 ± 7.8
Cassava	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Chard	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Cotton	285.8 ± 57.6	187.2 ± 37.8	495.3 ± 140.7	103.6 ± 23.1
Mint	68.7 ± 20.0	37.3 ± 11.1	126.0 ± 36.2	20.7 ± 6.4
Okra	22.0 ± 6.8	0.0 ± 0.0	178.0 ± 50.9	657.3 ± 186.5
Pepper	0.0 ± 0.0	176.7 ± 50.5	33.3 ± 10.0	0.0 ± 0.0
Squash	8.7 ± 3.0	105.7 ± 30.5	41.8 ± 10.7	569.3 ± 161.6
Sweet potato	142.3 ± 40.8	164.7 ± 47.2	358.7 ± 102.0	297.7 ± 84.8
Tobacco	6.1 ± 1.5	57.8 ± 12.0	7.0 ± 1.8	0.0 ± 0.0
Tomato	1.0 ± 0.6	0.3 ± 0.3	0.0 ± 0.0	0.3 ± 0.3
Mean	66.2 ± 20.2	138.3 ± 42.7	135.5 ± 41.8	163.9 ± 53

5.3.5 Proportion of females in F_1 progeny

Significant differences (P<0.05) in the proportion of female progeny among populations occurred on five out of 13 plants: aubergine, cotton, mint, pepper and sweet potato (Figure 5-10 and Appendix 20).

The largest differences in proportion of female F_1 progeny occurred on cotton (30.3 percentage points between Sudan Q1 and Uganda ASL) and pepper (29.4 percentage points between Spain Q1 and Sudan Q1) (Table 5-6), and were both statistically significant (P<0.001). In general, we observed the highest percentages of females on bean and aubergine, while the lowest occurred on mint and sweet potato. The mean proportion of females produced by the four populations ranged from 33.16% (Spain Q1) to 51.0% (Sudan Q1).

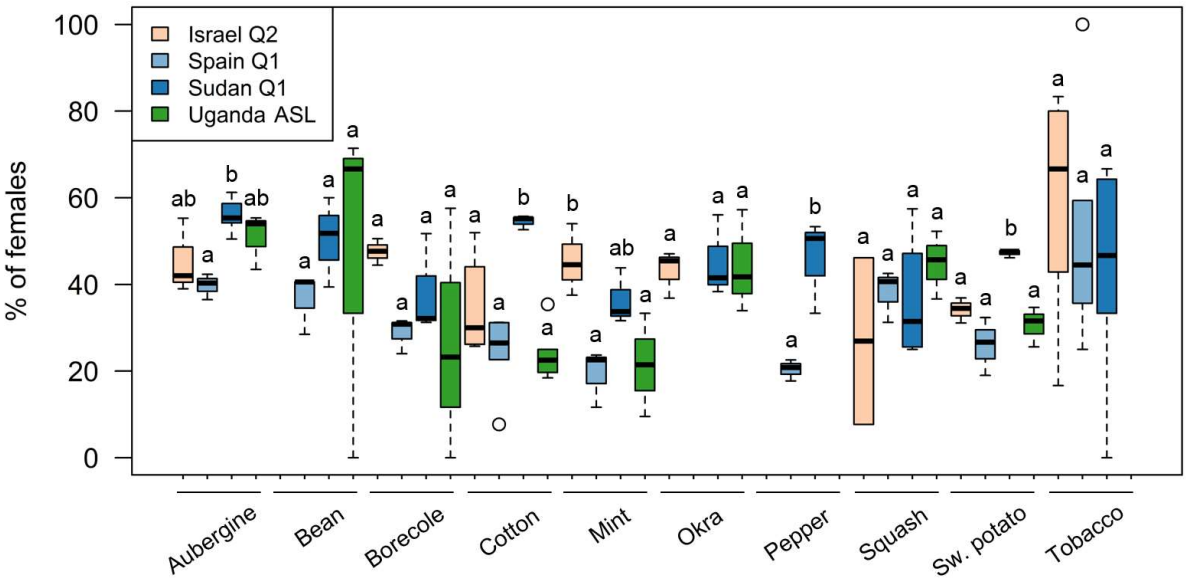


Figure 5-10: Percentages of female adults in F_1 progeny produced by four whitefly populations on 10 host plants. Cassava, chard and tomato are not shown due to the absence of F_1 generation in all four populations. Box length represents the interquartile range, whiskers express the full range, black bar corresponds to the median and circles represent outliers. Different letters above the boxes indicate significant differences between whitefly populations on each plant individually (Tukey's test, P<0.05).

Table 5-6: Mean percentages of F_1 female progeny ± standard errors for four whitefly populations on 10 host plants. Cassava, chard and tomato are not shown due to the absence of F_1 generation for all populations and dashes indicate the absence of F_1 generation for some populations on the remaining plants.

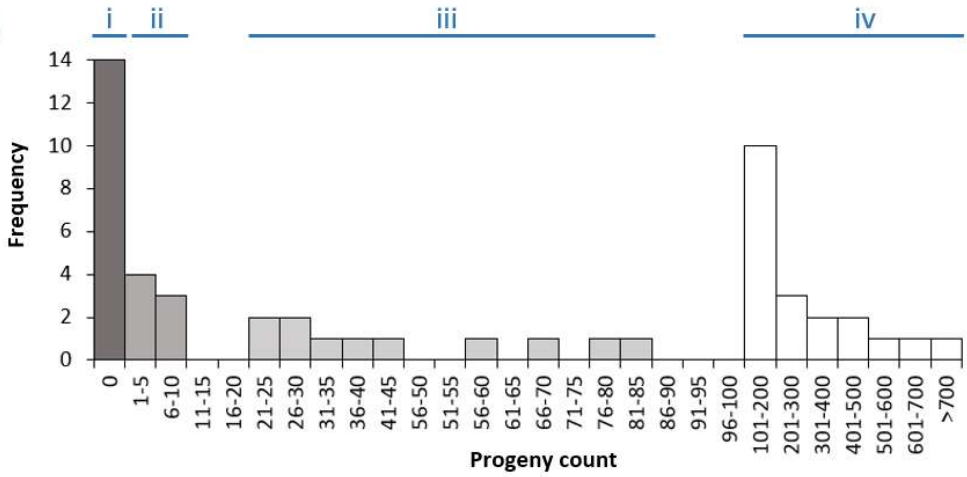
Host	Israel Q2	Spain Q1	Sudan Q1	Uganda ASL
Aubergine	45.5 ± 3.2	39.2 ± 1.0	56.3 ± 1.1	52.2 ± 1.3
Bean	-	35.6 ± 1.8	52.3 ± 3.3	58.3 ± 14.2
Borecole	47.5 ± 2.8	29.6 ± 2.1	38.2 ± 5.2	37.7 ± 5.5
Cotton	34.9 ± 1.2	27.7 ± 1.3	54.2 ± 1.3	23.9 ± 1.9
Mint	44.7 ± 3.5	18.8 ± 3.7	38.9 ± 2.5	22.6 ± 5.3
Okra	40.9 ± 6.1	-	-	45.6 ± 1.1
Pepper	-	-	50 ± 5.0	-
Squash	26.9 ± 0.1	39.7 ± 2.7	35.9 ± 3.7	43.3 ± 1.2
Sweet potato	34.2 ± 2.3	25.9 ± 2.0	47.2 ± 1.5	30.2 ± 1.5
Tobacco	55.8 ± 0.1	41.5 ± 2.6	50.0 ± 7.7	-
Mean	37.9 ± 0.9	33.16 ± 0.6	51.0 ± 0.6	42.5 ± 0.6

5.3.6 Classification of host plants

The frequency histogram plotting mean progeny counts (Figure 5-11 A) revealed three clusters: (i) 0–10, (ii) 21–85 and (iii) more than 100 F_1 adults. Based on these values, we classified the plants into three categories expressing the level of suitability as whitefly reproductive hosts: (i) unsuitable, (ii) suitable and (iii) favourable. The full list of plants categorised for each population is summarised in Figure 5-11 B.

Comparing the ranges of favourable reproductive hosts among the four populations (Figure 5-11 C) revealed that Spain Q1 had the largest host range with seven favourable hosts out of the 13 tested plants. In contrast, this number was five for Sudan Q1 and Uganda ASL, and only three for Israel Q2. Hosts favourable for all populations were cotton and sweet potato; aubergine was favourable for all but Israel Q2. Three hosts were not suitable for any population: cassava, chard and tomato.

A



B

Rank	Israel Q2	Spain Q1	Sudan Q1	Uganda "ASL"
1	cotton	aubergine	cotton	okra
2	sweet potato	bean	aubergine	squash
3	borecole	cotton	sweet potato	aubergine
4	aubergine	pepper	okra	sweet potato
5	mint	sweet potato	mint	cotton
6	okra	borecole	bean	borecole
7	squash	squash	squash	mint
8	tobacco	tobacco	pepper	bean
9	tomato	mint	borecole	tomato
10	bean	tomato	tobacco	cassava
11	cassava	cassava	cassava	chard
12	chard	chard	chard	pepper
13	pepper	okra	tomato	tobacco

C

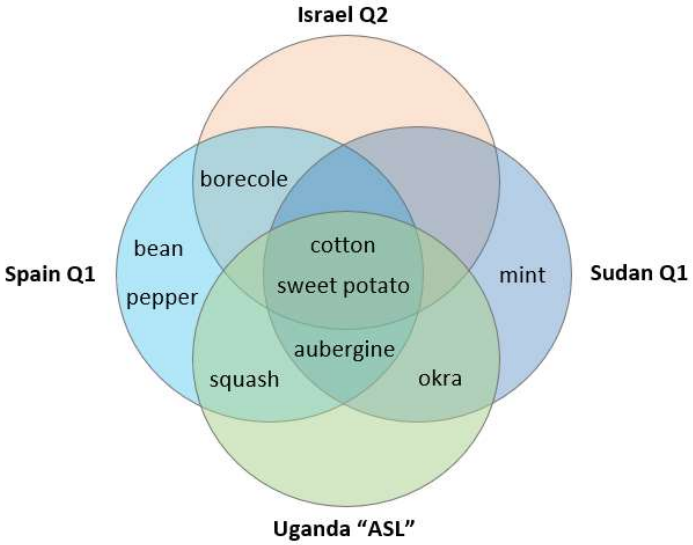


Figure 5-11: Classification of host-plants' suitability for the four MED populations based on the number of F_1 adults produced. **A:** Histogram showing the distribution of mean F_1 progeny counts of all populations on all plants with four identified clusters i–iv. **B:** Host plants ranked in descending order of suitability for each whitefly population, except for plants with zero F_1 adults that appear in alphabetical order. Colour-coding follows the histogram above. **C:** Distribution of favourable hosts (cluster iv) among the four whitefly populations. Plants not shown are cassava, chard and tomato that were unsuitable for all populations, and tobacco that fell into the cluster iii for Spain Q1.

5.3.7 Silver-leafing bioassay

The ability of the Uganda ASL population to induce squash silver-leafing was tested in a bioassay with MEAM1 as a positive control. Squash plants infested with MEAM1 whiteflies developed silver-leafing symptoms within two weeks after whitefly infestation (Figure 5-12 A). The ASL population failed to induce the symptoms (Figure 5-12 B), up to the end point of experiment (five weeks post-infestation).

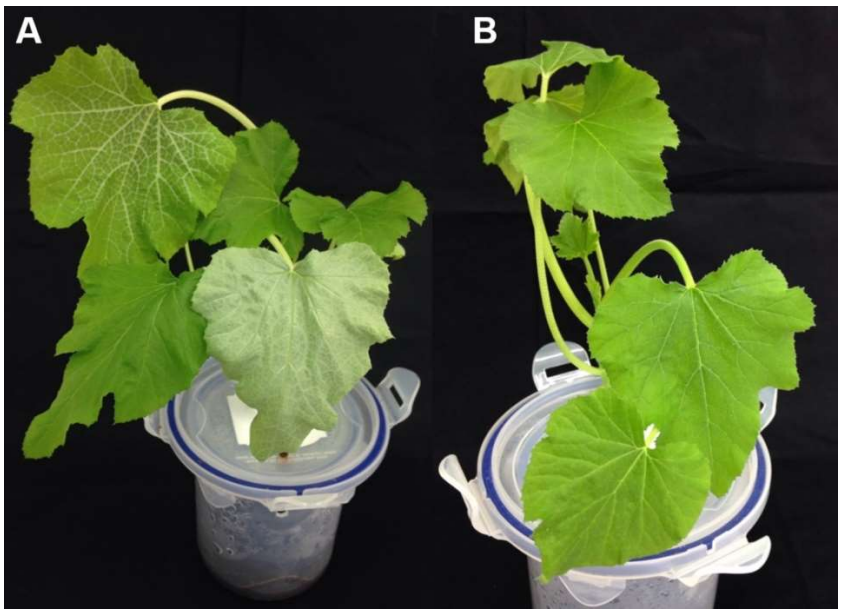


Figure 5-12: Squash plants after feeding by MEAM1 (**A**) and Uganda ASL (**B**). Photographs were taken two weeks after infestation. The plant infested with MEAM1 showed leaf silvering symptoms, while the plant fed on by ASL population did not show the symptoms.

5.3.8 Sample collection for a transcriptomic study and trial RNA extraction

Populations chosen for the study of differences in molecular mechanisms of host utilisation were Spain Q1 because of its polyphagy, and Uganda ASL because of its overall distinctness from the Q1 and Q2 groups. The selected host plants involved cotton as a favourable host for both populations; pepper as a favourable host for Spain Q1 but not for Uganda ASL; and okra as a favourable host for Uganda ASL but not for Spain Q1.

Twenty cages, labelled A to T, were set up between 2nd March and 22nd June 2017 for the purpose of collecting samples for a future transcriptomics study. A total of 92 samples contained in 102 microtubes (Appendix 21) were collected between 8th March and 28th June for Spain Q1 and Uganda ASL in five life stages on three plants. Between one and eight replicates were collected for each treatment (Table 5-7). The exception was Uganda ASL on pepper, collection of which was not possible beyond the 2nd instar because the population failed to continue its development on this unsuitable host.

Table 5-7: Number of samples collected for transcriptomic analyses for each population, plant and life stage. One sample represents about 200 eggs, 100 nymphs or 100 adults.

Population	Plant	Eggs	2nd	3rd	4th	Adults
Uganda ASL	okra	4	4	6	5	5
Uganda ASL	pepper	3	2	-	-	-
Uganda ASL	cotton	3	4	3	2	3
Spain Q1	okra	3	5	4	2	2
Spain Q1	pepper	6	8	3	2	4
Spain Q1	cotton	2	1	2	2	2

Six samples, one per each life stage and one extra for eggs, were selected for a trial RNA extraction (Table 5-2). Agarose gel electrophoresis of the extracted RNA after denaturation revealed that the extraction was successful in all samples except for 2nd instar that contained a barely detectable amount of RNA (Appendix 22). Samples extracted from the 4th instar nymphs and adults contained the highest amount of RNA. In addition, the well-defined bands in the gel indicated a low level of RNA degradation.

The integrity of RNA was further analysed by chip electrophoresis. In five out of six samples there were two well-defined peaks corresponding to the small and large subunits of ribosomal RNA (Figure 5-13). The sample originating from 2nd instar nymph, however, appeared empty. In the 4th instar nymph sample, the two peaks occurred at a lower molecular size than other samples, however, the result from agarose electrophoresis did not indicate such a difference (Appendix 22).

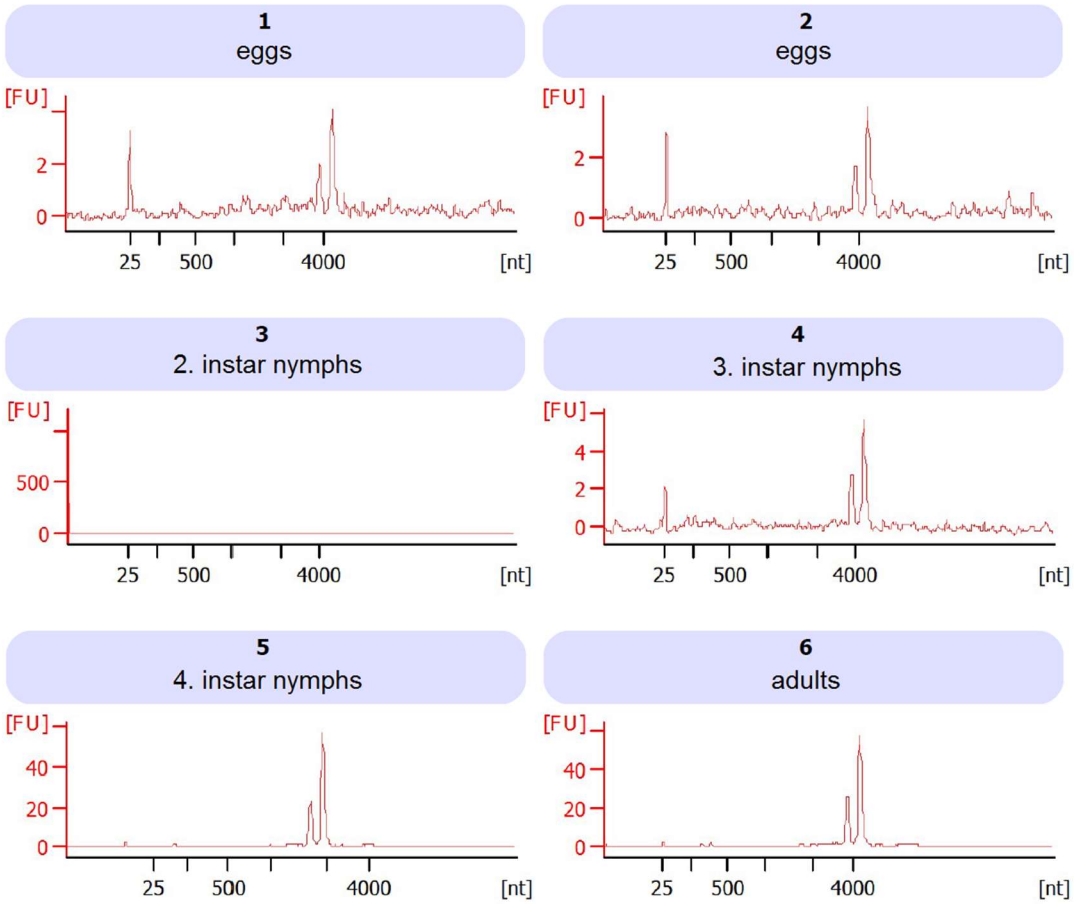


Figure 5-13: RNA electropherogram plots showing the nucleotide (nt) size distribution of molecules in the samples of extracted RNA and their relative quantity expressed in fluorescence units (FU).

The purity of our samples was analysed by the spectroscopic method of UV absorbance ratios. The 260/280 ratio of all six samples ranged from 1.68 to 2.04 (Table 5-8), which was close to the target purity value for RNA of 1.8–2.0. On the other hand, the 260/230 values reached the recommended range of 1.8–2.2 only in samples from 4th instars and adults. The remaining four samples had a low 260/230 ratios between 0.48–1.23, which indicated the presence of residual chemicals, such as carbohydrates or guanidine.

Table 5-8: Purity of the extracted RNA assessed by the ratios of sample absorbance at 260 and 280 nm, and 260 and 230 nm, measured with NanoDrop.

No.	Sample	260/280	260/230
1	eggs SpQ1 pepper	1.83	0.77
2	eggs SpQ1 cotton	1.68	0.85
3	2th instar SpQ1 pepper	1.76	0.48
4	3rd instar UgASL okra	1.81	1.23
5	4th instar UgASL okra	2.04	1.80
6	adults SpQ1 pepper	2.00	1.94

The concentration of RNA in the eluates was measured by both NanoDrop (UV spectroscopy) and Bioanalyzer (fluorometrics) and the total yield in 60 µl was calculated. Concentration measurements by Bioanalyzer were 1.6–2.7 times higher for eggs and 2nd instar nymphs than NanoDrop values, but the remaining three samples had congruent results from both methods (Table 5-9). Based on these measurements, the amount of material needed to obtain the minimum necessary amount of RNA (2 µg) was calculated (Table 5-9). For 4th instars and adults, only 16–24 specimens were needed, however, the Nanodrop values indicated that over 500 eggs or 2nd instar nymphs might be required to get 2 µg of RNA.

Table 5-9: Concentration of the extracted RNA measured with NanoDrop and Bioanalyzer and the calculated total yield in 60 µl. The last column shows how many specimens (eggs, nymphs or adults) is needed to acquire 2 µg of total RNA based on the yield calculations.

		NanoDrop		Bioanalyzer		Material used	Material needed
No.	**Sample**	**Conc.**	**Yield**	**Conc.**	**Yield**		
1	eggs	11.8 ng/µl	0.71 µg	24 ng/µl	1.44 µg	200	278–565
2	eggs	15.4 ng/µl	0.92 µg	24 ng/µl	1.44 µg	200	278–433
3	2nd instar	6.2 ng/µl	0.37 µg	17 ng/µl	1.02 µg	98	192–527
4	3rd instar	20.1 ng/µl	1.21 µg	23 ng/µl	1.38 µg	100	145–166
5	4th instar	126.5 ng/µl	7.59 µg	112 ng/µl	6.72 µg	60	16–18
6	adults	141.1 ng/µl	8.47 µg	143 ng/µl	8.58 µg	100	23–24

5.4 Discussion

Our fifth research question involved the biological differences among the Q1, Q2 and ASL groups with relation to the host-plants. We addressed this question by measuring the performance of our model populations of Q1, Q2 and ASL on a range of 13 plants and assessing the size and overlap of their host-plant ranges. The results showed significant differences in host range between ASL and the other MED groups, providing biological evidence of their distinctness. Furthermore, some significant differences also occurred between Q1 and Q2 groups and even within the Q1 group. Additionally, by measuring multiple life history traits we were able to see how these traits relate to each other and which of them could be useful for making decisions in pest management.

Our assay for measuring the oviposition rate was the least time-consuming, taking up only two days and counting relatively few (0–145) eggs. The experiment could be even more time-efficient if performed on leaf discs or using clip-cages to decrease and standardise the area available for laying eggs. However, the oviposition rate recorded after one or two days after release might not be representative for the full life span of a female, partly because of the adjustment time required after being handled by a researcher and forced into a host shift. More importantly, numbers of laid eggs do not automatically indicate the reproductive success on the plant, if the percentages of hatched eggs and nymphs reaching adulthood remain unknown.

The adult survival data served as a proxy of adult feeding success and provided an indication about the plant's potential to support the survival of adult whitefly populations in the field. The method adopted in this study might be too time-consuming for a practical application, as it required nearly daily counting for almost three weeks. The monitoring period could be shortened; however, the dynamics of population decline can differ between the periods directly after the host shift and after settling on the host. For example, the populations of Q1 and Q2 on borecole declined rapidly during the first 4–6 days, but then the population size stabilised around 30–50% for the rest of the experiment. Thus, extrapolating the adult survival time exclusively based on data from the first few days after a host shift could lead to inaccurate predictions.

We consider our results from F_1 progeny counts to be the most informative about the suitability of host plants, because the pest's ability to produce further generations is a key component of population dynamics for their infestation and spread. The assay only involved releasing parental adults with a standardized sex ratio and counting their progeny; periodic monitoring was not essential. This method could be modified to be more

time-efficient by shortening the ovipositing period and not studying the sex ratio of the progeny, as the differences in sex ratios were not as significant nor indicative of the reproductive success.

It is possible that the sex ratio of offspring was affected by secondary endosymbiotic bacteria, namely *Wolbachia*, *Cardinium* or *Rickettsia*, which can manipulate the reproductive system of their insect hosts (Perlman *et al.*, 2006; Engelstädter and Hurst, 2009). This might be the reason behind the slightly female-biased offspring of Sudan Q1 (Table 5-6), which was fixed for both *Rickettsia* and *Wolbachia*, or male-biased offspring of Spain Q1, which probably did not contain any of the three reported reproductive manipulators (Table 4-5).

Our host assessment divided the plants into three clusters based on the number of F_1 adults that successfully completed their life cycle on the plant. Data presented in Figure 5-11 B and 5-11 C suggested that Spain Q1 was the most polyphagous among the populations studied, with seven favourable and two suitable hosts out of 13 plant studied (Figure 5-11). Moreover, okra should also be considered as a suitable reproductive host, because during the sample collection for RNA extraction we managed to obtain adult whiteflies of Spain Q1 from okra. The reason for this discrepancy was that the adults took over 40 days to develop (Appendix 21), while in our host-performance assay the data collection ended at day 35.

Uganda ASL was the second least polyphagous population in this study, as it only successfully completed a life cycle on seven plant hosts. In contrast, a previous study in Uganda suggested that the ASL group was more polyphagous than Q1, because adults were collected from 37 and 10 plant species, respectively (Mugerwa, 2018). However, the ASL adults occurred five times more frequently than Q1 in the sampled data, which could have affected the perceived level of polyphagy in favour of the ASL group.

While conducting our experiments, we noticed that the squash plants fed on by Uganda ASL did not develop silver-leafing symptoms. In the additional silver-leafing bioassay with MEAM1 as a control, none of the plants infested with the ASL population developed silver-leafing symptoms five weeks after whitefly infestation, while the positive control with MEAM1 in our study and in Yokomi *et al.* (1990) showed symptoms within two weeks. A subsequent literature search revealed that there was no published evidence of the capability of the sub-Saharan MED-like populations to induce silver-leafing in squash. The original paper by Sseruwagi *et al.* (2005) clearly stated that only groups Ug6 and Ug7 (corresponding to MEAM1 and IO, respectively) induced squash silver-leafing, while Ug4

(similar to “okra biotype” from Ivory Coast) did not. The “silver-leafing” group of MED was previously linked to the former J biotype (De Barro *et al.*, 2011) reported in Nigeria, Ghana, Cameroon, Ivory Coast and Zimbabwe (De la Rúa *et al.*, 2006). However, De la Rúa *et al.* (2006) did not study the capacity to induce silver-leafing and in earlier studies it was explicit that J biotype failed to induce this phytotoxic disorder (Bedford *et al.*, 1994; Brown *et al.*, 1995a). Despite this lack of evidence, multiple subsequent studies about African silver-leafing populations of MED referred either to Sseruwagi *et al.* (2005) or De la Rúa *et al.* (2006), or to other papers that have not demonstrated the silver-leafing capacity (e.g. to Boykin *et al.*, 2007, De Barro *et al.*, 2011, or Dinsdale *et al.*, 2010). It is possible that the results of Sseruwagi *et al.* (2005) were misinterpreted due to the mixed “Ug4/Ug6” population reported to induce silver-leafing according to Table 1 in that study, and/or a confusion caused by the “Ug” naming system, and this misinterpretation was then cited in the subsequent literature.

The performance of Uganda ASL was more polarized than the other populations. It did poorly on many hosts (0–25 F_1 adults), but on the favourable ones it performed very well (298–657 F_1 adults). Israel Q2, on the other hand, was the only population that did not outperform other populations on any plant. This result could be due to a mismatch between the assayed plants and the actual range of favourable plants for Israel Q2, or an overall lower fitness of this population, as indicated by the low mean fecundity (Table 5-5). A limitation of this study is that only one population was tested from each locality on one or two varieties per plant. Ideally, several related populations would be tested on multiple plant varieties in more replicates in order to make a better assessment about the host suitability.

The Sudan Q1 population provided a model for comparison of how genetic relatedness and geographic proximity of whitefly populations relate to their host plant ranges. At the mitochondrial DNA level, Sudan Q1 was the most closely related to Spain Q1 and the least close to Uganda ASL. Geographically, however, it is the opposite. Comparing the host use of these three populations, Spain Q1 and Sudan Q1 shared similar preferences for bean, pepper and squash. However, on borecole, okra and sweet potato, the preferences of Sudan Q1 resembled more closely those of Uganda ASL. In addition, differences in biological adaptations are further evident in that Sudan Q1 outperformed all other populations on cotton and mint.

Our results not only allow a comparison between whitefly groups, but also allow a comparison between the parental fitness (oviposition rate, survival) and the fitness of their immature offspring (progeny counts). According to the optimal oviposition theory,

the oviposition preference of the female should correlate with host suitability for offspring development (Jaenike, 1978). However, this correlation in published studies ranges from very good to poor (Mayhew, 1997; Scheirs, 2002). An alternative theory is the optimal foraging model which predicts that adults prefer to feed on those hosts that give the highest adult performance, which exhibits itself in realized fecundity (Stephens and Krebs, 1986; Scheirs *et al.*, 2000). For the *B. tabaci* species complex, the current evidence does not support the optimal oviposition theory (Jiao *et al.*, 2012).

Discrepancies between the oviposition rate and progeny success were also observed in this study. The most striking was that of Uganda ASL on bean, where oviposition rate was the second highest among all measurements (Table 5-3); however, the number of adult offspring was near zero (Table 5-5). Other examples of such discrepancies included Israel Q2 on pepper and squash, Spain Q1 and Uganda ASL on tomato, and Uganda ASL and Sudan Q1 on borecole. In most of these examples the adult survival was also poor (Table 5-4), but with exceptions (e.g. Israel Q2 on squash). Conversely, in some cases the number of F_1 progeny exceeded the expected number of F_1 adults (mean number of eggs/female/48 h multiplied by fifteen and nine, giving an optimistic estimate of progeny counts after 15 females ovipositing for 18 days with no mortality). The largest difference was observed for Israel Q2 and Sudan Q1 on cotton, progeny counts of which were 1.7-fold greater than the optimistic estimate. These two cases can be partially explained by the likely underestimation of the oviposition rate measured in the first 48 h.

We also found that the adult survival could not be easily linked to offspring survival. For example, Israel Q2 adults survived very well on bean, although the offspring failed to develop. Similar, less extreme examples include Israel Q2 on borecole and Uganda ASL and Spain Q1 on sweet potato. An opposite trend, in which the large offspring success was not reflected in good parental survival, was also observed (Sudan Q1 on cotton, Sudan Q1 on mint, Uganda ASL on squash).

Taken together, our results suggest that there is no straightforward relationship between adult feeding success, female oviposition efforts and the offspring developmental success. There are likely additional factors contributing to a female's decision to oviposit, other than the quality of her feed. For example, the leaf surface morphology can play a role (Chu *et al.*, 1995; Mcauslane, 1996). Possible explanations include interference with the air movement and capturing humidity by the leaf pubescence (Butler *et al.*, 1986; Inbar and Gerling, 2008) and evolutionary changes stemming from multi-trophic interactions between whiteflies, plants and natural enemies (van Lenteren *et al.*, 1995; Guershon and Gerling, 1999; Meekes *et al.*, 2000; Queiroz and Oliveira, 2001; Head

et al., 2004). Therefore, the preference of a female whitefly to oviposit on a particular host, and potentially also on leaves in different growth stages, could reflect an evolutionary trade-off between host nutritional suitability and the risk of being attacked by natural enemies (Dicke, 2000; Murphy, 2004; Vosteen *et al.*, 2016).

No clear link was observed between the level of leaf hairiness and general oviposition rate or fecundity in our study. For example, three hosts with the highest oviposition rates (sweet potato, aubergine and bean) and four hosts on which the whiteflies reached the highest fecundity (aubergine, cotton, sweet potato and okra) carried leaves with a variety of qualities. However, we did observe a pattern in the oviposition preference for different leaf stages. The strongest oviposition preference for older leaves occurred on tomato, tobacco, aubergine and okra, all of which had relatively smooth old leaves but the young ones were densely covered with trichomes. In contrast, five out of six plants on which relatively young leaves were preferred were glabrous (sweet potato, pepper, cassava, cotton and borecole). The only exception was squash, young leaves of which were moderately hairy. We hypothesise that there were additional leaf properties playing a large role in progeny success other than the leaf surface morphology, such as the chemical defence mechanisms of plants against pests.

The capability of whiteflies to feed and reproduce on a plant requires mechanisms dealing with the plant defences. For example, borecole and other *Brassica* species produce glucosinolates which, upon feeding damage, are broken down into bioactive compounds by the enzyme myrosinase (Bones and Rossiter, 1996). The capability of Israel Q2 to feed on borecole was expected, because its ability to circumvent this defence mechanism by glucosinolate desulfation has been reported (Malka *et al.*, 2016). As all four populations in this study were able to feed and reproduce on borecole (26–151 F_1 adults), it is possible that all of them can employ this deactivating mechanism, albeit with varying level of efficiency.

The unsuitability of tomato for all populations was surprising, because tomato is considered a good host for MED (Bonato *et al.*, 2007; Jiao *et al.*, 2012) and the variety ‘Moneymaker’ has previously been shown susceptible to MED in Spain (Nombela *et al.*, 2003; Rodríguez-López *et al.*, 2011). The oviposition rate on tomato was relatively high for all four populations (Figure 5-5), so the adverse effects of the plant defence probably took place during the development of eggs or nymphs. Tomato leaves are covered with eight possible types of glandular and non-glandular trichomes (Channarayappa *et al.*, 1992). The glandular trichomes produce zingiberene and acylsugars that can have a repelling, fumigant or toxic effect to whiteflies, or the released exudates can trap whitefly

adults (Kisha, 1981; Muigai *et al.*, 2002). The glandular trichomes (type IV) were negatively correlated with whitefly oviposition and density and positively correlated with trapping and mortality (Muigai *et al.*, 2003; Oriani and Vendramim, 2010). The tomato trichome density varied between studies, which was attributed to the variation in experimental factors, such as the photoperiod, plant age, fertilisers and the growing environment (greenhouses or the field) (Oriani and Vendramim, 2010 and references within). In our study, the experimental conditions with tomato plants confined within a small Lock&Lock cage could have potentially influenced the leaf surface morphology.

Tobacco leaves contain nicotine, which is a highly toxic alkaloid for herbivores (Appel and Martin, 1992; Steppuhn *et al.*, 2004; du Rand *et al.*, 2015). Tobacco was not a favourable host for any of the populations in this study, however, Spain Q1 was the only population for which tobacco was at least suitable (58 F_1 adults). This was expected, as the mtCOI of the Q1 group corresponds to the original *B. tabaci* syntype collected from tobacco in Greece (Gennadius, 1889; Tay *et al.*, 2012) and because the invasive Q1 populations were reported to develop resistance to neonicotinoids (Fernández *et al.*, 2009; Dennehy *et al.*, 2010; Luo *et al.*, 2010). Neonicotinoids are pesticides with a similar mode of action to nicotine (Tomizawa and Casida, 2003), and so the detoxification mechanisms of these two compounds are likely similar (du Rand *et al.*, 2015; Magesh *et al.*, 2017). In contrast, Sudan Q1 did not reproduce well on tobacco (7 F_1 adults), despite the close genetic link to the invasive Q1 populations. This was not surprising, because the Sudan Q1 population, also called SUD-S or BTS, has been used frequently as a susceptible control in toxicological studies (Byrne *et al.*, 2000; Nauen *et al.*, 2002; Rauch and Nauen, 2003; Nauen and Denholm, 2005; Ma *et al.*, 2007). This low ability to detoxify nicotine could partially explain why the sub-Saharan populations are not invasive.

The detoxification capacity of whiteflies could also be affected by the composition of endosymbiotic bacteria. For example, *Rickettsia*-positive MEAM1 whiteflies were more susceptible to pyriproxyfen, acetamiprid, thiamethoxam and spiromesifen than their *Rickettsia*-negative counterparts (Kontsedalov *et al.*, 2008). Similarly, the presence of *Rickettsia* in a double infection with *Arsenophonus* increased the mortality of MED whiteflies in Israel after treatment with pyriproxyfen, thiamethoxam, spiromesifen and imidacloprid, when compared with MED whiteflies with single *Arsenophonus* infection (Ghanim and Kontsedalov, 2009). Interestingly, the most polyphagous Spain Q1 was the only population in which *Rickettsia* was not detected by the PCR test in our study (Table 4-5), and the least polyphagous Israel Q2 seemed to have a very high relative abundance of *Rickettsia* (Figure 4-12). Further investigation, involving quantitative PCR and

isofemale lines of whiteflies with different infection status, would be needed to establish the role of endosymbionts in the detoxification mechanisms.

This study has implications for the pest management of MED. The polyphagy and ability to detoxify nicotine could serve as a pre-requisite to the invasiveness of Spain Q1 and related populations. More specifically, polyphagy makes pests more flexible and challenging to control, while the nicotine detoxification capacity makes them predisposed to develop resistance to neonicotinoid pesticides. Fighting the invasive MED, therefore, requires an alternative and combined approach.

Our findings could contribute to making decisions about plant choices in cultural practices, such as intercropping, trap crops or barrier crops. For example, intercropping with plants supporting long adult survival and development of large numbers of offspring, such as sweet potato, cotton and aubergine, could have detrimental effects to adjacent susceptible crops. Such plants would support the multiplication and facilitate the spread of MED or ASL populations in the field. More suitable plants for intercropping would be those attractive for feeding and oviposition, but toxic for adults or nymphs, such as bean for Uganda ASL or pepper and squash for Israel Q2. However, an experiment with choice conditions would be needed to ascertain whether there is a preference for the trap crop over the susceptible crop. Moreover, understanding the relationships between the intercropped plants and whitefly natural enemies, and subsequent choice of plants that support the populations of natural enemies, could lead to an enhanced overall efficiency of pest control (Landis *et al.*, 2000).

Plants unattractive or toxic to whiteflies could be used as a physical barrier around susceptible crops. For example, it has been suggested to use maize as a barrier crop to protect cotton fields from MED whiteflies in China, because *B. tabaci* do not attack maize (Zhang *et al.*, 2014). Our results suggest that chard could serve as a barrier crop against MED groups, but the low height of these plants needs to be taken into consideration. Cassava was similarly unsuitable for MED as chard, however, its use as a barrier crop would be impossible in sub-Saharan Africa, because multiple other African species of the *B. tabaci* complex colonise and seriously damage cassava (Macfadyen *et al.*, 2018).

Another component of future pest management could be the development of whitefly-resistant crop varieties or trap crops containing a whitefly-specific lethal factor. For example, a whitefly-resistant transgenic tobacco was developed by inserting a dsRNA precursor for RNAi specifically designed to silence the vital insect gene for vacuolar ATPase enzyme, subunit A (Thakur *et al.*, 2014). Our planned transcriptomic study could

help in understanding the molecular mechanisms conferring the whiteflies' ability to utilise host plants and so facilitate the identification of targets for RNAi. Our preliminary work in this chapter resulted in nearly complete set of samples. Based on the trial RNA extraction, there is enough material to obtain 2 µg RNA from 24 out of 30 treatments. The only missing parts are 200 more eggs of Spain Q1 on cotton and 100–400 more 2nd instar nymphs of all treatments except Spain Q1 on pepper. Upon extraction and purification, the RNA could be sequenced and data analysed in the search for differentially expressed genes while developing on suitable and unsuitable hosts.

CONCLUSIONS

This study aimed to investigate biological species diversity within a subset of the *B. tabaci* species complex by taking an integrative approach for species delimitation. We combined data from partial and full mitogenome sequences, reproductive incompatibilities, bacterial endosymbionts and host-plant adaptations to gain a more comprehensive view on the species status of the genetic groups Q1, Q2 and ASL within the putative MED species. We contrasted the results of our combined approach with the current view of MED as a single putative cryptic species delimited on the basis of the barcoding mitochondrial marker and the 3.5% nucleotide distance threshold.

Question 1: How closely are the MED groups related at the mtDNA level?

Phylogenetic analysis based on 15 mitochondrial genes from 12 different populations of seven putative species (MED, MEAM1, IO, NW1, Asia II-7, Australia and Asia I) in the *B. tabaci* complex provided an insight into evolutionary relationships within the Africa/Middle East/Asia Minor clade and among the groups of MED putative species. **All Q1 and Q2 populations displayed close genetic relationships and nucleotide distances between 0% and 1.37% across all four mitogenome regions. The ASL population was more distinct from Q1 and Q2 (1.98–4%) and was placed between the MED and MEAM1 species.** The phylogeny and nucleotide distances were calculated from HTS-derived mitogenomes, which provided us with confidence that only true mitochondrial sequences and no NUMTs were included in our analyses. This approach contrasted with previous studies inferring phylogenetic relationships and sequence divergence among MED populations from 3′ partial mtCOI sequences, because these contained poor-quality sequences and/or NUMTs (see Question 2 below).

Question 2: How reliable are the commonly used barcoding methods for identifying biological species?

Several issues were identified with the barcoding method used for *B. tabaci.* While the standard barcoding region of the mtCOI gene is near the 5’ end, the whitefly community has adopted the use of a segment near the 3’ end of the gene. **Our results demonstrated conflicting outcomes from an arbitrary choice of the barcoding region, as the ASL would be placed inside or outside MED based on the nucleotide distance under or over 3.5% in the 3’ or 5’ mtCOI region, respectively.** This incongruity emphasised the perils of over-reliance on a nucleotide distance threshold from a single partial mitochondrial sequence in species delimitation. Despite being only 109 bp apart and their divergence being highly correlated with the overall divergence

in the mitogenome, the 3' and 5' partial mtCOI regions provided discordant results when studied separately. **Moreover, our quality assessment of 289 unique haplotypes of 3' partial mtCOI sequences of MED groups from GenBank revealed that over 53% of them contained errors, which is potentially due to their pseudogene origin.** Failure to recognise and exclude such sequences from downstream analyses can lead to inflated values of intra- and inter-species nucleotide divergence, confound phylogenetic analyses and overestimate the species richness (van der Kuyl *et al.*, 1995; Thalmann *et al.*, 2004; Song *et al.*, 2008). Because NUMTs have already been reported in the order Hemiptera (Sunnucks and Hales, 1996), including in the MEAM1 putative species of the *B. tabaci* complex (Tay *et al.*, 2017a), there is no reason to assume that NUMTs are not present in the rest of the *B. tabaci* complex and care should be taken in future analyses to detect and discard such sequences.

Question 3: Are the MED groups reproductively compatible?

The reciprocal crossing experiments among MED populations were carried out with a rigorous approach. This involved using only virgin females, testing F_1 hybrid fertility and developing a molecular test for confirming the parentage of F_1 females emerged from the experiments. **This is the first time such a specific nuclear marker has been used for *B. tabaci* research.**

The documented genetic distinctness of ASL from the MED groups Q1 and Q2 (Question 1) was further supported by the evidence of reproductive isolation between them. Our results were in accordance with the previous field population genetics study on ASL and Q1 (Mouton *et al.*, 2015). **The Q1 and Q2 populations from the Mediterranean Basin were compatible and produced fertile hybrids**, which also confirmed previous field population genetics data (Gauthier *et al.*, 2014; Terraz *et al.*, 2014; Thierry *et al.*, 2015; Hadjistylli *et al.*, 2016). **Of a particular note was that the sub-Saharan Q1 population was incompatible with both Mediterranean Q1 and Q2 populations, which has not been reported before.** This provides further evidence that the proposed 3.5% boundary in partial mtCOI nucleotide distance is inconsistent with the presence of reproductive barriers between or within species. The presence of these reproductive barriers indicates that MED (*sensu* Dinsdale *et al.*, 2010) comprises more than one biological species.

Question 4: Do endosymbionts play a role in the reproductive relationships among MED groups?

Bacterial endosymbionts may influence reproductive relationships within MED. Each of the four populations in this study harboured a different combination of endosymbiotic bacteria, so the relationship between any two populations could have been affected by endosymbionts. Two main reasons for considering endosymbiont effects were: (i) the reproductive incompatibility between the closely related Q1 populations from Spain and Sudan and (ii) the asymmetry observed in F_2 progeny from back-crossing of Spain Q1 x Israel Q2 hybrid females with parental type males. Firstly, we hypothesise that **the incompatibility between the two Q1 populations might have been caused by *Wolbachia* and/or *Cardinium***, because they are both recognised reproductive manipulators (Werren *et al.*, 2008; Mann *et al.*, 2017) and both are present in Sudan Q1 but absent from Spain Q1. Secondly, **the asymmetry in F_2 progeny from Spain Q1 and Israel Q2 cross could have been influenced by *Rickettsia* and/or *Hamiltonella***, as *Rickettsia* has been linked to a higher percentage of females in the progeny and *Hamiltonella* with higher fitness in whiteflies (Himler *et al.*, 2011; Su *et al.*, 2013a; Asiimwe *et al.*, 2014). The involvement of these bacteria could explain the more numerous F_2 progeny with Q1 maternal background and the stronger female bias in the F_2 progeny with Q2 maternal background. Further experiments are needed to test these hypotheses rigorously.

Question 5: Do the MED groups share the same host-plant range and do they perform equally well on the host-plants?

We observed statistically significant differences in host-plant performance among all four populations in adult survival times and progeny counts. Some populations also differed in oviposition rate and progeny sex ratio. None of the populations shared the same host range, as classified on the basis of their achieved fecundity. The Spain Q1 population was the most polyphagous population and the Israel Q2 population the least, although the latter colony achieved a relatively low fecundity on all plant-hosts. **The Uganda ASL population displayed the most marked differences in host-plant performance compared with the Q1 and Q2 populations.** The ASL population exhibited a uniquely high fecundity on okra and squash and a very long adult survival on sweet potato, but the adult death rate was very rapid on unsuitable plants, such as tobacco, bean, tomato and pepper.

This work contributed to our knowledge about the range of plants attacked specifically by populations of different groups of the putative MED species. Such knowledge can help devising efficient strategies in the integrated pest management and more specifically in cultural practices. Moreover, we advocate against the use of neonicotinoid pesticides against the invasive MED populations (Q1 and Q2), because they have the potential to detoxify nicotine.

Our findings clearly demonstrate that the general host range data about *B. tabaci* complex cannot be universally applied, because populations belonging to the same putative species (MED), and even to the same putative group (Q1) can significantly differ in their host preferences.

Overall aim: An integrative approach to evaluate the species status of groups within the putative MED species.

An accurate delimitation of species is important for biodiversity conservation, but in the case of pest organisms, it is crucial to "know the enemy" in order to protect farmers' livelihoods effectively and prevent crops from damage. Ideally, the study of multiple factors, such as reproductive compatibility, morphology, ethology, ecology and molecular markers, would be conducted to delimit the species (Smith *et al.*, 2008; Leaché *et al.*, 2009). In this study, therefore, we have combined molecular markers, reproductive incompatibility and ecology to assess the species status of groups of the putative MED species.

From the biological species concept definition, the Q1 and Q2 groups belong to the same species because they can interbreed successfully. Furthermore, as the partial mtCOI sequence of the *B. tabaci* (Gennadius) syntype from Greece corresponded to MED-Q1 (Tay *et al.*, 2012), we conclude that **the Q1 and Q2 both belong to the species *B. tabaci sensu stricto.*** At the same time, however, we recognise their genetic differentiation in allopatry. It is possible that Q1 and Q2 represent diverging populations of incipient species that have retained the ability to interbreed when occurring in sympatry. This, however, relates only to the Mediterranean population of Q1. The reproductive incompatibility of Sudan Q1 and Spain Q1 indicated that Sudan Q1 might belong to a different species, but more evidence is needed to draw firm conclusions.

By combining multiple lines of evidence about the evolutionary relationships, reproductive incompatibility and host-plant range from our study and the published literature, we conclude that **the ASL group belongs to a separate, non-MED species within the Africa/Middle East/Asia Minor clade** of the *B. tabaci* species complex. If all populations

listed in Table 3-8 belong to the same species, the current informal synonyms of “ASL” include: “okra biotype” (Burban *et al.*, 1992; Omondi *et al.*, 2005), “Ug4” (Sseruwagi *et al.*, 2005), “J biotype” (De la Rúa *et al.*, 2006), “Sub-Saharan Africa silverleafing (Q-related)” (Boykin *et al.*, 2007), “SubSaharan Africa Silverleaf” (Dinsdale *et al.*, 2010), “Silverleafing East and West African Mediterranean populations” (De Barro, 2012) and “Q4” (Chu *et al.*, 2012a). In addition, we argue that **the “ASL” group does not cause silver-leafing in squash**, because it failed to induce these symptoms in our bioassays and no published evidence of this phenomenon was found. This species should, therefore, be renamed.

It is possible that the “ASL” species had already been described and later synonymised with *B*. *tabaci* (Russell, 1957). The synonymised species collected in sub-Saharan Africa include *B*. *gossypiperda* var. *mosaicivectura* Ghesquière (Congo, 1934), *B*. *vayssierei* Frappa (Madagascar, 1939), *B*. *goldingi* Corbett (Nigeria, 1935), *B*. *nigeriensis* Corbett (Nigeria, 1935) and *B*. *rhodesiaensis* Corbett (Rhodesia = Zimbabwe, 1936) (Mound and Halsey, 1978). DNA sequencing from museum syntypes of these species would be needed to establish whether any of the above names apply to the “ASL” species. If not, a formal description of this new species will be required to accompany a new species name.

The previous inclusion of the “ASL” population in the MED species demonstrated a fundamental problem with species delimitation based exclusively on an arbitrarily chosen barcoding sequence in mitochondrial DNA. Furthermore, we showed a case of serial mis-citations that led to the perpetuation of the erroneous assignment of the diagnostic trait of squash silver-leafing to the “ASL” species. Overall, we advocate the importance of an integrative approach to cryptic species identification and delimitation, most importantly including the biological species concept and using multiple genes or genomic data to infer the phylogeny of the *B*. *tabaci* complex. With the increasing ease with which genome-wide SNPs can be obtained, combining nucleotide polymorphisms from both mitochondrial DNA and nuclear DNA genomes (e.g. Elfekih *et al.*, 2018) will offer greater power for inference of *B*. *tabaci* cryptic species phylogenetic relationships. Understanding the species diversity in pest-species complexes is the first step towards designing novel and targeted management strategies, as well as bringing consistency and clarity to communication within the scientific community and the wider public.

The combined contents of Chapter 3 and Chapter 4 have been published in Scientific Reports in July 2018 (Publications, page 231). A manuscript based on Chapter 5 has been accepted for publication in April 2019 in the Journal of Pest Science (Publications, page 244).

FUTURE WORK

This study worked with four populations belonging to three genetic groups previously identified within the MED putative species. However, there are more MED groups that have not been included in this research and these should be studied to gain a more complete picture of the evolutionary relationships within MED. Firstly, there is the Q3 (Gueguen *et al.*, 2010) or Q5 (Chu *et al.*, 2012a) group identified in Burkina Faso, and secondly, the Q3 (Chu *et al.*, 2012a) found in Croatia. Our comparison of their 3' partial mtCOI sequences revealed close genetic relationships among them, as well as with the Israel Q2 population from this study. However, as live colonies of these groups were unavailable, we could not perform experiments with them, nor generate HTS data and include these groups in our mitogenome phylogeny. Most importantly, reciprocal crossing experiments between Q3 groups and Q1, Q2 and ASL should be performed to ascertain the species status of these population by the biological species concept. Their close genetic relationships with Israel Q2 and each other in the mtCOI gene do not mean that they are necessarily reproductively compatible, as we have shown in our study.

A reciprocal crossing experiment between an "ASL" population from West Africa and the East African "ASL" population from this study would help elucidate the systematics and the geographic spread of "ASL". In addition, because only an East African "ASL" population was studied in this work, it would be useful to perform squash silver-leafing bioassays with Western African populations of "ASL" to confirm whether or not they induce silver-leafing; a phenomenon which is often cited but undocumented for "ASL".

We suggest that the "ASL" putative species should be officially described and assigned a Latin binomial. This effort is currently ongoing in collaboration with experts on Sternorrhyncha systematics from Natural History Museum in London, who are also developing a novel technique for comparing the morphometrics of the 4th instar whitefly nymphs. Such data will aid the formal description of the "ASL" species, together with the molecular, biological and ecological evidence from this study.

The species status of our Sudan Q1 population remains unclear and further research needs to be done. Because our colony was originally established in the late 70's, it is possible that through a series of bottlenecks and inbreeding its biology became different to its relatives in the field. A new colony should, therefore, be established from current field populations and reciprocal crossing experiments should be repeated with Mediterranean Q1 and Q2 populations. Ideally, isofemale lines with different composition of endosymbionts, particularly *Cardinium* and *Wolbachia*, would also be developed to

study their role in reproductive compatibility among the populations. Furthermore, it would be interesting to include data from nuclear DNA in phylogenetic analyses of the evolutionary relationships within MED and the Africa/Middle East/Asia Minor clade. If Sudan Q1 and Spain Q1 belonged to different species, it would provide a useful insight into the differences between nuclear and mitochondrial evolution and the nature of reproductive barriers preventing the gene flow between sister species.

Our host plant study could be further expanded by conducting experiments with host choice, in which the whitefly preference for different plants would be studied. However, it would be laborious and impractical to conduct all 78 pairwise combinations, assuming the choice between two hosts and testing all 13 plants, which would come to a total of 312 experiments if all four populations were to be studied. Instead, host preference assays would arise from specific needs in the field, for example in making decisions about crop choice.

A continuation of the transcriptomic analysis of Spain Q1 and Uganda ASL developing on three plants species would enable identification of genes that contribute significantly to the whitefly's ability to detoxify plant-defence chemicals. This line of research could lead to the development of crops specifically resistant to whiteflies, minimising the need to spray with insecticides and thus reducing the negative impact on the environment.

REFERENCES

Abdel-Raheem, M.A., Al-Keridis, L.A., 2017. Virulence of three entomopathogenic fungi against whitefly, *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) in tomato crop. J. Entomol. 14, 155–159.

Abd-Rabou, S., Simmons, A.M., 2010. Survey of reproductive host plants of *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Egypt, including new host records. Entomol. News 121, 456–465.

Agapow, P.M., Bininda-Emonds, O.R., Crandall, K.A., Gittleman, J.L., Mace, G.M., Marshall, J.C., Purvis, A., 2004. The impact of species concept on biodiversity studies. Q. Rev. Biol. 79, 161–179.

Ahmed, M.Z., De Barro, P.J., Ren, S.X., Greeff, J.M., Qiu, B.L., 2013. Evidence for horizontal transmission of secondary endosymbionts in the *Bemisia tabaci* cryptic species complex. PLoS One 8, e53084.

Ahmed, M.Z., Li, S.J., Xue, X., Yin, X.J., Ren, S.X., Jiggins, F.M., Greeff, J.M., Qiu, B.L., 2015. The intracellular bacterium *Wolbachia* uses parasitoid wasps as phoretic vectors for efficient horizontal transmission. PLoS Pathog. 10, e1004672.

Ahmed, M.Z., Naveed, M., Noor ul Ane, M., Ren, S.X., De Barro, P.J., Qiu, B.L., 2014. Host suitability comparison between the MEAM1 and AsiaII 1 cryptic species of *Bemisia tabaci* in cotton-growing zones of Pakistan. Pest Manag. Sci. 70, 1531–1537.

Ahmed, M.Z., Ren, S.X., Mandour, N.S., Greeff, J.M., Qiu, B.L., 2010a. Prevalence of *Wolbachia* supergroups A and B in *Bemisia tabaci* (Hemiptera: Aleyrodidae) and some of its natural enemies. J. Econ. Entomol. 103, 1848–1859.

Ahmed, M.Z., Ren, S.X., Mandour, N.S., Maruthi, M.N., Naveed, M., Qiu, B.L., 2010b. Phylogenetic analysis of *Bemisia tabaci* (Hemiptera: Aleyrodidae) populations from cotton plants in Pakistan, China, and Egypt. J. Pest Sci. 83, 135–141.

Ahmed, M.Z., Shatters, R.G., Ren, S.X., Jin, G.H., Mandour, N.S., Qiu, B.L., 2009. Genetic distinctions among the Mediterranean and Chinese populations of *Bemisia tabaci* Q biotype and their endosymbiont *Wolbachia* populations. J. Appl. Entomol. 133, 733–741.

Alemandri, V., De Barro, P.J., Bejerman, N., Argüello Caro, E.B., Dumón, A.D., Mattio, M.F., Rodriguez, S.M., Truoli, G., 2012. Species within the *Bemisia tabaci* (Hemiptera: Aleyrodidae) complex in soybean and bean crops in Argentina. J. Econ. Entomol. 105, 48–53.

Ali, S., Zhang, C., Wang, Z.Q., Wang, X.M., Wu, J.H., Cuthbertson, A.G.S., Shao, Z.F., Qiu, B.L., 2017. Toxicological and biochemical basis of synergism between the entomopathogenic fungus *Lecanicillium muscarium* and the insecticide matrine against *Bemisia tabaci* (Gennadius). Sci. Rep. 7, 46558.

Alicai, T., Omongo, C.A., Maruthi, M.N., Hillocks, R.J., Baguma, Y., Kawuki, R., Bua, A., Otim-Nape, G.W., Colvin, J., 2007. Re-emergence of cassava brown streak disease in Uganda. Plant Dis. 91, 24–29.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Al-Zyoud, F., Sengonca, C., Al-Abbadi, S., 2007. Evaluation of *Serangium parcesetosum* (Col.: Coccinellidae) for biological control of *Bemisia tabaci* under greenhouse conditions. J. Pest Sci. 80, 85–92.

Amichot, M., Tarès, S., Brun-Barale, A., Arthaud, L., Bride, J.M., Bergé, J.P., 2004. Point mutations associated with insecticide resistance in the *Drosophila* cytochrome P450 *Cyp6a2* enable DDT metabolism. Eur. J. Biochem. 271, 1250–1257.

Anderson, P.K., Martin, J.H., Hernandez, P., Lagnaoui, A., 2001. *Bemisia afer sens. lat.* (Homoptera: Aleyrodidae) outbreak in the Americas. Fla. Entomol. 84, 316.

Andrews, S., 2014. FastQC A quality control tool for high throughput sequence data. Babraham Institute, Babraham, UK. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Antignus, Y., Mor, N., Joseph, R.B., Lapidot, M., Cohen, S., 1996. Ultraviolet-absorbing plastic sheets protect crops from insect pests and from virus diseases vectored by insects. Environ. Entomol. 25, 919–924.

Appel, H.M., Martin, M.M., 1992. Significance of metabolic load in the evolution of host specificity of *Manduca sexta*. Ecology 73, 216–228.

Arias, B.V., Bellotti, A.C., Vargas, H.L.B., 2004. Nataima-31, a cassava (*Manihot esculenta*) variety resistant to the whitefly, *Aleurotrachelus socialis*. Presented at the Sixth International Scientific Meeting of the Cassava Biotechnology Network, CIAT, Cali, Colombia, p. 66.

Asiimwe, P., Ecaat, J.S., Guershon, M., Kyamanywa, S., Gerling, D., Legg, J.P., 2007. Evaluation of *Serangium* n. sp. (Col., Coccinellidae), a predator of *Bemisia tabaci* (Hom., Aleyrodidae) on cassava. J. Appl. Entomol. 131, 76–80.

Asiimwe, P., Kelly, S.E., Hunter, M.S., 2014. Symbiont infection affects whitefly dynamics in the field. Basic Appl. Ecol. 15, 507–515.

Augustinos, A.A., Santos-Garcia, D., Dionyssopoulou, E., Moreira, M., Papapanagiotou, A., Scarvelakis, M., Doudoumis, V., Ramos, S., Aguiar, A.F., Borges, P.A.V., Khadem, M., Latorre, A., Tsiamis, G., Bourtzis, K., 2011. Detection and characterization of *Wolbachia* infections in natural populations of aphids: Is the hidden diversity fully unraveled? PLoS One 6, e28695.

Avise, J.C., 2004. Molecular markers, natural history and evolution, 2nd edition. Sinauer Associates, Sunderland, Massachusetts.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C., 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18, 489–522.

Balakirev, E.S., Saveliev, P.A., Ayala, F.J., 2017. Complete mitochondrial genomes of the Cherskii's sculpin *Cottus czerskii* and Siberian taimen *Hucho taimen* reveal GenBank entry errors: Incorrect species identification and recombinant mitochondrial genome. Evol. Bioinforma. Online 13.

Baldauf, S.L., 2003. Phylogeny for the faint of heart: A tutorial. Trends Genet. TIG 19, 345–351.

Baldin, E.L.L., Vendramim, J.D., Lourenção, A.L., 2005. Resistance of tomato genotypes to the whitefly *Bemisia tabaci* (Gennadius) biotype B (Hemiptera: Aleyrodidae). Neotrop. Entomol. 34, 435–441.

Baldo, L., Werren, J.H., 2007. Revisiting *Wolbachia* supergroup typing based on WSP: Spurious lineages and discordance with MLST. Curr. Microbiol. 55, 81–87.

Barbosa, L.F., Yuki, V.A., Marubayashi, J.M., De Marchi, B.R., Perini, F.L., Pavan, M.A., de Barros, D.R., Ghanim, M., Moriones, E., Navas-Castillo, J., Krause-Sakate, R., 2015. First report of *Bemisia tabaci* Mediterranean (Q biotype) species in Brazil. Pest Manag. Sci. 71, 501–504.

Barinaga, M., 1993. Is devastating whitefly invader really a new species? Science 259, 30–30.

Bartlett, P.W., 1992. Experience of polyphagous alien pests of protected crops in Great Britain. EPPO Bull. 22, 337–346.

Bas, N., Mollema, C., Lindhout, P., 1992. Resistance in *Lycopersicon hirsutum* f. *glabratum* to the greenhouse whitefly (*Trialeurodes vaporariorum*) increases with plant age. Euphytica 64, 189–195.

Basu, A.N., 1995. *Bemisia tabaci* (Gennadius): Crop pest and principal whitefly vector of plant viruses. Westview Press, Boulder, Colorado.

Baumann, P., 2005. Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. Annu. Rev. Microbiol. 59, 155–189.

Baumann, P., Moran, N.A., Baumann, L., 2006. Bacteriocyte-associated endosymbionts of insects, in: Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H., Stackebrandt, E. (Eds.), The prokaryotes. Springer, New York, pp. 403–438.

Becerra, J.X., Noge, K., Venable, D.L., 2009. Macroevolutionary chemical escalation in an ancient plant–herbivore arms race. Proc. Natl. Acad. Sci. 106, 18062–18066.

Bedford, I.D., Briddon, R.W., Brown, J.K., Rosell, R.C., Markham, P.G., 1994. Geminivirus transmission and biological characterisation of *Bemisia tabaci* (Gennadius) biotypes from different geographic regions. Ann. Appl. Biol. 125, 311–325.

Bedford, I.D., Briddon, R.W., Markham, P.G., Brown, J.K., Rosell, R.C., 1992. *Bemisia tabaci* - biotype characterisation and the threat of this whitefly species to agriculture, in: Proceedings of Brighton Crop Protection Conference Pests and Diseases. pp. 1235–1240.

Bel-Kadhi, M.S., Onillon, J.C., Cenis, J.L., 2008. Molecular characterisation of *Bemisia tabaci* biotypes in southern Tunisia. Tunis. J. Plant Prot. 3, 79–85.

Bellotti, A.C., Arias, B.V., 2001. Host plant resistance to whiteflies with emphasis on cassava as a case study. Crop Prot., Challenges and opportunities for pest management of *Bemisia tabaci* in the new century 20, 813–823.

Bellows, T.S., Perring, T.M., Gill, R.J., Headrick, D.H., 1994. Description of a species of *Bemisia* (Homoptera: Aleyrodidae). Ann. Entomol. Soc. Am. 87, 195–206.

Benirschke, K., Kumamoto, A.T., 1991. Mammalian cytogenetics and conservation of species. J. Hered. 82, 187–191.

Bensasson, D., Zhang, D.X., Hartl, D.L., Hewitt, G.M., 2001. Mitochondrial pseudogenes: Evolution's misplaced witnesses. Trends Ecol. Evol. 16, 314–321.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: Improved *de novo* metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69, 313–319.

Berry, S.D., Fondong, V.N., Rey, C., Rogan, D.T., Fauquet, C.M., Brown, J.K., 2004. Molecular evidence for five distinct *Bemisia tabaci* (Homoptera: Aleyrodidae) geographic haplotypes associated with cassava plants in sub-Saharan Africa. Ann. Entomol. Soc. Am. 97, 852–859.

Bethke, J.A., Byrne, F.J., Hodges, G.S., McKenzie, C.L., Shatters, R.G., 2009. First record of the Q biotype of the sweetpotato whitefly, *Bemisia tabaci*, in Guatemala. Phytoparasitica 37, 61–64.

Bethke, J.A., Paine, T.D., Nuessly, G.S., 1991. Comparative biology, morphometrics, and development of two populations of *Bemisia tabaci* (Homoptera: Aleyrodidae) on cotton and poinsettia. Ann. Entomol. Soc. Am. 84, 407–411.

Bing, X.L., Xia, W.Q., Gui, J.D., Yan, G.H., Wang, X.W., Liu, S.S., 2014. Diversity and evolution of the *Wolbachia* endosymbionts of *Bemisia* (Hemiptera: Aleyrodidae) whiteflies. Ecol. Evol. 4, 2714–2737.

Bing, X.L., Yang, J., Zchori-Fein, E., Wang, X.W., Liu, S.S., 2013. Characterization of a newly discovered symbiont of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae). Appl. Environ. Microbiol. 79, 569–575.

Bird, J., 1957. A whitefly-transmitted mosaic of *Jatropha gossypifolia*, 22nd edition. Agricultural Experiment Station, University of Puerto Rico.

Bird, J., Maramorosch, K., 1978. Viruses and virus diseases associated with whiteflies, in: Lauffer, M.A., Bang, F.B., Maramorosch, K., Smith, K.M. (Eds.), Advances in virus research. Academic Press, pp. 55–110.

Blua, M.J., Perring, T.M., Nuessly, G.S., Duffus, J.E., Toscano, N.C., 1994. Seasonal cropping pattern effects on abundance of *Bemisia tabaci* (Homoptera: Aleyrodidae) and incidence of lettuce infectious yellows virus. Environ. Entomol. 23, 1422–1427.

Bofkin, L., Goldman, N., 2007. Variation in evolutionary processes at different codon positions. Mol. Biol. Evol. 24, 513–521.

Bonato, O., Lurette, A., Vidal, C., Fargues, J., 2007. Modelling temperature-dependent bionomics of *Bemisia tabaci* (Q-biotype). Physiol. Entomol. 32, 50–55.

Bones, A.M., Rossiter, J.T., 1996. The myrosinase-glucosinolate system, its organisation and biochemistry. Physiol. Plant. 97, 194–208.

Bonfield, J.K., Smith, K.F., Staden, R., 1995. A new DNA sequence assembly program. Nucleic Acids Res. 23, 4992–4999.

Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27, 1767–1780.

Boykin, L.M., 2014. *Bemisia tabaci* nomenclature: lessons learned. Pest Manag. Sci. 70, 1454–1459.

Boykin, L.M., Armstrong, K.F., Kubatko, L., De Barro, P.J., 2012. Species delimitation and global biosecurity. Evol. Bioinforma. Online 8, 1–37.

Boykin, L.M., Bell, C.D., Evans, G.A., Small, I., De Barro, P.J., 2013. Is agriculture driving the diversification of the *Bemisia tabaci* species complex (Hemiptera: Sternorrhyncha: Aleyrodidae)?: Dating, diversification and biogeographic evidence revealed. BMC Evol. Biol. 13, 228.

Boykin, L.M., Shatters, R.G., Rosell, R.C., McKenzie, C.L., Bagnall, R.A., De Barro, P.J., Frohlich, D.R., 2007. Global relationships of *Bemisia tabaci* (Hemiptera: Aleyrodidae) revealed using Bayesian analysis of mitochondrial COI DNA sequences. Mol. Phylogenet. Evol. 44, 1306–1319.

Brown, J.K., 2000. Molecular markers for the identification and global tracking of whitefly vector-Begomovirus complexes. Virus Res. 71, 233–260.

Brown, J.K., Bird, J., 1992. Whitefly-transmitted geminiviruses and associated disorders in the Americas and the Caribbean Basin. Plant Dis. 76, 220–225.

Brown, J.K., Coats, S.A., Bedford, I.D., Markham, P.G., Bird, J., Frohlich, D.R., 1995a. Characterization and distribution of esterase electromorphs in the whitefly, *Bemisia tabaci* (Genn.) (Homoptera: Aleyrodidae). Biochem. Genet. 33, 205–214.

Brown, J.K., Frohlich, D.R., Rosell, R.C., 1995b. The sweetpotato or silverleaf whiteflies: Biotypes of *Bemisia tabaci* or a species complex? Annu. Rev. Entomol. 40, 511–534.

Brown, T.A., 2002. Molecular phylogenetics, in: Genomes, 2nd edition. Wiley-Liss, Oxford.

Brumin, M., Kontsedalov, S., Ghanim, M., 2011. *Rickettsia* influences thermotolerance in the whitefly *Bemisia tabaci* B biotype. Insect Sci. 18, 57–66.

Buchner, P., 1965. Endosymbiosis of animals with plant microorganisms. Interscience Publishers, New York.

Buckley, T.R., Simon, C., Chambers, G.K., 2001. Exploring among-site rate variation models in a maximum likelihood framework using empirical data: Effects of model assumptions on estimates of topology, branch lengths, and bootstrap support. Syst. Biol. 50, 67–86.

Buckner, J.S., Freeman, T.P., Ruud, R.L., Chu, C.C., Henneberry, T.J., 2002. Characterization and functions of the whitefly egg pedicel. Arch. Insect Biochem. Physiol. 49, 22–33.

Burban, C., Fishpool, L.D.C., Fauquet, C., Fargette, D., Thouvenel, J.C., 1992. Host-associated biotypes within West African populations of the whitefly *Bemisia tabaci* (Genn.), (Hom., Aleyrodidae). J. Appl. Entomol. 113, 416–423.

Butler, G.D., Henneberry, T.J., Wilson, F.D., 1986. *Bemisia tabaci* (Homoptera: Aleyrodidae) on cotton: Adult activity and cultivar oviposition preference. J. Econ. Entomol. 79, 350–354.

Butter, N.S., Vir, B.K., 1989. Morphological basis of resistance in cotton to the whitefly *Bemisia tabaci*. Phytoparasitica 17, 251.

Byrne, D.N., Hadley, N.F., 1988. Particulate surface waxes of whiteflies: Morphology, composition and waxing behaviour. Physiol. Entomol. 13, 267–276.

Byrne, D.N., Miller, W.B., 1990. Carbohydrate and amino acid composition of phloem sap and honeydew produced by *Bemisia tabaci*. J. Insect Physiol. 36, 433–439.

Byrne, F.J., Bedford, I.D., Devonshire, A.L., Markham, P.G., 1995a. Esterase variation and squash induction in B-type *Bemisia tabaci* (Homoptera: Aleyrodidae). Bull. Entomol. Res. 85, 175–179.

Byrne, F.J., Cahill, M., Denholm, I., Devonshire, A.L., 1995b. Biochemical identification of interbreeding between B-type and non B-type strains of the tobacco whitefly *Bemisia tabaci*. Biochem. Genet. 33, 13–23.

Byrne, F.J., Devonshire, A.L., 1996. Biochemical evidence of haplodiploidy in the whitefly *Bemisia tabaci*. Biochem. Genet. 34, 93–107.

Byrne, F.J., Devonshire, A.L., 1993. Insensitive acetylcholinesterase and esterase polymorphism in susceptible and resistant populations of the tobacco whitefly *Bemisia tabaci* (Genn.). Pestic. Biochem. Physiol. 45, 34–42.

Byrne, F.J., Gorman, K.J., Cahill, M., Denholm, I., Devonshire, A.L., 2000. The role of B-type esterases in conferring insecticide resistance in the tobacco whitefly, *Bemisia tabaci* (Genn). Pest Manag. Sci. 56, 867–874.

CABI, 2018. Bemisia tabaci. In: Invasive species compendium. Wallingford, UK: CAB International. www.cabi.org/isc.

Cahill, M., Denholm, I., Ross, G., Gorman, K., Johnston, D., 1996. Relationship between bioassay data and the simulated field performance of insecticides against susceptible and resistant adult *Bemisia tabaci* (Homoptera: Aleyrodidae). Bull. Entomol. Res. 86, 109–116.

Calvo, F.J., Bolckmans, K., Belda, J.E., 2011. Control of *Bemisia tabaci* and *Frankliniella occidentalis* in cucumber by Amblyseius swirskii. BioControl 56, 185–192.

Calvo, J., Bolckmans, K., Stansly, P.A., Urbaneja, A., 2009. Predation by *Nesidiocoris tenuis* on *Bemisia tabaci* and injury to tomato. BioControl 54, 237–246.

Cameron, S.L., 2014. Insect mitochondrial genomics: Implications for evolution and phylogeny. Annu. Rev. Entomol. 59, 95–117.

Campbell, B.C., 1993. Congruent evolution between whiteflies (Homoptera: Aleyrodidae) and their bacterial endosymbionts based on respective 18S and 16S rDNAs. Curr. Microbiol. 26, 129–132.

Campbell, B.C., Duffus, J.E., Baumann, P., 1993. Determining whitefly species. Science 261, 1333; author reply 1334-1335.

Campbell, B.C., Steffen-Campbell, J.D., Gill, R.J., 1994. Evolutionary origin of whiteflies (Hemiptera: Sternorrhyncha: Aleyrodidae) inferred from 18S rDNA sequences. Insect Mol. Biol. 3, 73–88.

Carabalí, A., Bellotti, A.C., Montoya-Lerma, J., Fregene, M., 2010a. *Manihot flabellifolia* Pohl, wild source of resistance to the whitefly *Aleurotrachelus socialis* Bondar (Hemiptera: Aleyrodidae). Crop Prot. 29, 34–38.

Carabalí, A., Bellotti, A.C., Montoya-Lerma, J., Fregene, M., 2010b. Resistance to the whitefly, *Aleurotrachelus socialis*, in wild populations of cassava, *Manihot tristis*. J. Insect Sci. Online 10, 170.

Castle, S.J., 1999. Agricultural intensification and pest outbreaks: A reappraisal of events in the Sudan Gezira. Ann. Entomol. Soc. Am. 92, 840–852.

Castle, S.J., Henneberry, T.J., Toscano, N.C., 1996. Suppression of *Bemisia tabaci* (Homoptera: Aleyrodidae) infestations in cantaloupe and cotton with sprinkler irrigation. Crop Prot. 15, 657–663.

Caterino, M.S., Cho, S., Sperling, F.A., 2000. The current state of insect molecular systematics: A thriving Tower of Babel. Annu. Rev. Entomol. 45, 1–54.

Cathrin, P.B., Ghanim, M., 2014. Recent advances on interactions between the whitefly *Bemisia tabaci* and begomoviruses, with emphasis on *Tomato yellow leaf curl virus*, in: Gaur, R.K.,

Hohn, T., Sharma, P. (Eds.), Plant virus-host interaction. Academic Press, Boston, pp. 79–103.

Cermak, R., 2008. Effect of dietary flavonoids on pathways involved in drug metabolism. Expert Opin. Drug Metab. Toxicol. 4, 17–35.

Channarayappa, C., Shivashankar, G., Muniyappa, V., Frist, R.H., 1992. Resistance of *Lycopersicon* species to *Bemisia tabaci*, a tomato leaf curl virus vector. Can. J. Bot. 70, 2184–2192.

Chen, W., Hasegawa, D.K., Kaur, N., Kliot, A., Pinheiro, P.V., Luan, J., Stensmyr, M.C., Zheng, Y., Liu, W., Sun, H., Xu, Y., Luo, Y., Kruse, A., Yang, X., Kontsedalov, S., Lebedev, G., Fisher, T.W., Nelson, D.R., Hunter, W.B., Brown, J.K., Jander, G., Cilia, M., Douglas, A.E., Ghanim, M., Simmons, A.M., Wintermantel, W.M., Ling, K.S., Fei, Z., 2016. The draft genome of whitefly *Bemisia tabaci* MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. BMC Biol. 14, 110.

Chen, Y., Ye, W., Zhang, Y., Xu, Y., 2015. High speed BLASTN: An accelerated MegaBLAST search tool. Nucleic Acids Res. 43, 7762–7768.

Chermitti, B., Braham, M., Alonso, C., Beitia, F., Cenis, J.L., 1997. Sur la presence en Tunisie des biotypes 'B' et 'non-B' de *Bemisia tabaci* (Homoptera: Aleyrodidae) et de leurs parasitoids associés. IOBCWPRS Bull. 20, 108–113.

Chernomor, O., von Haeseler, A., Minh, B.Q., 2016. Terrace aware data structure for phylogenomic inference from supermatrices. Syst. Biol. 65, 997–1008.

Chew, F.S., Renwick, J.A.A., 1995. Host plant choice in *Pieris* butterflies, in: Carde, R.T., Bell, W.J. (Eds.), Chemical ecology of insects 2. Springer, Boston, Massachusetts, pp. 214–238.

Chiel, E., Gottlieb, Y., Zchori-Fein, E., Mozes-Daube, N., Katzir, N., Inbar, M., Ghanim, M., 2007. Biotype-dependent secondary symbiont communities in sympatric populations of *Bemisia tabaci*. Bull. Entomol. Res. 97, 407–413.

Chiel, E., Inbar, M., Mozes-Daube, N., White, J.A., Hunter, M.S., Zchori-Fein, E., 2009. Assessments of fitness effects by the facultative symbiont *Rickettsia* in the sweetpotato whitefly (Hemiptera: Aleyrodidae). Ann. Entomol. Soc. Am. 102, 413–418.

Chowda-Reddy, R.V., Kirankumar, M., Seal, S.E., Muniyappa, V., Valand, G.B., Govindappa, M.R., Colvin, J., 2012. *Bemisia tabaci* phylogenetic groups in India and the relative transmission efficacy of *Tomato leaf curl Bangalore virus* by an indigenous and an exotic population. J. Integr. Agric. 11, 235–248.

Chu, C.C., Henneberry, T.J., Cohen, A.C., 1995. *Bemisia argentifolii* (Homoptera: Aleyrodidae): Host preference and factors affecting oviposition and feeding site preferences. Environ. Entomol.

Chu, D., Gao, C.S., De Barro, P., Zhang, Y.J., Wan, F.H., Khan, I.A., 2011. Further insights into the strange role of bacterial endosymbionts in whitefly, *Bemisia tabaci*: comparison of secondary symbionts from biotypes B and Q in China. Bull. Entomol. Res. 101, 477–486.

Chu, D., Hu, X., Gao, C., Zhao, H., Nichols, R.L., Li, X., 2012a. Use of mitochondrial cytochrome oxidase I polymerase chain reaction-restriction fragment length polymorphism for identifying subclades of *Bemisia tabaci* Mediterranean group. J. Econ. Entomol. 105, 242–251.

Chu, D., Tao, Y.L., Zhang, Y.J., Wan, F.H., Brown, J.K., 2012b. Effects of host, temperature and relative humidity on competitive displacement of two invasive *Bemisia tabaci* biotypes [Q and B]. Insect Sci. 19, 595–603.

Chu, D., Wan, F.H., Tao, Y.L., Liu, G.X., Fan, Z.X., Bi, Y.P., 2008. Genetic differentiation of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) biotype Q based on mitochondrial DNA markers. Insect Sci. 15, 115–123.

Chu, D., Wan, F.H., Zhang, Y.J., Brown, J.K., 2010a. Change in the biotype composition of *Bemisia tabaci* in Shandong Province of China from 2005 to 2008. Environ. Entomol. 39, 1028–1036.

Chu, D., Zhang, Y.J., Brown, J.K., Cong, B., Xu, B.Y., Wu, Q.J., Zhu, G.R., 2006. The introduction of the exotic Q biotype of *Bemisia tabaci* from the Mediterranean region into China on ornamental crops. Fla. Entomol. 89, 168–174.

Chu, D., Zhang, Y.J., Wan, F.H., 2010b. Cryptic invasion of the exotic *Bemisia tabaci* biotype Q occurred widespread in Shandong province of China. Fla. Entomol. 93, 203–207.

Claridge, M.F., Hollander, J.D., 1983. The biotype concept and its application to insect pests of agriculture. Crop Prot. 2, 85–95.

Clark, E.L., Karley, A.J., Hubbard, S.F., 2010. Insect endosymbionts: Manipulators of insect herbivore trophic interactions? Protoplasma 244, 25–51.

Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2016. GenBank. Nucleic Acids Res. 44, D67–D72.

Claudianos, C., Russell, R.J., Oakeshott, J.G., 1999. The same amino acid substitution in orthologous esterases confers organophosphate resistance on the house fly and a blowfly. Insect Biochem. Mol. Biol. 29, 675–686.

Cline, J., Braman, J.C., Hogrefe, H.H., 1996. PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. Nucleic Acids Res. 24, 3546–3551.

Cock, M.J.W., 1986. *Bemisia tabaci*: A literature survey on the cotton whitefly with an annotated bibliography. C.A.B. International Institute of Biological Control, Silwood Park, UK, pp. 17–26.

Cohen, A.C., Henneberry, T.J., Chu, C.C., 1996. Geometric relationships between whitefly feeding behavior and vascular bundle arrangements. Entomol. Exp. Appl. 78, 135–142.

Cohen, S., Antignus, Y., 1994. Tomato yellow leaf curl virus, a whitefly-borne geminivirus of tomatoes, in: Harris, K.F. (Ed.), Advances in disease vector research. Springer, New York, pp. 259–288.

Cohen, S., Berlinger, M.J., 1986. Transmission and cultural control of whitefly-borne viruses, in: Proceedings of a Symposium on *Bemisia tabaci* - Ecology and Control XVIIth International Congress of Entomology. pp. 89–97.

Collins, R.A., Boykin, L.M., Cruickshank, R.H., Armstrong, K.F., 2012. Barcoding's next top model: An evaluation of nucleotide substitution models for specimen identification. Methods Ecol. Evol. 3, 457–465.

Collins, R.A., Cruickshank, R.H., 2013. The seven deadly sins of DNA barcoding. Mol. Ecol. Resour. 13, 969–975.

Cosmides, L.M., Tooby, J., 1981. Cytoplasmic inheritance and intragenomic conflict. J. Theor. Biol. 89, 83–129.

Costa, H.S., Brown, J.K., 1991. Variation in biological characteristics and esterase patterns among populations of *Bemisia tabaci*, and the association of one population with silverleaf symptom induction. Entomol. Exp. Appl. 61, 211–219.

Costa, H.S., Toscano, N.C., Hendrix, D.L., Henneberry, T.J., 1999. Patterns of honeydew droplet production by nymphal stages of *Bemisia argentifolii* (Homoptera: Aleyrodidae) and relative composition of honeydew sugars. J. Entomol. Sci. 34, 305–313.

Costa, H.S., Toscano, N.C., Henneberry, T.J., 1996. Mycetocyte inclusion in the oocytes of *Bemisia argentifolii* (Homoptera: Aleyrodidae). Ann. Entomol. Soc. Am. 89, 694–699.

Costa, H.S., Ullman, D.E., Johnson, M.W., Tabashnik, B.E., 1993a. Squash silverleaf symptoms induced by immature, but not adult, *Bemisia tabaci*. Phytopathol. USA.

Costa, H.S., Westcot, D.M., Ullman, D.E., Johnson, M.W., 1993b. Ultrastructure of the endosymbionts of the whitefly, *Bemisia tabaci* and *Trialeurodes vaporariorum*. Protoplasma 176, 106–115.

Coudriet, D.L., Prabhaker, N., Kishaba, A.N., Meyerdirk, D.E., 1985. Variation in developmental rate on different hosts and overwintering of the sweetpotato whitefly, *Bemisia tabaci* (Homoptera: Aleyrodidae). Environ. Entomol. 14, 516–519.

Coyne, J.A., 1994. Ernst Mayr and the origin of species. Evol. Int. J. Org. Evol. 48, 19–30.

Coyne, J.A., Orr, H.A., 1989. Patterns of speciation in *Drosophila*. Evol. Int. J. Org. Evol. 43, 362–381.

Cradock, K.R., da Graça, J.V., Laing, M.D., 2002. Studies on the control of virus diseases in zucchini crops. South African J. Sci. 98: 225-227. South Afr. J. Sci. 98, 225–227.

Crowder, D.W., Horowitz, A.R., De Barro, P.J., Liu, S.S., Showalter, A.M., Kontsedalov, S., Khasdan, V., Shargal, A., Liu, J., Carrière, Y., 2010a. Mating behaviour, life history and adaptation to insecticides determine species exclusion between whiteflies. J. Anim. Ecol. 79, 563–570.

Crowder, D.W., Sitvarin, M.I., Carrière, Y., 2010b. Plasticity in mating behaviour drives asymmetric reproductive interference in whiteflies. Anim. Behav. 79, 579–587.

Cui, X., Wan, F., Xie, M., Liu, T., 2008. Effects of heat shock on survival and reproduction of two whitefly species, *Trialeurodes vaporariorum* and *Bemisia tabaci* biotype B. J. Insect Sci. 8.

Cuthbertson, A.G.S., 2014. The feeding rate of predatory mites on life stages of *Bemisia tabaci* Mediterranean species. Insects 5, 609–614.

Cuthbertson, A.G.S., Blackburn, L.F., Eyre, D.P., Cannon, R.J.C., Miller, J., Northing, P., 2010. *Bemisia tabaci*: The current situation in the UK and the prospect of developing strategies for eradication using entomopathogens. Insect Sci. 18, 1–10.

Cuthbertson, A.G.S., Buxton, J.H., Blackburn, L.F., Mathers, J.J., Robinson, K.A., Powell, M.E., Fleming, D.A., Bell, H.A., 2012. Eradicating *Bemisia tabaci* Q biotype on poinsettia plants in the UK. Crop Prot. 42, 42–48.

Cuthbertson, A.G.S., Collins, D.A., 2015. Tri-Tek (petroleum horticultural oil) and *Beauveria bassiana*: Use in eradication strategies for *Bemisia tabaci* Mediterranean species in UK glasshouses. Insects 6, 133–140.

Cuthbertson, A.G.S., Head, J., Walters, K.F.A., Gregory, S.A., 2003a. The efficacy of the entomopathogenic nematode, *Steinernema feltiae*, against the immature stages of *Bemisia tabaci*. J. Invertebr. Pathol. 83, 267–269.

Cuthbertson, A.G.S., Head, J., Walters, K.F.A., Murray, A.W.A., 2003b. The integrated use of chemical insecticides and the entomopathogenic nematode, *Steinernema feltiae*, for the control of sweetpotato whitefly, *Bemisia tabaci*. Nematology 5, 713–720.

Cuthbertson, A.G.S., Mathers, J.J., Northing, P., Prickett, A.J., Walters, K.F.A., 2008. The integrated use of chemical insecticides and the entomopathogenic nematode, *Steinernema carpocapsae* (Nematoda: Steinernematidae), for the control of sweetpotato whitefly, *Bemisia tabaci* (Hemiptera: Aleyrodidae). Insect Sci. 15, 447–453.

Cuthbertson, A.G.S., Mathers, J.J., Northing, P., WeiQi, L., Walters, K.F.A., 2007a. The susceptibility of immature stages of *Bemisia tabaci* to infection by the entomopathogenic nematode *Steinernema carpocapsae*. Russ. J. Nematol. 15, 153–156.

Cuthbertson, A.G.S., Vänninen, I., 2015. The importance of maintaining protected zone status against *Bemisia tabaci*. Insects 6, 432–441.

Cuthbertson, A.G.S., Walters, K.F.A., Northing, P., Luo, W., 2007b. Efficacy of the entomopathogenic nematode, *Steinernema feltiae*, against sweetpotato whitefly *Bemisia tabaci* (Homoptera: Aleyrodidae) under laboratory and glasshouse conditions. Bull. Entomol. Res. 97, 9–14.

Czosnek, H., 2007. *Tomato yellow leaf curl virus* disease: Management, molecular biology, breeding for resistance. Springer Netherlands.

Dadd, R.H., 1985. Nutrition: Organisms, in: Kerkut, G.A., Gilbert, L.I. (Eds.), Comprehensive insect physiology, biochemistry and pharmacology. Pergamon Press, Elmsford, New York, pp. 315–319.

Dalén, L., Lagerholm, V.K., Nylander, J.A.A., Barton, N., Bochenski, Z.M., Tomek, T., Rudling, D., Ericson, E., P.G.P., Irestedt, M., Stewart, J.R., 2017. Identifying bird remains using ancient DNA barcoding. Genes 8.

Dalmon, A., Halkett, F., Granier, M., Delatte, H., Peterschmitt, M., 2008. Genetic structure of the invasive pest *Bemisia tabaci*: Evidence of limited but persistent genetic differentiation in glasshouse populations. Heredity 100, 316–325.

Dalton, R., 2006. Whitefly infestations: The Christmas invasion. Nature 443, 898–900.

De Barro, P.J., 2012. The *Bemisia tabaci* species complex: Questions to guide future research. J. Integr. Agric. 11, 187–196.

De Barro, P.J., Ahmed, M.Z., 2011. Genetic networking of the *Bemisia tabaci* cryptic species complex reveals pattern of biological invasions. PLoS One 6.

De Barro, P.J., Bourne, A., Khan, S.A., Brancatini, V. a. L., 2006. Host plant and biotype density interactions – their role in the establishment of the invasive B biotype of *Bemisia tabaci*. Biol. Invasions 8, 287–294.

De Barro, P.J., Driver, F., 1997. Use of RAPD PCR to distinguish the B biotype from other biotypes of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae). Aust. J. Entomol. 36, 149–152.

De Barro, P.J., Driver, F., Trueman, J.W., Curran, J., 2000. Phylogenetic relationships of world populations of *Bemisia tabaci* (Gennadius) using ribosomal ITS1. Mol. Phylogenet. Evol. 16, 29–36.

De Barro, P.J., Hart, P.J., 2000. Mating interactions between two biotypes of the whitefly, *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Australia. Bull. Entomol. Res. 90, 103–112.

De Barro, P.J., Liebregts, W., Carver, M., 1998. Distribution and identity of biotypes of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) in member countries of the Secretariat of the Pacific Community. Aust. J. Entomol. 37, 214–218.

De Barro, P.J., Liu, S.S., Boykin, L.M., Dinsdale, A.B., 2011. *Bemisia tabaci*: A statement of species status. Annu. Rev. Entomol. 56, 1–19.

De Barro, P.J., Trueman, J.W.H., Frohlich, D.R., 2005. *Bemisia argentifolii* is a race of *B. tabaci* (Hemiptera: Aleyrodidae): The molecular genetic differentiation of B. tabaci populations around the world. Bull. Entomol. Res. 95, 193–203.

De la Rúa, P., Simón, B., Cifuentes, D., Martinez-Mora, C., Cenis, J.L., 2006. New insights into the mitochondrial phylogeny of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) in the Mediterranean Basin. J. Zool. Syst. Evol. Res. 44, 25–33.

De Ponti, O.M.B., Romanow, L.R., Berlinger, M.J., 1990. Whitefly-plant relationships: Plant resistance, in: Gerling, D. (Ed.), Whiteflies: Their bionomics, pest status and management. Intercept, Andover, UK, pp. 91–106.

De Queiroz, K., 2007. Species concepts and species delimitation. Syst. Biol. 56, 879–886.

Delatte, H., David, P., Granier, M., Lett, J.M., Goldbach, R., Peterschmitt, M., Reynaud, B., 2006. Microsatellites reveal extensive geographical, ecological and genetic contacts between invasive and indigenous whitefly biotypes in an insular environment. Genet. Res. 87, 109–124.

Delatte, H., Reynaud, B., Granier, M., Thornary, L., Lett, J.M., Goldbach, R., Peterschmitt, M., 2005. A new silverleaf-inducing biotype Ms of *Bemisia tabaci* (Hemiptera: Aleyrodidae) indigenous to the islands of the south-west Indian Ocean. Bull. Entomol. Res. 95, 29–35.

Demichelis, S., Arnò, C., Bosco, D., Marian, D., Caciagli, P., 2005. Characterization of biotype T of *Bemisia tabaci* associated with *Euphorbia characias* in Sicily. Phytoparasitica 33, 196–208.

Demichelis, S., Bosco, D., Manino, A., Marian, D., Caciagli, P., 2000. Distribution of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes in Italy. Can. Entomol. 132, 519–527.

Dennehy, T.J., DeGain, B.A., Harpold, V.S., Brown, J.K., Morin, S., Fabrick, J.A., Byrne, F.J., Nichols, R.L., 2005. New challenges to management of whitefly resistance to insecticides in Arizona. Univ. Ariz. Coop. Ext. Veg. Rep. 31.

Dennehy, T.J., Degain, B.A., Harpold, V.S., Zaborac, M., Morin, S., Fabrick, J.A., Nichols, R.L., Brown, J.K., Byrne, F.J., Li, X., 2010. Extraordinary resistance to insecticides reveals exotic Q biotype of *Bemisia tabaci* in the New World. J. Econ. Entomol. 103, 2174–2186.

Denno, R.F., McClure, M.S., Ott, J.R., 1995. Interspecific interactions in phytophagous insects: Competition reexamined and resurrected. Annu. Rev. Entomol. 40, 297–331.

DeSalle, R., Egan, M.G., Siddall, M., 2005. The unholy trinity: Taxonomy, species delimitation and DNA barcoding. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 1905–1916.

Després, L., David, J.P., Gallet, C., 2007. The evolutionary ecology of insect resistance to plant chemicals. Trends Ecol. Evol. 22, 298–307.

Dicke, M., 2000. Chemical ecology of host-plant selection by herbivorous arthropods: A multitrophic perspective. Biochem. Syst. Ecol. 28, 601–617.

Diehl, S.R., Bush, G.L., 1984. An evolutionary and applied perspective of insect biotypes. Annu. Rev. Entomol. 29, 471–504.

Dinsdale, A., Cook, L., Riginos, C., Buckley, Y.M., De Barro, P.J., 2010. Refined global analysis of *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. Ann. Entomol. Soc. Am. 103, 196–208.

Dittrich, V., Ernst, G.H., Ruesch, O., Uk, S., 1990. Resistance mechanisms in sweetpotato whitefly (Homoptera: Aleyrodidae) populations from Sudan, Turkey, Guatemala, and Nicaragua. J. Econ. Entomol. 83, 1665–1670.

Dobbs E.M., 2011. SketchAndCalc. http://www.SketchAndCalc.com. Accessed 8 March 2018.

Dong, T.G., Zhang, B., Jiang, Y.F., Hu, Q.B., 2016. Isolation and classification of fungal whitefly entomopathogens from soils of Qinghai-Tibet Plateau and Gansu Corridor in China. PLoS One 11, e0156087.

Douglas, A.E., 1998. Nutritional interactions in insect-microbial symbioses: Aphids and their symbiotic bacteria *Buchnera*. Annu. Rev. Entomol. 43, 17–37.

Downie, D.A., 2010. Baubles, bangles, and biotypes: A critical review of the use and abuse of the biotype concept. J. Insect Sci. Online 10, 176.

du Rand, E.E., Smit, S., Beukes, M., Apostolides, Z., Pirk, C.W.W., Nicolson, S.W., 2015. Detoxification mechanisms of honey bees (*Apis mellifera*) resulting in tolerance of dietary nicotine. Sci. Rep. 5, 11779.

Duffus, J.E., Flock, R.A., 1982. Whitefly-transmitted disease complex of the desert southwest. Calif. Agric. 36, 4–6.

Eakteiman, G., Moses-Koch, R., Moshitzky, P., Mestre-Rincon, N., Vassão, D.G., Luck, K., Sertchook, R., Malka, O., Morin, S., 2018. Targeting detoxification genes by phloem-mediated RNAi: A new approach for controlling phloem-feeding insect pests. Insect Biochem. Mol. Biol. 100, 10–21.

Eastop, V.F., 1973. Biotypes of aphids. Bull. Entomol. Soc. N. Z. 2, 40–51.

Elbaz, M., Lahav, N., Morin, S., 2010. Evidence for pre-zygotic reproductive barrier between the B and Q biotypes of *Bemisia tabaci* (Hemiptera: Aleyrodidae). Bull. Entomol. Res. 100, 581–590.

Elbert, A., Nauen, R., 2000. Resistance of *Bemisia tabaci* (Homoptera: Aleyrodidae) to insecticides in southern Spain with special reference to neonicotinoids. Pest Manag. Sci. 56, 60–64.

Elfekih, S., Etter, P., Tay, W.T., Fumagalli, M., Gordon, K.H.J., Johnson, E., De Barro, P.J., 2018. Genome-wide analyses of the *Bemisia tabaci* species complex reveal contrasting patterns of admixture and complex demographic histories. PLoS One 13, e0190555.

Ellsworth, P., Tronstad, R., Leser, J., Goodell, P.B., Godfrey, L.D., Henneberry, T.J., Hendrix, D., Brushwood, D., Naranjo, S.E., Castle, S., Nichols, R.L., 1999. Sticky cotton sources & solutions. Univ. Ariz. Coop. Extention IPM Ser., 13 1–4.

Enders, L.S., Miller, N.J., 2016. Stress-induced changes in abundance differ among obligate and facultative endosymbionts of the soybean aphid. Ecol. Evol. 6, 818–829.

Engelstädter, J., Hurst, G.D.D., 2009. The ecology and evolution of microbes that manipulate host reproduction. Annu. Rev. Ecol. Evol. Syst. 40, 127–149.

Esterhuizen, L.L., Mabasa, K.G., Heerden, V., Willem, S., Czosnek, H., Brown, J.K., Van Heerden, H., Rey, M.E.C., 2013. Genetic identification of members of the *Bemisia tabaci* cryptic species complex from South Africa reveals native and introduced haplotypes. J. Appl. Entomol. 137, 122–135.

Eveleens, K.G., 1983. Cotton-insect control in the Sudan Gezira: Analysis of a crisis. Crop Prot. 2, 273–287.

Fancelli, M., Vendramim, J.D., 2002. Development of *Bemisia tabaci* (Gennadius,1889) biotype B on *Lycopersicon* spp. genotypes. Sci. Agric. 59, 665–669.

Fang, Y.W., Liu, L.Y., Zhang, H.L., Jiang, D.F., Chu, D., 2014. Competitive ability and fitness differences between two introduced populations of the invasive whitefly *Bemisia tabaci* Q in China. PLoS One 9, e100423.

Farrell, B., Mitter, C., 1990. Phylogenesis of insect/plant interactions: Have *Phyllobrotica* leaf beetles (Chrysomelidae) and the Lamiales diversified in parallel? Evolution 44, 1389–1403.

Fauquet, C., Fargette, D., 1990. African cassava mosaic virus : Etiology, epidemiology, and control. Plant Dis. 74, 404–411.

Fekrat, L., Shishehbor, P., 2007. Some biological features of cotton whitefly, *Bemisia tabaci* (Homoptera: Aleyrodidae) on various host plants. Pak. J. Biol. Sci. PJBS 10, 3180–3184.

Felsenstein, J., 1983. Statistical inference of phylogenies. J. R. Stat. Soc. Ser. Gen. 146, 246–272.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 17, 368–376.

Fenn, J.D., Song, H., Cameron, S.L., Whiting, M.F., 2008. A preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. Mol. Phylogenet. Evol. 49, 59–68.

Fernández, E., Grávalos, C., Haro, P.J., Cifuentes, D., Bielza, P., 2009. Insecticide resistance status of *Bemisia tabaci* Q-biotype in south-eastern Spain. Pest Manag. Sci. 65, 885–891.

Field, L.M., 2000. Methylation and expression of amplified esterase genes in the aphid *Myzus persicae* (Sulzer). Biochem. J. 349, 863–868.

Firdaus, S., van Heusden, A.W., Hidayati, N., Supena, E.D.J., Visser, R.G.F., Vosman, B., 2012. Resistance to *Bemisia tabaci* in tomato wild relatives. Euphytica 187, 31–45.

Firdaus, S., Vosman, B., Hidayati, N., Jaya Supena, E.D., Visser, R.G.F., van Heusden, A.W., 2013. The *Bemisia tabaci* species complex: Additions from different parts of the world. Insect Sci. 20, 723–733.

Flint, H.M., Naranjo, S.E., Leggett, J.E., Henneberry, T.J., 1996. Cotton water stress, arthropod dynamics, and management of *Bemisia tabaci* (Homoptera: Aleyrodidae). J. Econ. Entomol. 89, 1288–1300.

Fontes, F.H.M., Colombo, C.A., Lourenção, A.L., 2012. Structure of genetic diversity of *Bemisia tabaci* (Genn.) (Hemiptera: Aleyrodidae) populations in Brazilian crops and locations. Sci. Agric. 69, 47–53.

Fransen, J.J., 1990. Natural enemies of whiteflies: Fungi, in: Gerling, D. (ed.), Whiteflies: Their bionomics, pest status and management. Intercept, Andover, UK, pp. 187–210.

Frohlich, D.R., Torres-Jerez, I., Bedford, I.D., Markham, P.G., Brown, J.K., 1999. A phylogeographical analysis of the *Bemisia tabaci* species complex based on mitochondrial DNA markers. Mol. Ecol. 8, 1683–1691.

Fujiwara, A., Maekawa, K., Tsuchida, T., 2015. Genetic groups and endosymbiotic microbiota of the *Bemisia tabaci* species complex in Japanese agricultural sites. J. Appl. Entomol. 139, 55–66.

Garcia-del-Pino, F., Morton, A., Shapiro-Ilan, D., 2018. Entomopathogenic nematodes as biological control agents of tomato pests, in: Wakil, W., Brust, G.E., Perring, T.M. (Eds.), Sustainable management of arthropod pests of tomato. Academic Press, San Diego, pp. 269–282.

Gaur, R.K., Hohn, T., Sharma, P., 2013. Plant virus-host interaction: Molecular approaches and viral evolution. Elsevier.

Gauthier, N., Clouet, C., Perrakis, A., Kapantaidaki, D., Peterschmitt, M., Tsagkarakou, A., 2014. Genetic structure of *Bemisia tabaci* MED populations from home-range countries, inferred by nuclear and cytoplasmic markers: Impact on the distribution of the insecticide resistance genes. Pest Manag. Sci. 70, 1477–1491.

Gawel, N.J., Bartlett, A.C., 1993. Characterization of differences between whiteflies using RAPD-PCR. Insect Mol. Biol. 2, 33–38.

Gelman, D.B., Blackburn, M.B., Hu, J.S., Gerling, D., 2002. The nymphal-adult molt of the silverleaf whitefly (*Bemisia argentifolii*): Timing, regulation, and progress. Arch. Insect Biochem. Physiol. 51, 67–79.

Gennadius, P., 1889. Disease of tobacco plantations in the Trikonia. The aleurodid of tobacco. Ellenike Ga. 5, 1–3.

Gerling, D., 1984. The overwintering mode of *Bemisia tabaci* and its parasitoids in Israel. Phytoparasitica 12, 109.

Gerling, D., Alomar, Ò., Arnó, J., 2001. Biological control of *Bemisia tabaci* using predators and parasitoids. Crop Prot., Challenges and opportunities for pest management of *Bemisia tabaci* in the new century 20, 779–799.

Gerling, D., Horowitz, A.R., Baumgaertner, J., 1986. Autecology of *Bemisia tabaci*, in: Proceedings of a Symposium on *Bemisia tabaci* - Ecology and Control XVIIth International Congress of Entomology, Proceedings of a Symposium on *Bemisia tabaci* - Ecology and Control XVIIth International Congress of Entomology. pp. 5–19.

Gershoni, M., Templeton, A.R., Mishmar, D., 2009. Mitochondrial bioenergetics as a major motive force of speciation. BioEssays 31, 642–650.

Ghanim, M., Kontsedalov, S., 2009. Susceptibility to insecticides in the Q biotype of *Bemisia tabaci* is correlated with bacterial symbiont densities. Pest Manag. Sci. 65, 939–942.

Ghosh, S., 2016. Diversity and role of bacterial endosymbionts infecting the cassava whitefly, *Bemisia tabaci* (Gennadius) in sub-Saharan Africa (Dissertation). University of Greenwich, Chatham, UK.

Ghosh, S., Bouvaine, S., Maruthi, M.N., 2015. Prevalence and genetic diversity of endosymbiotic bacteria infecting cassava whiteflies in Africa. BMC Microbiol. 15.

Gilbertson, R.L., Batuman, O., Webster, C.G., Adkins, S., 2015. Role of the insect supervectors *Bemisia tabaci* and *Frankliniella occidentalis* in the emergence and global spread of plant viruses. Annu. Rev. Virol. 2, 67–93.

Gnankiné, O., Ketoh, G., Martin, T., 2013. Dynamics of the invasive *Bemisia tabaci* (Homoptera: Aleyrodidae) Mediterranean (MED) species in two West African countries. Int. J. Trop. Insect Sci. 33, 99–106.

Godfray, H.C.J., 1994. Parasitoids: Behavioral and evolutionary ecology. Princeton University Press.

Gottlieb, Y., Ghanim, M., Chiel, E., Gerling, D., Portnoy, V., Steinberg, S., Tzuri, G., Horowitz, A.R., Belausov, E., Mozes-Daube, N., Kontsedalov, S., Gershon, M., Gal, S., Katzir, N., Zchori-Fein, E., 2006. Identification and localization of a *Rickettsia* sp. in *Bemisia tabaci* (Homoptera: Aleyrodidae). Appl. Environ. Microbiol. 72, 3646–3652.

Gottlieb, Y., Zchori-Fein, E., Mozes-Daube, N., Kontsedalov, S., Skaljac, M., Brumin, M., Sobol, I., Czosnek, H., Vavre, F., Fleury, F., Ghanim, M., 2010. The transmission efficiency of *Tomato Yellow Leaf Curl Virus* by the whitefly *Bemisia tabaci* is correlated with the presence of a specific symbiotic bacterium species. J. Virol. 84, 9310–9317.

Greathead, A.H., 1986. Host plants, in: Cock, M.J.W. (Ed.), *Bemisia tabaci* - A literature survey on the cotton whitefly with an annotated bibliography. C.A.B International Institute of Biological Control, Silwood Park, UK, pp. 17–26.

Grille, G., Gauthier, N., Buenahora, J., Basso, C., Bonato, O., 2011. First report of the Q biotype of *Bemisia tabaci* in Argentina and Uruguay. Phytoparasitica 39, 235.

Groves, C., Grubb, P., 2011. Ungulate taxonomy. The John Hopkins University Press, Baltimore, Maryland.

Gueguen, G., Vavre, F., Gnankiné, O., Peterschmitt, M., Charif, D., Chiel, E., Gottlieb, Y., Ghanim, M., Zchori-Fein, E., Fleury, F., 2010. Endosymbiont metacommunities, mtDNA diversity and the evolution of the *Bemisia tabaci* (Hemiptera: Aleyrodidae) species complex. Mol. Ecol. 19, 4365–4376.

Guershon, M., Gerling, D., 1999. Predatory behavior of *Delphastus pusillus* in relation to the phenotypic plasticity of *Bemisia tabaci* nymphs. Entomol. Exp. Appl. 92, 239–248.

Guirao, P., Beitia, F., Cenis, J.L., 1997. Biotype determination of Spanish populations of *Bemisia tabaci* (Hemiptera: Aleyrodidae). Bull. Entomol. Res. 87, 587–593.

Guo, L., Xie, W., Wang, S., Wu, Q., Li, R., Yang, N., Yang, X., Pan, H., Zhang, Y., 2014. Detoxification enzymes of *Bemisia tabaci* B and Q: Biochemical characteristics and gene expression profiles. Pest Manag. Sci. 70, 1588–1594.

Guo, T., Guo, Q., Cui, X.Y., Liu, Y.Q., Hu, J., Liu, S.S., 2015. Comparison of transmission of *Papaya leaf curl China virus* among four cryptic species of the whitefly *Bemisia tabaci* complex. Sci. Rep. 5, 15432.

Hadjistylli, M., Roderick, G.K., Brown, J.K., 2016. Global population structure of a worldwide pest and virus vector: Genetic diversity and population history of the *Bemisia tabaci* sibling species group. PLoS One 11, e0165105.

Haegeman, A., Vanholme, B., Jacob, J., Vandekerckhove, T.T.M., Claeys, M., Borgonie, G., Gheysen, G., 2009. An endosymbiotic bacterium in a plant-parasitic nematode: Member of a new *Wolbachia* supergroup. Int. J. Parasitol. 39, 1045–1054.

Halon, E., Eakteiman, G., Moshitzky, P., Elbaz, M., Alon, M., Pavlidi, N., Vontas, J., Morin, S., 2015. Only a minority of broad-range detoxification genes respond to a variety of phytotoxins in generalist *Bemisia tabaci* species. Sci. Rep. 5, 17975.

Hanif-Khan, S., Bullock, R.C., Stoffella, P.J., Powell, C.A., Brecht, J.K., McAuslane, H.J., Yokomi, R.K., 1998. Tomato irregular-ripening symptom development and ripening of silverleaf whitefly-infested dwarf cherry tomatoes. J. Am. Soc. Hortic. Sci. 123, 119–125.

Hanssen, I.M., Lapidot, M., Thomma, B.P.H.J., 2010. Emerging viral diseases of tomato crops. Mol. Plant-Microbe Interact. MPMI 23, 539–548.

Harris, D.J., 2003. Can you bank on GenBank? Trends Ecol. Evol. 18, 317–319.

Harrison, R.G., 1998. Linking evolutionary pattern and process: The relevance of species concepts for the study of speciation, in: Howard, D.J., Berlocher, S.H. (Eds.), Endless Forms: Species and Speciation. Oxford University Press, Oxford, pp. 19–31.

Harrison, R.G., 1990. Hybrid zones: Windows on evolutionary process. Oxf. Surv. Evol. Biol. 7, 69–128.

Harrison, R.G., Larson, E.L., 2014. Hybridization, introgression, and the nature of species boundaries. J. Hered. 105 Suppl 1, 795–809.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.

Head, J., Lawrence, A.J., Walters, K.F.A., 2004. Efficacy of the entomopathogenic nematode, *Steinernema feltiae*, against *Bemisia tabaci* in relation to plant species. J. Appl. Entomol. 128, 543–547.

Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003a. Biological identifications through DNA barcodes. Proc R Soc Lond B 270, 313–321.

Hebert, P.D.N., Ratnasingham, S., deWaard, J.R., 2003b. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. Proc. R. Soc. B Biol. Sci. 270, S96–S99.

Heddi, A., Grenier, A.M., Khatchadourian, C., Charles, H., Nardon, P., 1999. Four intracellular genomes direct weevil biology: Nuclear, mitochondrial, principal endosymbiont, and *Wolbachia*. Proc. Natl. Acad. Sci. 96, 6814–6819.

Hendry, T.A., Hunter, M.S., Baltrus, D.A., 2014. The facultative symbiont *Rickettsia* protects an invasive whitefly against entomopathogenic *Pseudomonas syringae* strains. Appl. Environ. Microbiol. 80, 7161–7168.

Hennig, W., 1966. Phylogenetic systematics. University of Illinois Press.

Hilje, L., Costa, H.S., Stansly, P.A., 2001. Cultural practices for managing *Bemisia tabaci* and associated viral diseases. Crop Prot., Challenges and opportunities for pest management of *Bemisia tabaci* in the new century 20, 801–812.

Hillocks, R.J., Raya, M.D., Mtunda, K., Kiozia, H., 2001. Effects of brown streak virus disease on yield and quality of cassava in Tanzania. J. Phytopathol. 149, 389–394.

Himler, A.G., Adachi-Hagimori, T., Bergen, J.E., Kozuch, A., Kelly, S.E., Tabashnik, B.E., Chiel, E., Duckworth, V.E., Dennehy, T.J., Zchori-Fein, E., Hunter, M.S., 2011. Rapid spread of a bacterial symbiont in an invasive whitefly is driven by fitness benefits and female bias. Science 332, 254–256.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.

Holt, J., Colvin, J., 2001. Observation and theory of whitefly-borne virus disease epidemics, in: Jeger, M.J., Spence, N.J. (Eds.), Biotic Interactions in plant-pathogen associations. CABI, pp. 331–343.

Hopkins, R.J., van Dam, N.M., van Loon, J.J.A., 2009. Role of glucosinolates in insect-plant relationships and multitrophic interactions. Annu. Rev. Entomol. 54, 57–83.

Horowitz, A.R., Denholm, I., Gorman, K., Cenis, J.L., Kontsedalov, S., Ishaaya, I., 2003a. Biotype Q of *Bemisia tabaci* identified in Israel. Phytoparasitica 31, 94.

Horowitz, A.R., Gorman, K., Ross, G., Denholm, I., 2003b. Inheritance of pyriproxyfen resistance in the whitefly, *Bemisia tabaci* (Q biotype). Arch. Insect Biochem. Physiol. 54, 177–186.

Horowitz, A.R., Kontsedalov, S., Khasdan, V., Ishaaya, I., 2005. Biotypes B and Q of *Bemisia tabaci* and their relevance to neonicotinoid and pyriproxyfen resistance. Arch. Insect Biochem. Physiol. 58, 216–225.

Hosokawa, T., Koga, R., Kikuchi, Y., Meng, X.Y., Fukatsu, T., 2010. *Wolbachia* as a bacteriocyte-associated nutritional mutualist. Proc. Natl. Acad. Sci. 107, 769–774.

Hothorn, T., Bretz, F., Westfall, P., 2008. Simultaneous inference in general parametric models. Biom. J. Biom. Z. 50, 346–363.

Hsieh, C.H., Ko, C.C., Chung, C.H., Wang, H.Y., 2014. Multilocus approach to clarify species status and the divergence history of the *Bemisia tabaci* (Hemiptera: Aleyrodidae) species complex. Mol. Phylogenet. Evol. 76, 172–180.

Hu, J., De Barro, P.J., Zhao, H., Wang, J., Nardi, F., Liu, S.S., 2011. An extensive field survey combined with a phylogenetic analysis reveals rapid and widespread invasion of two alien whiteflies in China. PLoS One 6, e16061.

Hu, J., Jiang, Z.L., Nardi, F., Liu, Y.Y., Luo, X.R., Li, H.X., Zhang, Z.K., 2014. Members of *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species and the status of two invasive alien species in the Yunnan Province (China). J. Insect Sci. Online 14, 281.

Hu, J., Zhang, X., Jiang, Z., Zhang, F., Liu, Y., Li, Z., Zhang, Z., 2017. New putative cryptic species detection and genetic network analysis of *Bemisia tabaci* (Hempitera: Aleyrodidae) in China based on mitochondrial COI sequences. Mitochondrial DNA Part DNA Mapp. Seq. Anal. 29, 474–484.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Hunter, M.S., Perlman, S.J., Kelly, S.E., 2003. A bacterial symbiont in the *Bacteroidetes* induces cytoplasmic incompatibility in the parasitoid wasp *Encarsia pergandiella*. Proc. Biol. Sci. 270, 2185–2190.

Husain, M.A., Trehan, K.N., 1933. Observations on the life-history, bionomics and control of the whitefly on cotton. (*Bemisia gossypiperda* M. & L.). Indian J. Agric. Sci. 3, 701–753.

Huson, D.H., Scornavacca, C., 2012. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. Syst. Biol. 61, 1061–1067.

Iida, H., Kitamura, T., Honda, K.I., 2009. Comparison of egg-hatching rate, survival rate and development time of the immature stage between B- and Q-biotypes of *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae) on various agricultural crops. Appl. Entomol. Zool. 44, 267–273.

Inbar, M., Gerling, D., 2008. Plant-mediated interactions between whiteflies, herbivores, and natural enemies. Annu. Rev. Entomol. 53, 431–448.

Isaac, N.J.B., Mallet, J., Mace, G.M., 2004. Taxonomic inflation: Its influence on macroecology and conservation. Trends Ecol. Evol. 19, 464–469.

Jacq, C., Miller, J.R., Brownlee, G.G., 1977. A pseudogene structure in 5S DNA of *Xenopus laevis*. Cell 12, 109–120.

Jadhav, R., Nair, R.J., 2015. Pest blights India's GM cotton crop, fuelling debate over risks. Reuters.

Jaenike, J., 1978. On optimal oviposition behavior in phytophagous insects. Theor. Popul. Biol. 14, 350–356.

Jauset, A.M., Sarasúa, M.J., Avilla, J., Albajes, R., 2000. Effect of nitrogen fertilization level applied to tomato on the greenhouse whitefly. Crop Prot. 19, 255–261.

Jiang, H., Lei, R., Ding, S.W., Zhu, S., 2014. Skewer: A fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15, 182.

Jiang, Y.X., Lei, H., Collar, J.L., Martin, B., Muñiz, M., Fereres, A., 1999. Probing and feeding behavior of two distinct biotypes of *Bemisia tabaci* (Homoptera: Aleyrodidae) on tomato plants. J. Econ. Entomol. 92, 357–366.

Jiang, Y.X., Walker, G.P., 2007. Identification of phloem sieve elements as the site of resistance to silverleaf whitefly in resistant alfalfa genotypes. Entomol. Exp. Appl. 125, 307–320.

Jiao, X., Xie, W., Wang, S., Wu, Q., Zhou, L., Pan, H., Liu, B., Zhang, Y., 2012. Host preference and nymph performance of B and Q putative species of *Bemisia tabaci* on three host plants. J. Pest Sci. 85, 423–430.

Jiao, X.G., Xie, W., Wang, S.L., Wu, Q.J., Pan, H.P., Liu, B.M., Zhang, Y.J., 2013. Differences in host selection and performance between B and Q putative species of *Bemisia tabaci* on three host plants. Entomol. Exp. Appl. 147, 1–8.

Jones, D.R., 2003. Plant viruses transmitted by whiteflies. Eur. J. Plant Pathol. 109, 195–219.

Jukes, T.H., Cantor, C.R., 1969. Evolution of protein molecules, In: Munro, H.N. (Ed.), Mammalian protein metabolism. Academic Press, New York, pp. 21–132.

Kainer, D., Lanfear, R., 2015. The effects of partitioning on phylogenetic inference. Mol. Biol. Evol. 32, 1611–1627.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Karban, R., Myers, J.H., 1989. Induced plant responses to herbivory. Annu. Rev. Ecol. Syst. 20, 331–348.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinforma. Oxf. Engl. 28, 1647–1649.

Keohavong, P., Thilly, W.G., 1989. Fidelity of DNA polymerases in DNA amplification. Proc. Natl. Acad. Sci. 86, 9253–9257.

Khan, I.A., Wan, F.H., 2015. Life history of *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae) biotype B on tomato and cotton host plants. J. Entomol. Zool. Stud. 3, 117–121.

Khasdan, V., Levin, I., Rosner, A., Morin, S., Kontsedalov, S., Maslenin, L., Horowitz, A.R., 2005. DNA markers for identifying biotypes B and Q of *Bemisia tabaci* (Hemiptera: Aleyrodidae) and studying population dynamics. Bull. Entomol. Res. 95, 605–613.

Kikuchi, Y., 2009. Endosymbiotic bacteria in insects: Their diversity and culturability. Microbes Environ. 24, 195–204.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

King, A.M.Q., Lefkowitz, E.J., Mushegian, A.R., Adams, M.J., Dutilh, B.E., Gorbalenya, A.E., Harrach, B., Harrison, R.L., Junglen, S., Knowles, N.J., Kropinski, A.M., Krupovic, M., Kuhn, J.H., Nibert, M.L., Rubino, L., Sabanadzovic, S., Sanfaçon, H., Siddell, S.G., Simmonds, P., Varsani, A., Zerbini, F.M., Davison, A.J., 2018. Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2018). Arch. Virol.

Kisha, J.S.A., 1981. Observations on the trapping of the whitefly *Bemisia tabaci* by glandular hairs on tomato leaves. Ann. Appl. Biol. 97, 123–127.

Kliot, A., Cilia, M., Czosnek, H., Ghanim, M., 2014. Implication of the bacterial endosymbiont *Rickettsia* spp. in interactions of the whitefly *Bemisia tabaci* with *Tomato yellow leaf curl virus*. J. Virol. 88, 5652–5660.

Kontsedalov, S., Zchori-Fein, E., Chiel, E., Gottlieb, Y., Inbar, M., Ghanim, M., 2008. The presence of *Rickettsia* is associated with increased susceptibility of *Bemisia tabaci* (Homoptera: Aleyrodidae) to insecticides. Pest Manag. Sci. 64, 789–792.

Krall, S., 1995. Desert locusts in Africa: A disaster? Disasters 19, 1–7.

Kranthi, K.R., 2015. Whitefly - the black story. Cotton Stat. News 23.

Krieger, R.I., Feeny, P.P., Wilkinson, C.F., 1971. Detoxication enzymes in the guts of caterpillars: An evolutionary answer to plant defenses? Science 172, 579–581.

Kring, J.B., Schuster, D.J., Price, J.F., Simone, G.W., 1991. Sweetpotato whitefly-vectored geminivirus on tomato in Florida. Plant Dis. 75.

Kutuk, H., Alaoglu, O., Yigit, A., 2008. Evaluation of the lady beetle, *Serangium parcesetosum* for control of *Bemisia tabaci* on greenhouse eggplant in the Mediterranean region. Biocontrol Sci. Technol. 18, 675–683.

Lacey, L.A., Fransen, J.J., Carruthers, R.I., 1996. Global distribution of naturally occurring fungi of *Bemisia*, their biologies and use as biological control agents, in: Gerling, D., Mayer, R.T. (Eds.), *Bemisia*: 1995 taxonomy, biology, damage, control and management. Intercept, Andover, UK, pp. 401–433.

Landis, D.A., Wratten, S.D., Gurr, G.M., 2000. Habitat management to conserve natural enemies of arthropod pests in agriculture. Annu. Rev. Entomol. 45, 175–201.

Leaché, A.D., Koo, M.S., Spencer, C.L., Papenfuss, T.J., Fisher, R.N., McGuire, J.A., 2009. Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). Proc. Natl. Acad. Sci. 106, 12418–12423.

Leavitt, J.R., Hiatt, K.D., Whiting, M.F., Song, H., 2013. Searching for the optimal data partitioning strategy in mitochondrial phylogenomics: A phylogeny of Acridoidea (Insecta: Orthoptera: Caelifera) as a case study. Mol. Phylogenet. Evol. 67, 494–508.

Lee, W., Park, J., Lee, G.S., Lee, S., Akimoto, S., 2013. Taxonomic status of the *Bemisia tabaci* complex (Hemiptera: Aleyrodidae) and reassessment of the number of its constituent species. PLoS One 8, e63817.

Legaspi, J.C., Legaspi, B.C., Meagher, R.L., Ciomperlik, M.A., 1996. Evaluation of *Serangium parcesetosum* (Coleoptera: Coccinellidae) as a biological control agent of the silverleaf whitefly (Homoptera: Aleyrodidae). Environ. Entomol. 25, 1421–1427.

Legg, J.P., 1999. Emergence, spread and strategies for controlling the pandemic of cassava mosaic virus disease in east and central Africa. Crop Prot. 18, 627–637.

Legg, J.P., Fauquet, C.M., 2004. Cassava mosaic geminiviruses in Africa. Plant Mol. Biol. 56, 585–599.

Legg, J.P., Jeremiah, S.C., Obiero, H.M., Maruthi, M.N., Ndyetabula, I., Okao-Okuja, G., Bouwmeester, H., Bigirimana, S., Tata-Hangy, W., Gashaka, G., Mkamilo, G., Alicai, T., Lava Kumar, P., 2011. Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in Africa. Virus Res. 159, 161–170.

Legg, J.P., Shirima, R., Tajebe, L.S., Guastella, D., Boniface, S., Jeremiah, S., Nsami, E., Chikoti, P., Rapisarda, C., 2014. Biology and management of *Bemisia* whitefly vectors of cassava virus pandemics in Africa. Pest Manag. Sci. 70, 1446–1453.

Lenné, J., 2000. Pests and poverty: The continuing need for crop protection research. Outlook Agric. 29, 235–250.

Li, M., Hu, J., Xu, F.C., Liu, S.S., 2010. Transmission of *Tomato Yellow Leaf Curl Virus* by two invasive biotypes and a Chinese indigenous biotype of the whitefly *Bemisia tabaci*. Int. J. Pest Manag. 56, 275–280.

Li, T.Y., Vinson, S.B., Gerling, D., 1989. Courtship and mating behavior of *Bemisia tabaci* (Homoptera: Aleyrodidae). Environ. Entomol. 18, 800–806.

Li, X., Baudry, J., Berenbaum, M.R., Schuler, M.A., 2004. Structural and functional divergence of insect CYP6B proteins: From specialist to generalist cytochrome P450. Proc. Natl. Acad. Sci. 101, 2939–2944.

Li, X., Schuler, M.A., Berenbaum, M.R., 2007. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. Annu. Rev. Entomol. 52, 231–253.

Liang, P., Cui, J.Z., Yang, X.Q., Gao, X.W., 2007. Effects of host plants on insecticide susceptibility and carboxylesterase activity in *Bemisia tabaci* biotype B and greenhouse whitefly, *Trialeurodes vaporariorum*. Pest Manag. Sci. 63, 365–371.

Librado, P., Rozas, J., 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinforma. Oxf. Engl. 25, 1451–1452.

Lima, L.H.C., Návia, D., Inglis, P.W., Oliveira, M.R.V., 2000. Survey of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) biotypes in Brazil using RAPD markers. Genet. Mol. Biol. 23, 781–785.

Lin, K.J., Lu, Y.H., Wan, P., Yang, Y.Z., Wyckhuys, K.A.G., Wu, K.M., 2015. Simultaneous reduction in incidence of *Bemisia tabaci* (Hemiptera: Aleyrodidae) and *Sylepta derogata* (Lepidoptera: Pyralidae) using velvetleaf, *Abutilon theophrasti* as a trap crop. J. Pest Sci. 88, 49–56.

Lindroth, R.L., 1989. Host plant alteration of detoxication activity in *Papilio glaucus glaucus*. Entomol. Exp. Appl. 50, 29–35.

Liu, B., Preisser, E.L., Jiao, X., Pan, H., Xie, W., Wang, S., Wu, Q., Zhang, Y., 2013. Plant-mediated changes in the feeding behavior of an invasive whitefly. Environ. Entomol. 42, 980–986.

Liu, B., Yan, F., Chu, D., Pan, H., Jiao, X., Xie, W., Wu, Q., Wang, S., Xu, B., Zhou, X., Zhang, Y., 2012a. Difference in feeding behaviors of two invasive whiteflies on host plants with different suitability: Implication for competitive displacement. Int. J. Biol. Sci. 8, 697–706.

Liu, S.S., Colvin, J., De Barro, P.J., 2012b. Species concepts as applied to the whitefly *Bemisia tabaci* systematics: How many species are there? J. Integr. Agric. 11, 176–186.

Liu, S.S., De Barro, P.J., Xu, J., Luan, J.B., Zang, L.S., Ruan, Y.M., Wan, F.H., 2007. Asymmetric mating interactions drive widespread invasion and displacement in a whitefly. Science 318, 1769–1772.

Liu, X., Zhang, Y.J., Xie, W., Wu, Q.J., Wang, S.L., 2016. The suitability of biotypes Q and B of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) at different nymphal instars as hosts for *Encarsia formosa* Gahan (Hymenoptera: Aphelinidae). PeerJ 4, e1863.

Lo, N., Casiraghi, M., Salati, E., Bazzocchi, C., Bandi, C., 2002. How many *Wolbachia* supergroups exist? Mol. Biol. Evol. 19, 341–346.

Lo, N., Paraskevopoulos, C., Bourtzis, K., O'Neill, S.L., Werren, J.H., Bordenstein, S.R., Bandi, C., 2007. Taxonomic status of the intracellular bacterium *Wolbachia pipientis*. Int. J. Syst. Evol. Microbiol. 57, 654–657.

Lopez, J.V., Yuhki, N., Masuda, R., Modi, W., O'Brien, S.J., 1994. *Numt*, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J. Mol. Evol. 39, 174–190.

Lowe, S., Browne, M., Boudjelas, S., De Poorter, M., 2004. 100 of the world's worst invasive alien species. A selection form the Global Invasive Species Database. Invasive Species Spec. Group World Conserv. Union 1–12.

Lü, Z.C., Wang, Y.M., Zhu, S.G., Yu, H., Guo, J.Y., Wan, F.H., 2014. Trade-offs between survival, longevity, and reproduction, and variation of survival tolerance in Mediterranean *Bemisia tabaci* after temperature stress. J. Insect Sci. 14.

Luan, J.B., Ruan, Y.M., Zhang, L., Liu, S.S., 2008. Pre-copulation intervals, copulation frequencies, and initial progeny sex ratios in two biotypes of whitefly, *Bemisia tabaci*. Entomol. Exp. Appl. 129, 316–324.

Luan, J.B., Wang, X.W., Colvin, J., Liu, S.S., 2014. Plant-mediated whitefly-begomovirus interactions: Research progress and future prospects. Bull. Entomol. Res. 104, 267–276.

Luan, J.B., Xu, J., Lin, K.K., Zalucki, M.P., Liu, S.S., 2012. Species exclusion between an invasive and an indigenous whitefly on host plants with differential levels of suitability. J. Integr. Agric. 11, 215–224.

Luo, C., Jones, C.M., Devine, G., Zhang, F., Denholm, I., Gorman, K., 2010. Insecticide resistance in *Bemisia tabaci* biotype Q (Hemiptera: Aleyrodidae) from China. Crop Prot. 29, 429–434.

Ma, D., Gorman, K., Devine, G., Luo, W., Denholm, I., 2007. The biotype and insecticide-resistance status of whiteflies, *Bemisia tabaci* (Hemiptera: Aleyrodidae), invading cropping

systems in Xinjiang Uygur Autonomous Region, northwestern China. Crop Prot. 26, 612–617.

Ma, D., Hadjistylli, M., Gorman, K., Denholm, I., Devine, G., 2004. Pre- and post-zygotic breeding incompatibilities between B and Q biotypes of *Bemisia tabaci*, in: Proceedings of the 2nd European Whitefly Symposium. Cavtat, Croatia.

Macfadyen, S., Paull, C., Boykin, L.M., De Barro, P.J., Maruthi, M.N., Otim, M., Kalyebi, A., Vassão, D.G., Sseruwagi, P., Tay, W.T., Delatte, H., Seguni, Z., Colvin, J., Omongo, C.A., 2018. Cassava whitefly, *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) in East African farming landscapes: a review of the factors determining abundance. Bull. Entomol. Res. 1–18.

Maffei, M., Chialva, F., Sacco, T., 1989. Glandular trichomes and essential oils in developing peppermint leaves. New Phytol. 111, 707–716.

Magesh, V., Zhu, Z., Tang, T., Chen, S., Li, L., Wang, L., Varma, K.K., Wu, Y., 2017. Toxicity of neonicotinoids to honey bees and detoxification mechanism in honey bees. IOSR J. Environ. Sci. Toxicol. Food Technol. 11, 102–110.

Maitra, S., Dombrowski, S.M., Basu, M., Raustol, O., Waters, L.C., Ganguly, R., 2000. Factors on the third chromosome affect the level of *Cyp6a2* and *Cyp6a8* expression in *Drosophila melanogaster*. Gene 248, 147–156.

Malka, O., Santos-Garcia, D., Feldmesser, E., Sharon, E., Krause-Sakate, R., Delatte, H., van Brunschot, S., Patel, M., Visendi Muhindira, P., Mugerwa, H., Seal, S., Colvin, J., Morin, S., 2018. Species-complex diversification and host-plant associations in *Bemisia tabaci*: A plant-defense, detoxification perspective revealed by RNAseq analyses. Mol. Ecol. 27, 4241–4256.

Malka, O., Shekhov, A., Reichelt, M., Gershenzon, J., Vassão, D.G., Morin, S., 2016. Glucosinolate desulfation by the phloem-feeding insect *Bemisia tabaci*. J. Chem. Ecol. 42, 230–235.

Mallet, J., 2008. Mayr's view of Darwin: Was Darwin wrong about speciation? Biol. J. Linn. Soc. 95, 3–16.

Mallet, J., 2007. Species, concepts of, in: Levin, S.A. (Ed.), Encyclopedia of biodiversity. Academic Press pp. 1–15.

Mann, E., Stouthamer, C.M., Kelly, S.E., Dzieciol, M., Hunter, M.S., Schmitz-Esser, S., 2017. Transcriptome sequencing reveals novel candidate genes for *Cardinium hertigii*-caused cytoplasmic incompatibility and host-cell interaction. mSystems 2, e00141-17.

Mansoor, S., Khan, S.H., Bashir, A., Saeed, M., Zafar, Y., Malik, K.A., Briddon, R., Stanley, J., Markham, P.G., 1999. Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. Virology 259, 190–199.

Manyong, V.M., Maeda, C., Kanju, E., Legg, J.P., 2012. Economic damage of cassava brown streak disease in sub-Saharan Africa, in: Tropical Root and Tuber Crops and the Challenges of Globalization and Climate Change. Presented at the 11th ISTRC-AB Symposium, Kinshasa, Democratic Republic of Congo, pp. 61–68.

Mao, W., Rupasinghe, S., Zangerl, A.R., Schuler, M.A., Berenbaum, M.R., 2006. Remarkable substrate-specificity of CYP6AB3 in *Depressaria pastinacella*, a highly specialized caterpillar. Insect Mol. Biol. 15, 169–179.

Markham, P.G., Bedford, I.D., Liu, S., Pinner, M.S., 1994. The transmission of geminiviruses by *Bemisia tabaci*. Pestic. Sci. 42, 123–128.

Martin, J.H., Mound, L.A., 2007. An annotated check list of the world's whiteflies (Insecta: Hemiptera: Aleyrodidae). Zootaxa 1492, 1–84.

Martinez-Carrillo, J.L., Brown, J.K., 2007. Note: First report of the Q biotype of *Bemisia tabaci* in Southern Sonora, Mexico. Phytoparasitica 35, 282.

Marubayashi, J.M., Kliot, A., Yuki, V.A., Rezende, J.A.M., Krause-Sakate, R., Pavan, M.A., Ghanim, M., 2014. Diversity and localization of bacterial endosymbionts from whitefly species collected in Brazil. PLoS One 9, e108363.

Maruthi, M.N., Colvin, J., Seal, S.E., 2001. Mating compatibility, life-history traits, and RAPD-PCR variation in *Bemisia tabaci* associated with the cassava mosaic disease pandemic in East Africa. Entomol. Exp. Appl. 99, 13–23.

Maruthi, M.N., Colvin, J., Seal, S.E., Gibson, G., Cooper, J., 2002a. Co-adaptation between cassava mosaic geminiviruses and their local vector populations. Virus Res. 86, 71–85.

Maruthi, M.N., Colvin, J., Seal, S.E., Thresh, J.M., 2002b. First report of a distinct begomovirus infecting cassava from Zanzibar. Plant Dis. 86, 187–187.

Maxwell, F.G., Jennings, P.R., 1980. Breeding plants resistant to insects. John Wiley & Sons Inc, New York.

Mayhew, P.J., 1997. Adaptive patterns of host-plant selection by phytophagous insects. Oikos 79, 417–428.

Mayr, E., 2002. Species, in: Brenner, S., Miller, J.H. (Eds.), Encyclopedia of Genetics. Academic, San Diego, California, USA, pp. 1684–1689.

Mayr, E., 1969. Principles of systematic zoology. McGraw-Hill, New York.

Mayr, E., 1942. Systematics and the origin of species, from the viewpoint of a zoologist. Harvard University Press.

Mcauslane, H.J., 1996. Influence of leaf pubescence on ovipositional preference of *Bemisia argentifolii* (Homoptera: Aleyrodidae) on soybean. Environ. Entomol. 25, 834–841.

McCoy, C.W., Samson, R.A., Boucias, D.G., 1988. Entomogenous fungi, in: Ignoffo, C.M. (Ed.), CRC handbook of natural pesticides, Vol. V.: Microbial insecticides, Part A: Entomogenous protozoa and fungi. CRC Press Inc, Boca Raton, Florida, pp. 151–236.

McInerney, P., Adams, P., Hadi, M.Z., 2014. Error rate comparison during polymerase chain reaction by DNA polymerase. Mol. Biol. Int. 2014, 1–8.

McKenzie, C.L., Bethke, J.A., Byrne, F.J., Chamberlin, J.R., Dennehy, T.J., Dickey, A.M., Gilrein, D., Hall, P.M., Ludwig, S., Oetting, R.D., Osborne, L.S., Schmale, L., Shatters, R.G., 2012. Distribution of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes in North America after the Q invasion. J. Econ. Entomol. 105, 753–766.

McKenzie, C.L., Hodges, G., Osborne, L.S., Byrne, F.J., Shatters, R.G., 2009. Distribution of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes in Florida-investigating the Q invasion. J. Econ. Entomol. 102, 670–676.

Meekes, E.T.M., van Voorst, S., Joosten, N.N., Fransen, J.J., van Lenteren, J.C., 2000. Persistence of the fungal whitefly pathogen, *Aschersonia aleyrodis*, on three different plant species. Mycol. Res. 104, 1234–1240.

Meier, R., Shiyang, K., Vaidya, G., Ng, P.K.L., 2006. DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. Syst. Biol. 55, 715–728.

Metzker, M.L., 2010. Sequencing technologies - the next generation. Nat. Rev. Genet. 11, 31–46.

Meyer, C.P., Paulay, G., 2005. DNA barcoding: Error rates based on comprehensive sampling. PLoS Biol. 3, e422.

Miller, W., Drautz, D.I., Janecka, J.E., Lesk, A.M., Ratan, A., Tomsho, L.P., Packard, M., Zhang, Y., McClellan, L.R., Qi, J., Zhao, F.P., Gilbert, M.T.P., Dalén, L., Arsuaga, J.L., Ericson, P.G.P., Huson, D.H., Helgen, K.M., Murphy, W.J., Götherström, A., Schuster, S.C., 2009. The mitochondrial genome sequence of the Tasmanian tiger (*Thylacinus cynocephalus*). Genome Res. 19, 213–220.

Miller, W.B., Peralta, E., Ellis, D.R., Perkins, H.H., 1994. Stickiness potential of individual insect honeydew carbohydrates on cotton lint. Text. Res. J. 64, 344–350.

Minh, B.Q., Nguyen, M.A.T., von Haeseler, A., 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30, 1188–1195.

Mohammed, I.U., Ghosh, S., Maruthi, M.N., 2017. Generating virus-free cassava plants by in vitro propagation with chemical and heat treatment. Afr. J. Biotechnol. 16, 1551–1560.

Mohanty, A.K., Basu, A.N., 1986. Effect of host plants and seasonal factors on intraspecific variations in pupal morphology of the whitefly vector, *Bemisia tabaci* (Genn.) (Homoptera: Aleyrodidae). J. Entomol. Res. 10, 19–26.

Monci, F., Navas-Castillo, J., Cenis, J.L., Lacasa, A., Benazoun, A., Moriones, E., 2000. Spread of *Tomato yellow leaf curl virus* Sar from the Mediterranean Basin: Presence in the Canary Islands and Morocco. Plant Dis. 84, 490.

Mor, U., 1987. *Bemisia tabaci* and cotton physiology: A 5-year summary of the influence of water-stressed plants on the pest population. Phytoparasitica 15, 261.

Morales, F.J., Anderson, P.K., 2001. The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America. Arch. Virol. 146, 415–441.

Morey, M., Fernández-Marmiesse, A., Castiñeiras, D., Fraga, J.M., Couce, M.L., Cocho, J.A., 2013. A glimpse into past, present, and future DNA sequencing. Mol. Genet. Metab., Special Issue: Diagnosis 110, 3–24.

Morrison, D.A., 2011. Review of species: A history of the idea. Syst. Biol. 60, 239–241.

Morrow, J.L., Hall, A.A.G., Riegler, M., 2017. Symbionts in waiting: The dynamics of incipient endosymbiont complementation and replacement in minimal bacterial communities of psyllids. Microbiome 5, 58.

Mound, L.A., 1965. Effect of whitefly (*Bemisia tabaci*) on cotton in the Sudan Gezira. Emp. Cotton Grow. Rev. 42, 290–294.

Mound, L.A., 1963. Host-correlated variation in *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae). Proc. R. Entomol. Soc. Lond. Ser. Gen. Entomol. 38, 171–180.

Mound, L.A., Halsey, S.H., 1978. Whitefly of the world. British Museum (natural History), London.

Mount, D.W., 2008. Maximum parsimony method for phylogenetic prediction. Cold Spring Harb. Protoc. 2008, pdb.top32.

Mouton, L., Gnankiné, O., Henri, H., Terraz, G., Ketoh, G., Martin, T., Fleury, F., Vavre, F., 2015. Detection of genetically isolated entities within the Mediterranean species of *Bemisia tabaci*: New insights into the systematics of this worldwide pest. Pest Manag. Sci. 71, 452–458.

Moya, A., Guirao, P., Cifuentes, D., Beitia, F., Cenis, J.L., 2001. Genetic diversity of Iberian populations of *Bemisia tabaci* (Hemiptera: Aleyrodidae) based on random amplified polymorphic DNA-polymerase chain reaction. Mol. Ecol. 10, 891–897.

Mugerwa, H., 2018. Molecular characterisation of cassava-colonising populations of the whitefly, *Bemisia tabaci* (Dissertation). University of Greenwich.

Mugerwa, H., Seal, S.E., Wang, H.L., Patel, M.V., Kabaalu, R., Omongo, C.A., Alicai, T., Tairo, F., Ndunguru, J., Sseruwagi, P., Colvin, J., 2018. African ancestry of New World, *Bemisia tabaci*-whitefly species. Sci. Rep. 8, 2734.

Muhire, B.M., Varsani, A., Martin, D.P., 2014. SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9, e108277.

Muigai, S.G., Bassett, M.J., Schuster, D.J., Scott, J.W., 2003. Greenhouse and field screening of wild *Lycopersicon* germplasm for resistance to the whitefly *Bemisia argentifolii*. Phytoparasitica 31, 27.

Muigai, S.G., Schuster, D.J., Snyder, J.C., Scott, J.W., Bassett, M.J., McAuslane, H.J., 2002. Mechanisms of resistance in *Lycopersicon* germplasm to the whitefly *Bemisia argentifolii*. Phytoparasitica 30, 347.

Muñiz, M., 2000. Host suitability of two biotypes of *Bemisia tabaci* on some common weeds. Entomol. Exp. Appl. 95, 63–70.

Muñiz, M., Nombela, G., 2001. Differential variation in development of the B- and Q-biotypes of *Bemisia tabaci* (Homoptera: Aleyrodidae) on sweet pepper at constant temperatures. Environ. Entomol. 30, 720–727.

Muñiz, M., Nombela, G., 1997. Development, oviposition and female longevity of two biotypes of *Bemisia tabaci* (Homioptera: Aleyrodidae) on three varieties of *Capsicum annuum* L. IOBCWPRS Bull. 20, 143–146.

Murphy, S.M., 2004. Enemy-free space maintains swallowtail butterfly host shift. Proc. Natl. Acad. Sci. 101, 18048–18052.

Nakhla, M.K., Maxwell, D.P., Martinez, R.T., Carvalho, M.G., Gilbertson, R.L., 1994. Widespread occurrence of the eastern Mediterranean strain of tomato yellow leaf curl geminivirus in tomatoes in the Dominican Republic. Plant Dis. 78, 926.

Narang, S.K., Seawright, J.A., Suarez, M.F., 1991. Genetic structure of natural populations of *Anopheles albimanus* in Colombia. J. Am. Mosq. Control Assoc. 7, 437–445.

Naranjo, S.E., 2001. Conservation and evaluation of natural enemies in IPM systems for *Bemisia tabaci*. Crop Prot., Challenges and opportunities for pest management of *Bemisia tabaci* in the new century 20, 835–852.

Naranjo, S.E., Ellsworth, P.C., 2009. The contribution of conservation biological control to integrated control of *Bemisia tabaci* in cotton. Biol. Control 51, 458–470.

Naranjo, S.E., Ellsworth, P.C., Chu, C.C., Henneberry, T.J., Riley, D.G., Watson, T.F., Nichols, R.L., 1998. Action thresholds for the management of *Bemisia tabaci* (Homoptera: Aleyrodidae) in cotton. J. Econ. Entomol. 91, 1415–1426.

Nauen, R., Denholm, I., 2005. Resistance of insect pests to neonicotinoid insecticides: current status and future prospects. Arch. Insect Biochem. Physiol. 58, 200–215.

Nauen, R., Stumpf, N., Elbert, A., 2002. Toxicological and mechanistic studies on neonicotinoid cross resistance in Q-type *Bemisia tabaci* (Hemiptera: Aleyrodidae). Pest Manag. Sci. 58, 868–875.

Nava-Camberos, U., Riley, D.G., Harris, M.K., 2001. Temperature and host plant effects on development, survival, and fecundity of *Bemisia argentifolii* (Homoptera: Aleyrodidae). Environ. Entomol. 30, 55–63.

Navas-Castillo, J., Camero, R., Bueno, M., Moriones, E., 2000. Severe yellowing outbreaks in tomato in Spain associated with infections of *Tomato chlorosis virus*. Plant Dis. 84, 835–837.

Navas-Castillo, J., Fiallo-Olivé, E., Sánchez-Campos, S., 2011. Emerging virus diseases transmitted by whiteflies. Annu. Rev. Phytopathol. 49, 219–248.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.

Nirgianaki, A., Banks, G.K., Frohlich, D.R., Veneti, Z., Braig, H.R., Miller, T.A., Bedford, I.D., Markham, P.G., Savakis, C., Bourtzis, K., 2003. *Wolbachia* infections of the whitefly *Bemisia tabaci*. Curr. Microbiol. 47, 93–101.

Nombela, G., Beitia, F., Muñiz, M., 2001. A differential interaction study of *Bemisia tabaci* Q-biotype on commercial tomato varieties with or without the *Mi* resistance gene, and comparative host responses with the B-biotype. Entomol. Exp. Appl. 98, 339–344.

Nombela, G., Williamson, V.M., Muñiz, M., 2003. The root-knot nematode resistance gene *Mi-1.2* of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. Mol. Plant-Microbe Interact. MPMI 16, 645–649.

Nováková, E., Moran, N.A., 2012. Diversification of genes for carotenoid biosynthesis in aphids following an ancient transfer from a fungus. Mol. Biol. Evol. 29, 313–323.

Nuessly, L.R., Meyerdirk, D.E., Coudriet, D.L., Henneberry, T.J., 1994. The effect of short season cotton production schedules on *Bemisia tabaci* (Gennadius). Southwest. Entomol. 209–217.

Ogden, T.H., Rosenberg, M.S., 2006. Multiple sequence alignment accuracy and phylogenetic inference. Syst. Biol. 55, 314–328.

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2018. Vegan: Community Ecology Package, R package version 2.5-2.

Oliveira, M.R.V., Henneberry, T.J., Anderson, P., 2001. History, current status, and collaborative research projects for *Bemisia tabaci*. Crop Prot., Challenges and opportunities for pest management of *Bemisia tabaci* in the new century 20, 709–723.

Omondi, B.A., Sseruwagi, P., Obeng-Ofori, D., Danquah, E.Y., Kyerematen, R.A., 2005. Mating interactions between okra and cassava biotypes of *Bemisia tabaci* (Homoptera: Aleyrodidae) on eggplant. Int. J. Trop. Insect Sci. 25, 159–167.

Omongo, C.A., Kawuki, R., Bellotti, A.C., Alicai, T., Baguma, Y., Maruthi, M.N., Bua, A., Colvin, J., 2012. African cassava whitefly, *Bemisia tabaci*, resistance in African and South American cassava genotypes. J. Integr. Agric. 11, 327–336.

Oriani, M.A.G., Vendramim, J.D., 2010. Influence of trichomes on attractiveness and ovipositional preference of *Bemisia tabaci* (Genn.) B biotype (Hemiptera: Aleyrodidae) on tomato genotypes. Neotrop. Entomol. 39, 1002–1007.

Otim-Nape, G.W., Thresh, J.M., 1998. The current pandemic of Cassava mosaic virus disease in Uganda, in: Cooke, B.M., Jones, D.G., Kaye, B. (Eds.), The epidemiology of plant diseases. Springer, Dordrecht, pp. 423–443.

Page, R.D.M., Holmes, E.C., 1998. Molecular evolution: A phylogenetic approach, 1st edition. Wiley-Blackwell, Oxford ; Malden, MA.

Pan, H.P., Chu, D., Ge, D.Q., Wang, S.L., Wu, Q.J., Xie, W., Jiao, X.G., Liu, B.M., Yang, X., Yang, N., Su, Q., Xu, B.Y., Zhang, Y.J., 2011. Further spread of and domination by *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotype Q on field crops in China. J. Econ. Entomol. 104, 978–985.

Pan, H.P., Chu, D., Liu, B.M., Xie, W., Wang, S.L., Wu, Q.J., Xu, B.Y., Zhang, Y.J., 2013. Relative amount of symbionts in insect hosts changes with host-plant adaptation and insecticide resistance. Environ. Entomol. 42, 74–78.

Pangnakorn, U., Chuenchooklin, S., 2016. Acute and chronic effect of biopesticide on infestation of whitefly *Bemisia tabaci* (Gennadius) on the culantro cultivation. Int. J. Agric. Biosyst. Eng. 10, 118–122.

Parrella, G., Nappo, A.G., Manco, E., Greco, B., Giorgini, M., 2014. Invasion of the Q2 mitochondrial variant of Mediterranean *Bemisia tabaci* in southern Italy: Possible role of bacterial endosymbionts. Pest Manag. Sci. 70, 1514–1523.

Parrella, G., Scassillo, L., Giorgini, M., 2012. Evidence for a new genetic variant in the *Bemisia tabaci* species complex and the prevalence of the biotype Q in southern Italy. J. Pest Sci. 85, 227–238.

Parsa, S., Medina, C., Rodríguez, V., 2015. Sources of pest resistance in cassava. Crop Prot. 68, 79–84.

Pascual, S., 2006. Short communication. Mechanisms in competition, under laboratory conditions, between Spanish biotypes B and Q of *Bemisia tabaci* (Gennadius). Span. J. Agric. Res. 4, 351–354.

Pascual, S., Callejas, C., 2004. Intra- and interspecific competition between biotypes B and Q of *Bemisia tabaci* (Hemiptera: Aleyrodidae) from Spain. Bull. Entomol. Res. 94, 369–375.

Patil, B.L., Legg, J.P., Kanju, E., Fauquet, C.M., 2015. Cassava brown streak disease: a threat to food security in Africa. J. Gen. Virol. 96, 956–968.

Paulson, G.S., Beardsley, J.W., 1985. Whitefly (Hemiptera: Aleyrodidae) egg pedicel insertion into host plant stomata. Ann. Entomol. Soc. Am. 78, 506–508.

Pennisi, E., 2010. Armed and dangerous. Science 327, 804–805.

Péréfarres, F., Thierry, M., Becker, N., Lefeuvre, P., Reynaud, B., Delatte, H., Lett, J.M., 2012. Biological invasions of geminiviruses: case study of TYLCV and *Bemisia tabaci* in Reunion Island. Viruses 4, 3665–3688.

Perlman, S.J., Hunter, M.S., Zchori-Fein, E., 2006. The emerging diversity of *Rickettsia*. Proc. Biol. Sci. 273, 2097–2106.

Perring, T.M., Cooper, A., Kazmer, D.J., 1992. Identification of the poinsettia strain of *Bemisia tabaci* (Homoptera: Aleyrodidae) on broccoli by electrophoresis. J. Econ. Entomol. 85, 1278–1284.

Perring, T.M., Cooper, A.D., Rodriguez, R.J., Farrar, C.A., Bellows, T.S., 1993. Identification of a whitefly species by genomic and behavioral studies. Science 259, 74–77.

Perring, T.M., Farrar, C.A., Cooper, A.D., 1994. Mating behaviour and competitive displacement in whiteflies, in: Henneberry, T.J. (Ed.), Silverleaf whitefly (formerly sweetpotato whitefly, strain B): 1994 supplement to the five-year national research and action plan. US Department of Agriculture, p. 25.

Perring, T.M., Symmes, E.J., 2006. Courtship behavior of *Bemisia argentifolii* (Hemiptera: Aleyrodidae) and whitefly mate recognition. Ann. Entomol. Soc. Am. 99, 598–606.

Phillips, J.R., Graves, J.B., Luttrell, R.G., 1989. Insecticide resistance management: Relationship to integrated pest management. Pestic. Sci. 27, 459–464.

Picó, B., Díez, M.J., Nuez, F., 1996. Viral diseases causing the greatest economic losses to the tomato crop. II. The tomato yellow leaf curl virus — a review. Sci. Hortic. 67, 151–196.

Poinar, G.O., 1990. Taxonomy and biology of Steinernematidae and Heterorhabditidae. Entomopathog. Nematodes Biol. Control 23–61.

Polston, J.E., McGovern, R.J., Brown, L.G., 1999. Introduction of tomato yellow leaf curl virus in Florida and implications for the spread of this and other geminiviruses of tomato. Plant Dis. 83, 984–988.

Powell Abel, P., Nelson, R.S., De, B., Hoffmann, N., Rogers, S.G., Fraley, R.T., Beachy, R.N., 1986. Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. Science 232, 738–743.

Powell, D.A., Bellows, T.S., 1992. Adult longevity, fertility and population growth rates for *Bemisia tabaci* (Genn.) (Hom., Aleyrodidae) on two host plant species. J. Appl. Entomol. 113, 68–78.

Powell, M.E., Cuthbertson, A.G.S., Bell, H.A., Boonham, N., Morris, J., Northing, P., 2012. First record of the Q Biotype of the sweetpotato whitefly, *Bemisia tabaci*, intercepted in the UK. Eur. J. Plant Pathol. 133, 797–801.

Pruthi, H.S., Samuel, C.K., 1942. Entomological investigations of the leaf curl disease of tobacco in Northern India. V. Biology and population of the whitefly vector, *Bemisia tabaci* (Genn.) in relation to the incidence of the disease. Indian J. Agric. Sci. 12, 37–57.

Qin, L., Pan, L.L., Liu, S.S., 2016. Further insight into reproductive incompatibility between putative cryptic species of the *Bemisia tabaci* whitefly complex. Insect Sci. 23, 215–224.

Qiu, B.L., Mandour, N.S., Xu, C.X., Ren, S.X., 2008. Evaluation of the entomopathogenic nematode *Steinernema feltiae* as a biological control agent of the whitefly, *Bemisia tabaci*. Int. J. Pest Manag. 54, 247–253.

Quaintance, A.L., 1900. Contribution towards a monograph of the American Aleurodidae. Tech. Ser. U. S. Dep. Agric. Bur. Entomol. 8, 9–64.

Quaintance, A.L., Baker, A.C., 1914. Classification of the Aleyrodidae part II. Tech. Ser. U. S. Dep. Agric. Bur. Entomol. 27, 95–109.

Queiroz, J.M., Oliveira, P.S., 2001. Tending ants protect honeydew-producing whiteflies (Homoptera: Aleyrodidae). Environ. Entomol. 30, 295–297.

Quintela, E.D., Abreu, A.G., dos S. Lima, J.F., Mascarin, G.M., dos Santos, J.B., Brown, J.K., 2016. Reproduction of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) B biotype in maize fields (Zea mays L.) in Brazil. Pest Manag. Sci. 72, 2181–2187.

R Core Team, 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rahman, M.U., Khan, A.Q., Rahmat, Z., Iqbal, M.A., Zafar, Y., 2017. Genetics and genomics of cotton leaf curl disease, its viral causal agents and whitefly vector: A way forward to sustain cotton fiber security. Front. Plant Sci. 8, 1157.

Raina, H.S., Rawal, V., Singh, S., Daimei, G., Shakarad, M., Rajagopal, R., 2015. Elimination of *Arsenophonus* and decrease in the bacterial symbionts diversity by antibiotic treatment leads to increase in fitness of whitefly, *Bemisia tabaci*. Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis. 32, 224–230.

Rana, V.S., Singh, S.T., Priya, N.G., Kumar, J., Rajagopal, R., 2012. *Arsenophonus* GroEL interacts with CLCuV and is localized in midgut and salivary gland of whitefly *B. tabaci*. PLoS One 7, e42168.

Rannala, B., Yang, Z., 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. J. Mol. Evol. 43, 304–311.

Rao, Q., Luo, C., Zhang, H., Guo, X., Devine, G.J., 2011. Distribution and dynamics of *Bemisia tabaci* invasive biotypes in central China. Bull. Entomol. Res. 101, 81–88.

Ratnasingham, S., Hebert, P.D.N., 2007. BOLD: The barcode of life data system (http://www.barcodinglife.org). Mol. Ecol. Notes 7, 355–364.

Ratte, H.T., 1985. Temperature and insect development, in: Hoffmann, K.H. (Ed.), Environmental physiology and biochemistry of insects. Springer-Verlag, Berlin, Heidelberg, pp. 33–66.

Rauch, N., Nauen, R., 2003. Identification of biochemical markers linked to neonicotinoid cross resistance in *Bemisia tabaci* (Hemiptera: Aleyrodidae). Arch. Insect Biochem. Physiol. 54, 165–176.

Régnière, J., Powell, J., Bentz, B., Nealis, V., 2012. Effects of temperature on development, survival and reproduction of insects: Experimental design, data analysis and modeling. J. Insect Physiol. 58, 634–647.

Reitz, S.R., Trumble, J.T., 2002. Competitive displacement among insects and arachnids. Annu. Rev. Entomol. 47, 435–465.

Ren, S.X., Wang, Z.Z., Qiu, B.L., Xiao, Y., 2001. The pest status of *Bemisia tabaci* in China and non-chemical control strategies. Insect Sci. 8, 279–288.

Roberts, P.A., Thomason, I.J., 1986. Variability in reproduction of isolates of *Meloidogyne incognita* and *M. javanica* on resistant tomato genotypes. Plant Dis. 70, 547–551.

Roditakis, E., Grispou, M., Morou, E., Kristoffersen, J.B., Roditakis, N., Nauen, R., Vontas, J., Tsagkarakou, A., 2009. Current status of insecticide resistance in Q biotype *Bemisia tabaci* populations from Crete. Pest Manag. Sci. 65, 313–322.

Rodríguez-López, M.J., Garzo, E., Bonani, J.P., Fereres, A., Fernández-Muñoz, R., Moriones, E., 2011. Whitefly resistance traits derived from the wild tomato *Solanum pimpinellifolium* affect the preference and feeding behavior of *Bemisia tabaci* and reduce the spread of *Tomato yellow leaf curl virus*. Phytopathology 101, 1191–1201.

Rogan, D.T., 2012. Reproductive competition of *Bemisia tabaci* biotypes on cotton (Dissertation). The University of Arizona, Tucson, Arizona.

Ronda, M., Adán, A., Beitia, D.F., Cifuentes, D., Cenis, J.L., 2000. Interbreeding between biotypes of *Bemisia tabaci*. Eur. Whitefly Stud. Netw. Newsl. 3.

Ros, V.I.D., Fleming, V.M., Feil, E.J., Breeuwer, J.A.J., 2009. How diverse is the genus *Wolbachia*? Multiple-gene sequencing reveals a putatively new *Wolbachia* supergroup

recovered from spider mites (Acari: Tetranychidae). Appl. Environ. Microbiol. 75, 1036–1043.

Rosell, R.C., Bedford, I.D., Frohlich, D.R., Gill, R.J., Brown, J.K., Markham, P.G., 1997. Analysis of morphological variation in distinct populations of *Bemisia tabaci* (Homoptera: Aleyrodidae). Ann. Entomol. Soc. Am. 90, 575–589.

Rossi, M., Goggin, F.L., Milligan, S.B., Kaloshian, I., Ullman, D.E., Williamson, V.M., 1998. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. Proc. Natl. Acad. Sci. 95, 9750–9754.

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. Clifton NJ 132, 365–386.

Ruan, Y.M., Luan, J.B., Zang, L.S., Liu, S.S., 2007. Observing and recording copulation events of whiteflies on plants using a video camera. Entomol. Exp. Appl. 124, 229–233.

Russell, L.M., 1957. Synonyms of *Bemisia tabaci* (Gennadius) (Homoptera, Aleyrodidae). Bull. Brooklyn Entomol. Soc. 52, 122–123.

Saleh, D., Laarif, A., Clouet, C., Gauthier, N., 2012. Spatial and host-plant partitioning between coexisting *Bemisia tabaci* cryptic species in Tunisia. Popul. Ecol. 54, 261–274.

Santos-Garcia, D., Farnier, P.A., Beitia, F., Zchori-Fein, E., Vavre, F., Mouton, L., Moya, A., Latorre, A., Silva, F.J., 2012. Complete genome sequence of "*Candidatus* Portiera aleyrodidarum" BT-QVLC, an obligate symbiont that supplies amino acids and carotenoids to *Bemisia tabaci*. J. Bacteriol. 194, 6654–6655.

Santos-Garcia, D., Silva, F.J., Moya, A., Latorre, A., 2014. No exception to the rule: *Candidatus* Portiera aleyrodidarum cell wall revisited. FEMS Microbiol. Lett. 360, 132–136.

Santos-Garcia, D., Vargas-Chavez, C., Moya, A., Latorre, A., Silva, F.J., 2015. Genome evolution in the primary endosymbiont of whiteflies sheds light on their divergence. Genome Biol. Evol. 7, 873–888.

Scheirs, J., 2002. Integrating optimal foraging and optimal oviposition theory in plant–insect research. Oikos 96, 187–191.

Scheirs, J., De Bruyn, L., Verhagen, R., 2000. Optimization of adult performance determines host choice in a grass miner. Proc. R. Soc. B Biol. Sci. 267, 2065–2069.

Schuler, M.A., 2011. P450s in plant-insect interactions. Biochim. Biophys. Acta 1814, 36–45.

Schuster, D.J., 2004. Squash as a trap crop to protect tomato from whitefly-vectored tomato yellow leaf curl. Int. J. Pest Manag. 50, 281–284.

Schuster, D.J., Funderburk, J.E., Stansly, P.A., 1996. IPM in tomatoes, in: Rosen, A., Bennett, F.D., Capinera, J.L. (Eds.), Pest management in the subtropics: Integrated pest management - A Florida perspective. Intercept, Andover, UK, pp. 387–411.

Schuster, D.J., Mueller, T.F., Kring, J.B., Price, J.F., 1990. Relationship of the sweetpotato whitefly to a new tomato fruit disorder in Florida. HortScience 25, 1618–1620.

Scott, I.A.W., Workman, P.J., Drayton, G.M., Burnip, G.M., 2007. First record of *Bemisia tabaci* biotype Q in New Zealand. N. Z. Plant Prot. 60, 264–270.

Seal, S.E., vandenBosch, F., Jeger, M.J., 2006. Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control. Crit. Rev. Plant Sci. 25, 23–46.

Secker, A.E., Bedford, I.D., Markham, P.G., De Williams, M.E.C., 1998. Squash, a reliable field indicator for the presence of the B biotype of tobacco whitefly, *Bemisia tabaci*. Presented at the Brighton Crop Protection Conference: Pests and Diseases, Brighton, UK.

Shah, M.M.R., Liu, T.X., 2013. Feeding experience of *Bemisia tabaci* (Hemiptera: Aleyrodidae) affects their performance on different host plants. PLoS One 8, e77368.

Shah, M.M.R., Zhang, S.Z., Liu, T.X., 2015. Whitefly, host plant and parasitoid: A review on their interactions. Asian J. Appl. Sci. Eng. 4, 47–60.

Sharaf, N., 1986. Chemical control of *Bemisia tabaci*, in: Proceedings of a Symposium on *Bemisia tabaci* - Ecology and Control XVIIth International Congress of Entomology. pp. 111–127.

Sharma, R.K., Gupta, V.K., Jindal, J., Dilawari, V.K., 2008. Host associated genetic variations in whitefly, *Bemisia tabaci* (Genn.). IJBT Vol73 July 2008.

Shi, X.B., Pan, H.P., Zhang, H.G., Jiao, X.G., Xie, W., Wu, Q.J., Wang, S.L., Fang, Y., Chen, G., Zhou, X.G., Zhang, Y.J., 2014. *Bemisia tabaci* Q carrying *tomato yellow leaf curl virus* strongly suppresses host plant defenses. Sci. Rep. 4.

Simmons, A.M., Harrison, H.F., Ling, K.S., 2008. Forty-nine new host plant species for *Bemisia tabaci* (Hemiptera: Aleyrodidae). Entomol. Sci. 11, 385–390.

Simón, B., Cenis, J.L., De la Rúa, P., 2007. Distribution patterns of the Q and B biotypes of *Bemisia tabaci* in the Mediterranean Basin based on microsatellite variation. Entomol. Exp. Appl. 124, 327–336.

Sloan, D.B., Moran, N.A., 2013. The evolution of genomic instability in the obligate endosymbionts of whiteflies. Genome Biol. Evol. 5, 783–793.

Sloan, D.B., Moran, N.A., 2012. Endosymbiotic bacteria as a source of carotenoids in whiteflies. Biol. Lett. 8, 986–989.

Smith, M.A., Rodriguez, J.J., Whitfield, J.B., Deans, A.R., Janzen, D.H., Hallwachs, W., Hebert, P.D.N., 2008. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. Proc. Natl. Acad. Sci. 105, 12359–12364.

Song, H., Buhay, J.E., Whiting, M.F., Crandall, K.A., 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc. Natl. Acad. Sci. 105, 13486–13491.

Srivathsan, A., Meier, R., 2012. On the inappropriate use of Kimura-2-parameter (K2P) divergences in the DNA-barcoding literature. Cladistics 28, 190–194.

Sseruwagi, P., Legg, J.P., Maruthi, M.N., Colvin, J., Rey, M.E.C., Brown, J.K., 2005. Genetic diversity of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) populations and presence of the B biotype and a non-B biotype that can induce silverleaf symptoms in squash, in Uganda. Ann. Appl. Biol. 147, 253–265.

Stansly, P.A., Naranjo, S.E., 2010. *Bemisia*: Bionomics and management of a global pest. Springer, Dodrecht, Netherlands.

Steenberg, T., Humber, R.A., 1999. Entomopathogenic potential of *Verticillium* and *Acremonium* species (Deuteromycotina: Hyphomycetes). J. Invertebr. Pathol. 73, 309–314.

Stephens, D.W., Krebs, J.R., 1986. Foraging theory. Princeton University Press.

Steppuhn, A., Gase, K., Krock, B., Halitschke, R., Baldwin, I.T., 2004. Nicotine's defensive function in nature. PLoS Biol. 2, e217.

Stouthamer, R., Breeuwer, J.A., Hurst, G.D., 1999. *Wolbachia pipientis*: Microbial manipulator of arthropod reproduction. Annu. Rev. Microbiol. 53, 71–102.

Su, Q., Oliver, K.M., Pan, H.P., Jiao, X.G., Liu, B.M., Xie, W., Wang, S.L., Wu, Q.J., Xu, B.Y., White, J.A., Zhou, X.G., Zhang, Y.J., 2013a. Facultative symbiont *Hamiltonella* confers benefits to *Bemisia tabaci* (Hemiptera: Aleyrodidae), an invasive agricultural pest worldwide. Environ. Entomol. 42, 1265–1271.

Su, Q., Pan, H.P., Liu, B.M., Chu, D., Xie, W., Wu, Q.J., Wang, S.L., Xu, B.Y., Zhang, Y.J., 2013b. Insect symbiont facilitates vector acquisition, retention, and transmission of plant virus. Sci. Rep. 3, 1367.

Suatoni, E., Vicario, S., Rice, S., Snell, T., Caccone, A., 2006. An analysis of species boundaries and biogeographic patterns in a cryptic species complex: The rotifer-*Brachionus plicatilis*. Mol. Phylogenet. Evol. 41, 86–98.

Sun, D.B., Liu, Y.Q., Qin, L., Xu, J., Li, F.F., Liu, S.S., 2013. Competitive displacement between two invasive whiteflies: Insecticide application and host plant effects. Bull. Entomol. Res. 103, 344–353.

Sun, D.B., Xu, J., Luan, J.B., Liu, S.S., 2011. Reproductive incompatibility between the B and Q biotypes of the whitefly *Bemisia tabaci* in China: Genetic and behavioural evidence. Bull. Entomol. Res. 101, 211–220.

Sunnucks, P., Hales, D.F., 1996. Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Mol. Biol. Evol. 13, 510–524.

Tahiri, A., Halkett, F., Granier, M., Gueguen, G., Peterschmitt, M., 2013. Evidence of gene flow between sympatric populations of the Middle East-Asia Minor 1 and Mediterranean putative species of *Bemisia tabaci*. Ecol. Evol. 3, 2619–2633.

Takahashi, R., 1936. Some Aleyrodidae, Aphididae, Coccidae (Homoptera), and Thysanoptera from Micronesia, Tenthredo. Acta Entomol. 1, 109–120.

Tan, X.L., Hu, N., Zhang, F., Ramirez-Romero, R., Desneux, N., Wang, S., Ge, F., 2016. Mixed release of two parasitoids and a polyphagous ladybird as a potential strategy to control the tobacco whitefly *Bemisia tabaci*. Sci. Rep. 6.

Tang, X.T., Cai, L., Shen, Y., Du, Y.Z., 2018. Diversity and evolution of the endosymbionts of *Bemisia tabaci* in China. PeerJ 6, e5516.

Tay, W.T., Elfekih, S., Court, L.N., Gordon, K.H.J., De Barro, P.J., 2016. Complete mitochondrial DNA genome of *Bemisia tabaci* cryptic pest species complex Asia I (Hemiptera: Aleyrodidae). Mitochondrial DNA Part DNA Mapp. Seq. Anal. 27, 972–973.

Tay, W.T., Elfekih, S., Court, L.N., Gordon, K.H.J., Delatte, H., De Barro, P.J., 2017a. The trouble with MEAM2: Implications of pseudogenes on species delimitation in the globally invasive *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species complex. Genome Biol. Evol. 9, 2732–2738.

Tay, W.T., Elfekih, S., Polaszek, A., Court, L.N., Evans, G.A., Gordon, K.H.J., De Barro, P.J., 2017b. Novel molecular approach to define pest species status and tritrophic interactions from historical *Bemisia* specimens. Sci. Rep. 7, 429.

Tay, W.T., Evans, G.A., Boykin, L.M., De Barro, P.J., 2013. Demonstrating the value of voucher specimens to help resolve modern whitefly taxonomy, in: Proceedings of the 1st International Whitefly Symposium. Kolymbari, Crete, Greece, p. 71.

Tay, W.T., Evans, G.A., Boykin, L.M., De Barro, P.J., 2012. Will the real *Bemisia tabaci* please stand up? PLoS One 7, e50550.

Teeter, K.C., Thibodeau, L.M., Gompert, Z., Buerkle, C.A., Nachman, M.W., Tucker, P.K., 2010. The variable genomic architecture of isolation between hybridizing species of house mice. Evolution 64, 472–485.

Terraz, G., Gueguen, G., Arnó, J., Fleury, F., Mouton, L., 2014. Nuclear and cytoplasmic differentiation among Mediterranean populations of *Bemisia tabaci*: Testing the biological relevance of cytotypes. Pest Manag. Sci. 70, 1503–1513.

Thakur, N., Upadhyay, S.K., Verma, P.C., Chandrashekar, K., Tuli, R., Singh, P.K., 2014. Enhanced whitefly resistance in transgenic tobacco plants expressing double stranded RNA of v-ATPase A gene. PLoS One 9, e87235.

Thalmann, O., Hebler, J., Poinar, H.N., Pääbo, S., Vigilant, L., 2004. Unreliable mtDNA data due to nuclear insertions: A cautionary tale from analysis of humans and other great apes. Mol. Ecol. 13, 321–335.

Thangaraj, S.R., McCulloch, G.A., Subbarayalu, M., Subramaniam, C., Walter, G.H., 2016. Development of microsatellite markers and a preliminary assessment of population structuring in the rice weevil, *Sitophilus oryzae* (L.). J. Stored Prod. Res. 66, 12–17.

Thao, M.L., Baumann, L., Baumann, P., 2004. Organization of the mitochondrial genomes of whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha). BMC Evol. Biol. 4, 25.

Thao, M.L., Baumann, L., Hess, J.M., Falk, B.W., Ng, J.C.K., Gullan, P.J., Baumann, P., 2003. Phylogenetic evidence for two new insect-associated chlamydia of the family *Simkaniaceae*. Curr. Microbiol. 47, 46–50.

Thao, M.L., Baumann, P., 2004. Evolutionary relationships of primary prokaryotic endosymbionts of whiteflies and their hosts. Appl. Environ. Microbiol. 70, 3401–3406.

Thierry, M., Becker, N., Hajri, A., Reynaud, B., Lett, J.M., Delatte, H., 2011. Symbiont diversity and non-random hybridization among indigenous (Ms) and invasive (B) biotypes of *Bemisia tabaci*. Mol. Ecol. 20, 2172–2187.

Thierry, M., Bile, A., Grondin, M., Reynaud, B., Becker, N., Delatte, H., 2015. Mitochondrial, nuclear, and endosymbiotic diversity of two recently introduced populations of the invasive *Bemisia tabaci* MED species in La Réunion. Insect Conserv. Divers. 8, 71–80.

Thompson, W.M.O., 2011. The whitefly, *Bemisia tabaci* (Homoptera: Aleyrodidae) interaction with geminivirus-infected host plants: *Bemisia tabaci*, host plants and geminiviruses. Springer, Dodrecht, Netherlands.

Thresh, J.M., Otim-Nape, G.W., Legg, J.P., Fargette, D., 1997. African cassava mosaic virus disease: The magnitude of the problem. Afr. J. Root Tuber Crops 2, 13–19.

Tindall, K.R., Kunkel, T.A., 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. Biochemistry 27, 6008–6013.

Tomizawa, M., Casida, J.E., 2003. Selective toxicity of neonicotinoids attributable to specificity of insect and mammalian nicotinic receptors. Annu. Rev. Entomol. 48, 339–364.

Toscano, L.C., Boiça Jr., A.L., Maruyama, W.I., 2002. Nonpreference of whitefly for oviposition in tomato genotypes. Sci. Agric. 59, 677–681.

Toscano, N., Castle, S., Henneberry, T., Castle, N., 1998. Persistent silverleaf whitefly exploits desert crop systems. Calif. Agric. 52, 29–33.

Treangen, T.J., Salzberg, S.L., 2011. Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat. Rev. Genet. 13, 36–46.

Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232–W235.

Tsagkarakou, A., Tsigenopoulos, C.S., Gorman, K., Lagnel, J., Bedford, I.D., 2007. Biotype status and genetic polymorphism of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Greece: Mitochondrial DNA and microsatellites. Bull. Entomol. Res. 97, 29–40.

Tsueda, H., Tsuchida, K., 2011. Reproductive differences between Q and B whiteflies, *Bemisia tabaci*, on three host plants and negative interactions in mixed cohorts. Entomol. Exp. Appl. 141, 197–207.

Tukey, J.W., 1949. Comparing individual means in the analysis of variance. Biometrics 5, 99–114.

Turner, R.S., 2005. After the famine: Plant pathology, *Phytophthora infestans*, and the late blight of potatoes, 1845–1960. Hist. Stud. Phys. Biol. Sci. 35, 341–370.

Ueda, S., Brown, J.K., 2006. First report of the Q biotype of *Bemisia tabaci* in Japan by mitochondrial cytochrome oxidase I sequence analysis. Phytoparasitica 34, 405.

Ueda, S., Kitamura, T., Kijima, K., Honda, K.I., Kanmiya, K., 2009. Distribution and molecular characterization of distinct Asian populations of *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Japan. J. Appl. Entomol. 133, 355–366.

Vähä, J.P., Primmer, C.R., 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Mol. Ecol. 15, 63–72.

van der Kuyl, A.C., Kuiken, C.L., Dekker, J.T., Perizonius, W.R.K., Goudsmit, J., 1995. Nuclear counterparts of the cytoplasmic mitochondrial 12S rRNA gene: A problem of ancient DNA and molecular phylogenies. J. Mol. Evol. 40, 652–657.

van Lenteren, J.C., Hua, L.Z., Kamerman, J.W., Rumei, X., 1995. The parasite-host relationship between *Encarsia formosa* (Hym., Aphelinidae) and *Trialeurodes vaporariorum* (Hom., Aleyrodidae) XXVI. Leaf hairs reduce the capacity of *Encarsia* to control greenhouse whitefly on cucumber. J. Appl. Entomol. 119, 553–559.

Vasudeva, V., 2015. GM cotton: Whitefly attack raises anxiety among farmers. The Hindu.

Venables, W.N., Ripley, B.D., 2002. Modern applied statistics with S, 4th edition, Statistics and Computing. Springer-Verlag, New York.

Vilgalys, R., 2003. Taxonomic misidentification in public DNA databases. New Phytol. 160, 4–5.

Villesen, P., 2007. FaBox: An online toolbox for fasta sequences. Mol. Ecol. Notes 7, 965–968.

Vincze, T., Posfai, J., Roberts, R.J., 2003. NEBcutter: A program to cleave DNA with restriction enzymes. Nucleic Acids Res. 31, 3688–3691.

Vosteen, I., Gershenzon, J., Kunert, G., 2016. Enemy-free space promotes maintenance of host races in an aphid species. Oecologia 181, 659–672.

Vyskočilová, S., Tay, W.T., van Brunschot, S., Seal, S., Colvin, J., 2018. An integrative approach to discovering cryptic species within the *Bemisia tabaci* whitefly species complex. Sci. Rep. 8, 10886.

Walker, G.P., Perring, T.M., Freeman, T.P., 2010. Life history, functional anatomy, feeding and mating behaviour, in: Stansly, P.A., Naranjo, S.E. (Eds.), *Bemisia*: Bionomics and management of a global pest. Springer, Dordrecht, pp. 109–160.

Walsh, P.S., Metzger, D.A., Higuchi, R., 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 10, 506–513.

Wan, F.H., Zhang, G.F., Liu, S.S., Luo, C., Chu, D., Zhang, Y.J., Zang, L.S., Jiu, M., Lü, Z.C., Cui, X.H., Zhang, L.P., Zhang, F., Zhang, Q.W., Liu, W.X., Liang, P., Lei, Z.R., Zhang, Y., 2009. Invasive mechanism and management strategy of *Bemisia tabaci* (Gennadius) biotype B: Progress report of 973 Program on invasive alien species in China. Sci. China C Life Sci. 52, 88–95.

Wang, C.Z., 2007. Interpretation of the biological species concept from interspecific hybridization of two *Helicoverpa* species. Chin. Sci. Bull. 52, 284–286.

Wang, H.L., Yang, J., Boykin, L.M., Zhao, Q.Y., Li, Q., Wang, X.W., Liu, S.S., 2013. The characteristics and expression profiles of the mitochondrial genome for the Mediterranean species of the *Bemisia tabaci* complex. BMC Genomics 14, 401.

Wang, J.Y., McCommas, S., Syvanen, M., 1991. Molecular cloning of a glutathione S-transferase overproduced in an insecticide-resistant strain of the housefly (*Musca domestica*). Mol. Gen. Genet. MGG 227, 260–266.

Wang, P., Ruan, Y.M., Liu, S.S., 2010. Crossing experiments and behavioral observations reveal reproductive incompatibility among three putative species of the whitefly *Bemisia tabaci*. Insect Sci. 17, 508–516.

Wang, P., Sun, D.B., Qiu, B.L., Liu, S.S., 2011a. The presence of six cryptic species of the whitefly *Bemisia tabaci* complex in China as revealed by crossing experiments. Insect Sci. 18, 67–77.

Wang, X.W., Luan, B., Li, J.M., Su, Y.L., Xia, J., Liu, S.S., 2011b. Transcriptome analysis and comparison reveal divergence between two invasive whitefly cryptic species. BMC Genomics 12, 458.

Wang, Y., Chen, J., Jiang, L.Y., Qiao, G.X., 2015. Hemipteran mitochondrial genomes: Features, structures and implications for phylogeny. Int. J. Mol. Sci. 16, 12382–12404.

Weeks, A.R., Marec, F., Breeuwer, J.A., 2001. A mite species that consists entirely of haploid females. Science 292, 2479–2482.

Weeks, A.R., Stouthamer, R., 2004. Increased fecundity associated with infection by a cytophaga-like intracellular bacterium in the predatory mite, *Metaseiulus occidentalis*. Proc. Biol. Sci. 271 Suppl 4, S193-195.

Wen, Z., Zeng, R.S., Niu, G., Berenbaum, M.R., Schuler, M.A., 2009. Ecological significance of induction of broad-substrate cytochrome P450s by natural and synthetic inducers in *Helicoverpa zea*. J. Chem. Ecol. 35, 183–189.

Werren, J.H., Baldo, L., Clark, M.E., 2008. *Wolbachia*: Master manipulators of invertebrate biology. Nat. Rev. Microbiol. 6, 741–751.

Wiley, E.O., 1978. The evolutionary species concept reconsidered. Syst. Zool. 27, 17–26.

Wilkins, J.S., 2009. Species: A history of the idea. University of California Press, Los Angeles, California.

Williams, W.G., Kennedy, G.G., Yamamoto, R.T., Thacker, J.D., Bordner, J., 1980. 2-tridecanone: A naturally occurring insecticide from the wild tomato *Lycopersicon hirsutum* f.*glabratum*. Science 207, 888–889.

Wilson, A.C.C., Duncan, R.P., 2015. Signatures of host/symbiont genome coevolution in insect nutritional endosymbioses. Proc. Natl. Acad. Sci. 112, 10255–10261.

Wilson, J.J., 2012. DNA barcodes for insects, in: Kress, W.J., Erickson, D.L. (Eds.), DNA barcodes: Methods and Protocols, methods in molecular biology. Humana Press, New Jersey, pp. 17–46.

Wu, C.I., 2001. The genic view of the process of speciation. J. Evol. Biol. 14, 851–865.

Xia, J., Zhang, C.R., Zhang, S., Li, F.F., Feng, M.G., Wang, X.W., Liu, S.S., 2013. Analysis of whitefly transcriptional responses to *Beauveria bassiana* infection reveals new insights into insect-fungus interactions. PLoS One 8.

Xie, W., Wang, S., Wu, Q., Feng, Y., Pan, H., Jiao, X., Zhou, L., Yang, X., Fu, W., Teng, H., Xu, B., Zhang, Y., 2011. Induction effects of host plants on insecticide susceptibility and detoxification enzymes of *Bemisia tabaci* (Hemiptera: Aleyrodidae). Pest Manag. Sci. 67, 87–93.

Xie, W., Wu, Q.J., Wang, S.L., Jiao, X.G., Guo, L.T., Zhou, X.G., Zhang, Y.J., 2014. Transcriptome analysis of host-associated differentiation in *Bemisia tabaci* (Hemiptera: Aleyrodidae). Front. Physiol. 5.

Xu, C., Qiu, B., Cuthbertson, A.G.S., Zhang, Y., Ren, S., 2012. Adaptability of sweetpotato whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) on seven marginal host plants. Int. J. Pest Manag. 58, 297–301.

Xu, H.X., Hong, Y., Zhang, M.Z., Wang, Y.L., Liu, S.S., Wang, X.W., 2015. Transcriptional responses of invasive and indigenous whiteflies to different host plants reveal their disparate capacity of adaptation. Sci. Rep. 5, 10774.

Xu, J., De Barro, P.J., Liu, S.S., 2010. Reproductive incompatibility among genetic groups of *Bemisia tabaci* supports the proposition that the whitefly is a cryptic species complex. Bull. Entomol. Res. 100, 359–366.

Xu, J., Lin, K.K., Liu, S.S., 2011. Performance on different host plants of an alien and an indigenous *Bemisia tabaci* from China. J. Appl. Entomol. 135, 771–779.

Xu, Q.Y., Chai, F.H., An, X.C., Han, S.C., 2014. Comparison of detoxification enzymes of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes B and Q after various host shifts. Fla. Entomol. 97, 715–723.

Xue, X., Li, S.J., Ahmed, M.Z., De Barro, P.J., Ren, S.X., Qiu, B.L., 2012. Inactivation of *Wolbachia* reveals its biological roles in whitefly host. PLoS One 7, e48148.

Ye, X.D., Su, Y.L., Zhao, Q.Y., Xia, W.Q., Liu, S.S., Wang, X.W., 2014. Transcriptomic analyses reveal the adaptive features and biological differences of guts from two invasive whitefly species. BMC Genomics 15, 370.

Yokomi, R.K., Hoelmer, K.A., Osborne, L.S., 1990. Relationships between the sweetpotato whitefly and the squash silverleaf disorder. Phytopathology 80, 895–900.

Yu, H., Wan, F.H., Guo, J.Y., 2012. Different thermal tolerance and hsp gene expression in invasive and indigenous sibling species of *Bemisia tabaci*. Biol. Invasions 14, 1587–1595.

Zachos, F.E., 2018. (New) Species concepts, species delimitation and the inherent limitations of taxonomy. J. Genet. 97, 811–815.

Zachos, F.E., 2016. Species concepts in biology: Historical development, theoretical foundations and practical relevance. Springer International Publishing.

Zachos, F.E., 2015. Taxonomic inflation, the phylogenetic species concept and lineages in the Tree of Life – a cautionary comment on species splitting. J. Zool. Syst. Evol. Res. 53, 180–184.

Zang, L.S., Chen, W.X., Liu, S.S., 2006. Comparison of performance on different host plants between the B biotype and a non-B biotype of *Bemisia tabaci* from Zhejiang, China. Entomol. Exp. Appl. 121, 221–227.

Zang, L.S., Liu, S.S., 2007. A comparative study on mating behaviour between the B Biotype and a non-B Biotype of *Bemisia tabaci* (Hemiptera: Aleyrodidae) from Zhejiang, China. J. Insect Behav. 20, 157.

Žanić, K., Cenis, J.L., Kačić, S., Katalinić, M., 2005. Current status of *Bemisia tabaci* in coastal Croatia. Phytoparasitica 33, 60.

Zchori-Fein, E., Bourtzis, K., 2011. Manipulative tenants: Bacteria associated with arthropods. CRC Press, Boca Raton, Florida.

Zchori-Fein, E., Brown, J.K., 2002. Diversity of prokaryotes associated with *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae). Ann. Entomol. Soc. Am. 95, 711–718.

Zchori-Fein, E., Gottlieb, Y., Kelly, S.E., Brown, J.K., Wilson, J.M., Karr, T.L., Hunter, M.S., 2001. A newly discovered bacterium associated with parthenogenesis and a change in host selection behavior in parasitoid wasps. Proc. Natl. Acad. Sci. 98, 12555–12560.

Zchori-Fein, E., Perlman, S.J., 2004. Distribution of the bacterial symbiont *Cardinium* in arthropods. Mol. Ecol. 13, 2009–2016.

Zhang, D.X., Hewitt, G.M., 1996. Nuclear integrations: Challenges for mitochondrial DNA markers. Trends Ecol. Evol. 11, 247–251.

Zhang, L.P., Zhang, Y.J., Zhang, W.J., Wu, Q.J., Xu, B.Y., Chu, D., 2005. Analysis of genetic diversity among different geographical populations and determination of biotypes of *Bemisia tabaci* in China. J. Appl. Entomol. 129, 121–128.

Zhang, X.M., Yang, N.W., Wan, F.H., Lövei, G.L., 2014. Density and seasonal dynamics of *Bemisia tabaci* (Gennadius) Mediterranean on common crops and weeds around cotton fields in northern China. J. Integr. Agric. 13, 2211–2220.

Zheng, Y.P., Nitin, M., Wang, L.L., Ruan, Y.M., Jiang, B., Shen, Z.G., Lu, H.F., 2013. Early detection of squash silverleaf and chemical control of *Bemisia tabaci* using imidacloprid. Int. J. Pest Manag. 59, 10–19.

Zubair, M., Zaidi, S.S.A., Shakir, S., Farooq, M., Amin, I., Scheffler, J.A., Scheffler, B.E., Mansoor, S., 2017. Multiple begomoviruses found associated with cotton leaf curl disease in Pakistan in early 1990 are back in cultivated cotton. Sci. Rep. 7, 680.

APPENDICES

Appendix 1: Evoutionary models used in the mitogenome phylogenetic analysis and their characteristics.

Abbreviation	Model name	Features
GTR	General time reversible	Unequal base frequencies, substitution rates AC, AG, AT, CG, CT, GT
HKY	Hasegawa, Kishino and Yano	Unequal base frequencies, substitution rates AC = AT = CG = GT, AG = CT
K3Pu	Kimura 3-parameter unequal	Unequal base frequencies, substitution rates AC = GT, AT = CG, AG = CT
TIM3	Transition model 3	Unequal base frequencies, substitution rates AC = CG, AT = GT, AG, CT
TPM2	Three parameter model 2	Equal base frequencies, substitution rates AC = AT, CG = GT, AG = CT
TPM2u	Three parameter model 2 unequal	Unequal base frequencies, substitution rates AC = AT, CG = GT, AG = CT
TPM3	Three parameter model 3	Equal base frequencies, substitution rates AC = CG, AT = GT, AG = CT
TPM3u	Three parameter model 3 unequal	Unequal base frequencies, substitution rates AC = CG, AT = GT, AG = CT
TVM	Transversion model	Unequal base frequencies, substitution rates AC, AT, CG, GT, AG = CT
	Function name	
+F	Amino acid frequencies	Indicates that amino acid frequencies can be modelled
+I	Invariant sites	Adds a proportion of extent of static, unchanging sites in a dataset
+G(4)	Gamma frequencies	Allows for varying rates across sites according to a discretised gamma distribution (specifies 4 gamma rate categories)

Appendix 2: Results from the sliding window analysis of sequence divergence across the full length of mtCOI gene (1,542 bp) from whiteflies of the Africa/Middle East/Asia Minor clade. The positions or ranges of positions are indicated in parentheses for each minimum and maximum observed nucleotide distance K(JC-total). Q1 consists of two sequences (Spain Q1 and Burkina Faso Q1).

Sequences compared	Minimum distance (window positions)	Maximum distance (window positions)	Average distance
Q1 / Q2	0.00611 (526–1182 … 546–1202)	0.01383 (81–737, … 146–802)	0.01012
ASL / Q1+Q2	0.02476 (781–1437, … 791–1447)	0.04765 (346–1002, 351–1007)	0.03564
ASL / MEAM1	0.04873 (771–1427, … 791–1447)	0.09565 (121–777, 126–782, 211–867)	0.06854
ASL / IO	0.06350 (871–1527)	0.11847 (211–867)	0.08216
Q1+Q2 / MEAM1	0.04765 (776–1432)	0.09392 (186–842, … 206–862)	0.06783
Q1+Q2 / IO	0.06019 (866–1522, 871–1527)	0.09680 (191–847, … 211–867)	0.07377
MEAM1 / IO	0.06682 (876–1532 … 886–1542)	0.11138 (106–762, 191–847 … 201–857)	0.08216

Appendix 3: List of unique published MED haplotypes (Q1, Q2, Q3/Q5, Q3 Croatia and ASL/Q4) and changes in DNA and amino acid sequences identified in them compared to the reference mtCOI from our HTS-derived mitogenomes. Positions in the last two columns refer to the nucleotide positions in which the change occurred. Haplotypes assigned to a particular criterion are colour-coded: criterion (i) INDELs, criterion (iii) outlying number of polymorphisms, criterion (iv) significant amino acid changes and an additional category sequence from a different species. Some haplotypes fall under two criterions. Premature STOP codons (criterion ii) were not detected outside clusters of non-synonymous mutations, which fall under criterion (iii). Key: n.s. = not studied, ins = insertion, del = deletion.

Haplotype	Length (bp)	Accessions	Reference	DNA changes	Amino acid changes
Q1_1	749	AM691051	Chu_Q1		
Q1_2	753	AM691057	Chu_Q1		
Q1_3	758	AM691058	Chu_Q1		
Q1_4	757	AM691080	Chu_Q1		
Q1_5	817	AM691084	Chu_Q1		
Q1_6	817	AM691052 DQ473394 EF694107 EF694109	Chu_Q1 Ahmed_MedBasin1 Chu_Q1 Chu_Q1		
Q1_7	756	AM691068	Chu_Q1		
Q1_8	752	AM691071	Chu_Q1		
Q1_9	756	AM691054	Chu_Q1		
Q1_10	758	EU760734	Gueguen_Q1	751 del A	N.s.
Q1_11	791	EU760746	Gueguen_Q1	751 del A	N.s.
Q1_12	754	AM691056	Chu_Q1		1021 T->M
Q1_14	684	AM691067	Chu_Q1		
Q1_15	768	DQ365874	Ahmed_MedBasin1	1512 del T	N.s.
Q1_16	768	DQ365875 DQ365876	Chu_Q1 Ahmed_MedBasin1	1512 del T	N.s.
Q1_17	838	EU427719	Chu_Q1		709 G->C
Q1_18	807	EU427724	Ahmed_MedBasin1		
Q1_19	776	AF342769	Chu_Q1	1522 del T	N.s.
Q1_20	729	EF398126 EF398114	Ahmed_MedBasin1	1539 ins T	N.s.
Q1_21	803	EU760724	Gueguen_Q1	751 del A	N.s.
Q1_22	803	EU760738	Gueguen_Q1	751 del A	N.s.
Q1_23	805	EU760747	Gueguen_Q1	751 del A	N.s.
Q1_24	676	HM807533	Chu_Q1		
Q1_25	761	DQ365859	Ahmed_MedBasin1	1512 del T	N.s.
Q1_26	809	FJ025793 FJ025794	Ahmed_MedBasin1 Ahmed_MedBAsin1	1518 ins A	N.s.

Q1_27	810	FJ025796	Ahmed_MedBasin1	1434 ins T, 1478 ins C	N.s.
Q1_28	808	FJ025797	Ahmed_MedBasin1	1518 ins A	N.s.
Q1_29	809	FJ025795	Ahmed_MedBasin1	738 ins G, 755 ins T, 765 ins T, 822 ins T, 1518 ins A	N.s.
Q1_30	817	HM590170	Chu_Q1	100% identical to *B. tabaci* Asia I KR020523, JN855568, HM590165 and KF790648 from India; HG918196, HG315654 and HF934996 from Pakistan; KJ778614 from Bangladesh	N.s.
Q1_31	778	AY057179	Chu_Q1	AY057179 = a cassava whitefly from Uganda	N.s.
Q1_32	777	AY057180	Chu_Q1	AY057180 = a cassava whitefly from Uganda 100% identical to KX570785 and AM040604 from Uganda; JQ286450 from Tanzania	N.s.
Q1_33	778	AY057169	Chu_Q1	AY057169 = a cassava whitefly from Uganda	N.s.
Q1_34	783	AY057170	Chu_Q1	802 ins T AY057170 = a cassava whitefly from Uganda	N.s.
Q1_35	779	AY057171	Chu_Q1	AY057171 = a cassava whitefly from Uganda	N.s.
Q1_36	783	AY057163	Chu_Q1	AY057163 = a cassava whitefly from Uganda 100% identical to KX570846, KX570845, AM040603 and AF418669 from Uganda	N.s.

Q1_37	778	AY057162	Chu_Q1	AY057162 = a cassava whitefly from Malawi	N.s.
Q1_38	657	GU086333	Chu_Q1		
Q1_39	781	GQ139499 GQ139501 GQ139502	Chu_Q1 Chu_Q1 Chu_Q1		
Q1_40	781	GQ139500	Chu_Q1		787 L->W
Q1_41	781	GQ139503	Chu_Q1		
Q1_42	721	GQ139504	Chu_Q1		
Q1_43	817	GQ371165	Chu_Q1		1480 K->S
Q1_44	728	EF398122	Chu_Q1		
Q1_45	729	EF398116	Chu_Q1	1539 ins T	N.s.
Q1_46	615	EF398118	Chu_Q1		
Q1_47	720	HM586106	Chu_Q1		
Q1_48	649	GU168791 GU168792 GU168794 GU168795 GU168796 GU168797 GU168798 GU168799 GU168800	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1		
Q1_49	649	GU168793	Chu_Q1		787 L->F
Q1_50	657	GU086329	Chu_Q1		
Q1_51	818	EU263633	Chu_Q1	1464 del C, 1483 del A	N.s.
Q1_52	831	EU263631	Ahmed_MedBasin1		
Q1_53	831	EU263629	Chu_Q1		
Q1_54	844	EU192071	Ahmed_ MedBasin1	1539 ins T	N.s.
Q1_55	842	EU192072	Chu_Q1	1528 del A	N.s.
Q1_56	811	EU263630	Chu_Q1	Last 14 bp cluster of mutations, 1415 del T	N.s.
Q1_57	657	GU086332	Chu_Q1		
Q1_58	807	EU263626	Chu_Q1	Last 22 bp cluster of mutations, 766 del A	N.s.
Q1_59	816	EF667477	Chu_Q1		736 L->F
Q1_60	814	EF566760	Chu_Q1		
Q1_61	818	EF667474	Chu_Q1		
Q1_62	789	EF694104	Chu_Q1		
Q1_63	788	EF694105 EF694106	Chu_Q1 Chu_Q1		

		FJ375358 FJ375350 FJ375354 FJ375355 FJ375351 FJ375348 FJ375353 FJ375352 FJ375347 FJ375357 FJ375356 FJ375346	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1		
Q1_64	759	HM137324	Chu_Q1		
Q1_65	759	HM137360	Chu_Q1		1477 W->C
Q1_66	759	HM137334	Chu_Q1		
Q1_67	859	HM597854	Chu_Q1	First 42 bp cluster of mutations, 763 ins G	N.s.
Q1_68	813	HM597855	Chu_Q1	751 del A	N.s.
Q1_69	836	HM597863	Chu_Q1		
Q1_70	836	HM597849	Chu_Q1	751 del A	N.s.
Q1_71	844	HM597865	Chu_Q1		
Q1_72	833	HM597869	Chu_Q1		
Q1_73	680	FJ594432 FJ594429 FJ594433 FJ594431 FJ594434 FJ594428 FJ594430	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1		
Q1_74	836	HM597870 HM597871	Chu_Q1 Chu_Q1		
Q1_75	854	HM597852	Chu_Q1	First 27 bp cluster of mutations, 738 del A	N.s.
Q1_76	851	HM626150	Chu_Q1	751 del A, 1404 ins T	N.s.
Q1_77	850	HM626151	Chu_Q1	740 del C, 751 del A	N.s.
Q1_78	720	HM137321	Chu_Q1		1477 W->C
Q1_79	856	HM626152	Chu_Q1	First 17 bp cluster of mutations, 751 del A	N.s.
Q1_80	759	HM137331	Chu_Q1		
Q1_81	863	HM626153	Chu_Q1	First 29 bp cluster of mutations	748 A->G 751 I->T 763 L->W

Q1_82	846	HM597864	Chu_Q1	751 del A	N.s.
Q1_83	837	HM597848	Chu_Q1	751 del A	N.s.
Q1_84	843	HM597850	Chu_Q1	751 del A	N.s.
Q1_85	835	HM597857	Chu_Q1		
Q1_86	855	HM597853	Chu_Q1	751 del A	N.s.
Q1_87	834	HM597868	Chu_Q1		
Q1_88	839	HM597847	Chu_Q1	751 del A	N.s.
Q1_89	788	FJ375349	Chu_Q1		1477 W->C
Q1_90	837	HM597851	Chu_Q1	751 del A	N.s.
Q1_91	835	HM597867	Chu_Q1		
Q1_92	839	HM597862	Chu_Q1		739 P->T
Q1_93	833	HM597859	Chu_Q1		754 V->A 769 S->I
Q1_94	838	HM597866	Chu_Q1		
Q1_95	871	HM802266	Chu_Q1		709 G->C
Q1_96	866	HM802267	Chu_Q1		709 G->C
Q1_97	870	HM802268	Chu_Q1		709 G->C
Q1_98	841	HM597856	Chu_Q1	751 del A	N.s.
Q1_99	736	FJ188504	Chu_Q1		
Q1_100	781	FJ188524	Chu_Q1		
Q1_101	773	FJ188507	Chu_Q1		
Q1_102	784	FJ188508	Chu_Q1		
Q1_103	784	FJ188552	Chu_Q1		
Q1_104	821	FJ188553	Chu_Q1		
Q1_105	814	EU427722	Chu_Q1		
Q1_106	1242	EU427723	Chu_Q1		
Q1_107	835	EF080823	Ahmed_MedBasin1		709 G->C
Q1_108	829	EU427725	Chu_Q1	1512 del T	N.s.
Q1_110	545	JN966761	Chu_Q1		
Q1_111	687	DQ989547 DQ989548 DQ989549 DQ989550	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1		
Q1_112	656	HM488309 HM488310 HM488311	Chu_Q1 Chu_Q1 Chu_Q1		
Q1_113	656	HM488312 HM488313 HM488314 HM488331	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1	751 del A	N.s.
Q1_114	656	HM488315	Chu_Q1	751 del A	N.s.
Q1_115	656	HM488324 HM488327 HM488332 HM488338	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1	751 del A	N.s.

Q1_116	658	HM488325	Chu_Q1	751 del A , 1378 ins T	N.s.
Q1_117	657	HM488326 HM488329 HM488330 HM488333 HM488334 HM488335 HM488336 HM488337 HM488339	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1	751 del A	N.s.
Q1_118	657	HM488328 HM488328	Chu_Q1 Chu_Q1	751 del A , 1358 ins T	N.s.
Q1_119	748	DQ462583 DQ462584 DQ462585	Chu_Q1 Chu_Q1 Chu_Q1		
Q1_120	748	DQ462586	Ahmed_MedBasin1		
Q1_121	676	HM807573 HM807574 HM807549 HM807580 HM807579	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1		
Q1_122	777	AB204586 AB204587 AB204588	Chu_Q1 Chu_Q1 Chu_Q1		
Q1_123	777	AB204579	Chu_Q1		
Q1_124	796	EU760723	Gueguen_Q1	751 del A	N.s.
Q1_125	741	FJ766417	Chu_Q1		
Q1_126	816	AJ517769	Ahmed_MedBasin1		1018 F->S
Q1_127	730	DQ174540	Chu_Q1		1186 F->S
Q1_128	779	AY057178	Chu_Q1	AY057178 = a cassava whitefly from Uganda	N.s.
Q1_129	730	DQ174539	Chu_Q1		1327 A->G 1399 F->S 1402 L->S 1414 L->F 1432 L->S 1453 L->P
Q1_130	799	AY057174	Chu_Q1	AY057174 = a sweet-potato whitefly from Uganda	N.s.
Q1_131	789	AM691053	Chu_Q1		
Q1_132	817	AM691050	Chu_Q1		
Q1_133	817	AM691055	Chu_Q1		937 P->L 1207 T->A

Q1_134	817	AM691063	Chu_Q1		910 T->I 1411 V->A
Q1_135	761	DQ365857 DQ365858 DQ365860 DQ365862 DQ365863 DQ365865 DQ365866 DQ365867 DQ365868 DQ365869 DQ365870	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1	1512 del T	N.s.
Q1_136	761	DQ365871	Chu_Q1	1512 del T	N.s.
Q1_137	676	HM807578	Chu_Q1		
Q1_138	800	AY903578	Chu_Q1		892 D->G
Q1_139	684	DQ302946	Chu_Q1		
Q1_140	757	AM691069 AM691074	Chu_Q1 Chu_Q1		
Q1_141	759	AM691070	Chu_Q1		
Q1_142	757	AM691075	Chu_Q1		
Q1_143	762	AM691076	Chu_Q1		
Q1_144	752	AM691077	Chu_Q1		
Q1_145	758	AM691079	Chu_Q1		
Q1_146	761	AM691081	Ahmed_MedBasin1		
Q1_147	750	AM691082	Chu_Q1		
Q1_148	724	AM691083	Chu_Q1		
Q1_149	801	EU760753	Gueguen_Q1	751 del A	N.s.
Q1_150	730	DQ174541	Ahmed_MedBasin1		1453 L->P
Q1_151	657	GU086330	Chu_Q1		
Q1_152	761	DQ365856 DQ365861 DQ365864	Chu_Q1 Chu_Q1 Chu_Q1	1512 del T	N.s.
Q1_153	780	EU760736	Gueguen_Q1		
Q1_154	676	HM807571 HM807563 HM807560 HM807561 HM807550	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1		
Q1_155	728	AY827614	Chu_Q1		766 I->M 865 H->P
Q1_156	739	AY827613	Chu_Q1		766 I->M 820 A->P
Q1_157	730	AY827612	Chu_Q1		862 H->P
Q1_158	739	AY827615	Chu_Q1		766 I->M 826 L->W

					829 T->P 844 G->S 871 F->I 916 A->G 1078 S->P
Q1_159	728	EF398125	Chu_Q1		859 G->R
Q1_160	728	EF398121	Chu_Q1		862 H->Q
Q1_161	729	EF398117	Chu_Q1	1539 ins T	N.s.
Q1_162	728	EF398120	Chu_Q1		862 H->N 901 A->V
Q1_163	728	EF398115	Chu_Q1		
Q1_164	729	EF398123	Chu_Q1		
Q1_165	778	AF342773	Chu_Q1		1021 T->M
Q1_166	687	DQ989546	Ahmed_MedBasin1		1021 T->M
Q1_167	822	EU099427	Chu_Q1	First 116 bp cluster of mutations, 750 del A	N.s.
Q1_168	789	DQ133378	Chu_Q1	First 8 bp cluster of mutations 750 del A, 765 del A, 1499 del T, 1510 del TT	N.s.
Q1_169	788	DQ133379	Chu_Q1	First 7 bp cluster of mutations 750 del A, 765 del A, 1501 del A, 1507 del G	N.s.
Q1_170	790	DQ133380	Chu_Q1	First 9 bp cluster of mutations 750 del A, 765 del A, 1501 del A, 1510 del TT	N.s.
Q1_171	806	EU760726	Gueguen_Q1	750 del A	N.s.
Q1_172	807	EU760730	Gueguen_Q1	750 del A	N.s.
Q1_173	812	EU760732	Gueguen_Q1	750 del A	N.s.
Q1_174	794	EU760722	Gueguen_Q1	1459 del TA	N.s.
Q1_175	811	EU760728	Gueguen_Q1	750 del A	N.s.
Q1_176	798	EU760729	Gueguen_Q1		
Q1_177	741	FJ766392 FJ766394 FJ766396 FJ766397 FJ766401 FJ766414 FJ766430 FJ766436	Gueguen_Q1 Gueguen_Q1 Gueguen_Q1 Gueguen_Q1 Gueguen_Q1 Gueguen_Q1 Gueguen_Q1 Gueguen_Q1		

		FJ766426 FJ766425	Gueguen_Q1 Gueguen_Q1		
Q1_178	741	FJ766381	Gueguen_Q1		
Q1_179	673	FJ766125	Chu_Q1	FJ766125 = AtpF gene in chloroplast DNA from *Cenchrus compressus*	N.s.
Q1_180	650	FJ766126	Chu_Q1	FJ766126 = AtpF gene in chloroplast DNA from *Phacelurus latifolius*	N.s.
Q1_181	676	HM807568 HM807566 HM807562 HM807551 HM807548 HM807543 HM807544 HM807545 HM807546 HM807538 HM807539	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1		
Q1_182	776	AY057138	Chu_Q1		766 I->M 1270 F->C
Q1_183	831	EU192061	Chu_Q1	First 10 bp cluster of mutations 710 del GT, 750 ins GA, 750 del A, 1511 del T	N.s.
Q1_184	831	EU192049	Ahmed_ MedBasin1	First 10 bp cluster of mutations 710 del GT, 750 ins GA, 750 del A, 1511 del T	N.s.
Q1_185	657	GU086337	Chu_Q1		1258 V->G
Q1_186	545	JN966876	Chu_Q1		1258 V->G
Q1_187	657	GU086338	Chu_Q1		1258 V->E
Q1_188	657	GU086336	Chu_Q1		805 L->W 922 M->K 973 L->W 1258 V->E
Q1_189	657	GU086339	Chu_Q1		
Q1_190	817	EF694108	Chu_Q1		1078 S->F 1360 S->F
Q1_191	716	AM691064	Chu_Q1		1219 Y->H

Q1_192	818	AM176575	Ahmed_MedBasin1		988 S->F
Q1_193	817	AM176574	Chu_Q1		1054 L->S 1447 F->L
Q1_194	817	AM176571	Chu_Q1		
Q1_195	817	AM180063	Ahmed_MedBasin1		769 S->G 988 S->P 1045 M->T
Q1_196	759	HM137320	Chu_Q1		1219 Y->C 1477 W->L
Q1_197	759	HM137333	Chu_Q1		117 V->D
Q1_198	759	AM691059	Chu_Q1		
Q1_199	793	EU760743	Gueguen_Q1	750 del A	N.s.
Q1_200	805	EU760744	Gueguen_Q1	750 del A	N.s.
Q1_201	545	JN966877 JN966878 JN966879 JN966880	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1		
Q1_202	755	AM691062	Chu_Q1		
Q1_203	795	EU760737	Gueguen_Q1	750 del A	N.s.
Q1_204	741	FJ766432 FJ766434	Gueguen_Q1 Gueguen_Q1		
Q1_205	807	EU760725	Gueguen_Q1	750 del A	N.s.
Q1_206	770	EU760721	Gueguen_Q1		
Q1_207	817	AM176573	Ahmed_MedBasin1		1153 I->T
Q1_208	545	JN966871 JN966872 JN966873 JN966874 JN966875	Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1 Chu_Q1		1369 S->G
Q1_209	798	EU760755	Gueguen_Q1	750 del A	N.s.
Q1_210	801	EU760749	Gueguen_Q1	750 del A	N.s.
Q1_211	795	EU760761	Gueguen_Q1	1459 del T	N.s.
Q1_212	790	EU760740	Gueguen_Q1	750 del A, 1459 del T, 1483 ins A, 1512 ins T	N.s.
Q2_1	777	AF342776	Gueguen_Q2		1361 S->F
Q2_2	772	AM944348 AM944347	Ahmed_MedBasin2		
Q2_3	501	AY518191	Ahmed_MedBasin2	787 ins A, short sequence	N.s.
Q2_4	768	DQ365877	Ahmed_MedBasin2	1512 del T	N.s.
Q2_5	747	DQ365878	Ahmed_MedBasin2		
Q2_6	852	EF080821	Ahmed_MedBasin2		709 G->C
Q2_7	729	EF398128	Ahmed_MedBasin2	1539 ins T	N.s.
Q2_8	838	EU427721	Ahmed_MedBasin2		709 G->C
Q2_10	803	EU760751	Gueguen_Q2		

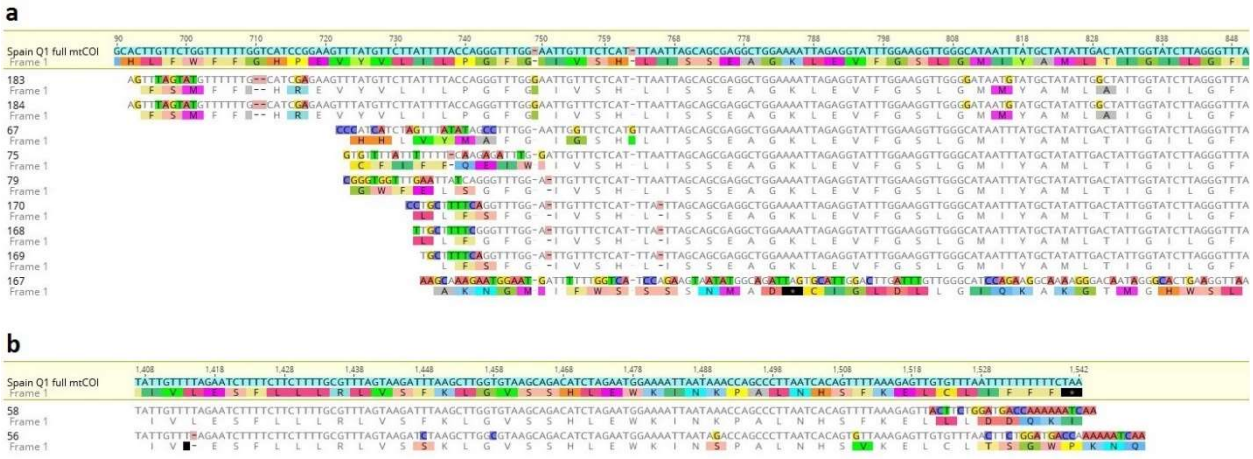
Q2_11	804	EU760754	Gueguen_Q2	751 del A	N.s.
Q2_12	797	EU760756	Gueguen_Q2	751 del A	N.s.
Q2_13	785	EU760757	Gueguen_Q2		
Q2_14	776	FJ998204	Ahmed_MedBasin1	843 ins T, 957 del T, 1498 del C	N.s.
Q2_15	779	AY766372	Chu_Q2	798-818 cluster of mutations; 804 ins A, 1496 ins GGACT, 1515 ins C, last 87 bp cluster of mutations	N.s.
Q2_16	676	HM807534	Chu_Q2		
Q2_17	687	DQ989554 DQ989553	Chu_Q2 Chu_Q2		
Q2_18	777	AB297895 AB297896	Chu_Q2 Chu_Q2		
Q2_19	740	AY827617 AY827618	Chu_Q2 Chu_Q2		766 I->M 769 S->G
Q2_20	811	FJ188567	Chu_Q2		
Q2_21	837	FJ188539	Chu_Q2		709 G->C
Q2_22	800	FJ188558	Chu_Q2		
Q2_23	823	FJ188480	Chu_Q2		
Q2_24	822	FJ188481 FJ188559	Chu_Q2 Chu_Q2		
Q2_25	821	FJ188482	Chu_Q2		
Q2_26	745	FJ188483	Chu_Q2		
Q2_27	803	FJ188543	Chu_Q2		
Q2_28	802	FJ188541	Chu_Q2		
Q2_29	545	JN966827 JN966870	Chu_Q2 Chu_Q2		
Q2_30	740	AY827619	Chu_Q2		766 I->M 769 S->G 829 T->A 838 I->M 946 I->M
Q2_31	657	GU086331	Chu_Q2		1361 S->F
Q2_32	780	AY766371	Chu_Q2	798-816 cluster of mutations; 798 ins A, 1483 ins AC, last 89 bp cluster of mutations	N.s.
Q2_33	780	AY766370	Chu_Q2	798-819 cluster of mutations; 798 ins A, 1485 del T, last 88 bp cluster of mutations	N.s.

Q2_34	777	AB297897	Chu_Q2		1183 W->R 1468 H->R
Q2_35	777	AB297898	Chu_Q2		796 F->V 1297 G->E
Q2_36	727	EU760756	Chu_Q2		
Q2_37	715	EU760757	Chu_Q2		
Q3_Croatia_1	657	GU086334	Chu_Q2		1259 V->E
Q3_Croatia_2	657	GU086335	Chu_Q2		
Q3/Q5_1	741	FJ766382 FJ766386 FJ766387 FJ766395 FJ766409 FJ766424 FJ766427 FJ766417	Gueguen_Q3 Gueguen_Q3 Gueguen_Q3 Gueguen_Q3 Gueguen_Q3 Gueguen_Q3 Gueguen_Q3 Gueguen_Q3		
Q3/Q5_2	741	FJ766405 FJ766419	Gueguen_Q3 Gueguen_Q3		
Q3/Q5_3	741	FJ766420 FJ766421 FJ766422 FJ766423 FJ766400 FJ766428	Gueguen_Q3 Gueguen_Q3 Gueguen_Q3 Gueguen_Q3 Chu_Q5 Gueguen_Q3		
Q3/Q5_4	741	FJ766429	Gueguen_Q3		1204 L->V
Q3/Q5_5	741	FJ766384	Gueguen_Q3		1204 L->V
Q3/Q5_6	737	FJ766431	Gueguen_Q3		1204 L->V
Q3/Q5_7	740	FJ766408	Gueguen_Q3	1515 del A	N.s.
Q3/Q5_8	741	FJ766385	Gueguen_Q3		
ASL/Q4_1	800	AY903549 AY903533 AY903555 AY903572 AY903574	Chu_Q4 Chu_Q4 Chu_Q4 Chu_Q4 Chu_Q4		
ASL/Q4_2	800	AY903532	Chu_Q4		
ASL/Q4_3	800	AY903556	Chu_Q4		
ASL/Q4_4	800	AY903551	Chu_Q4		1531 F->S 1393 I->M
ASL/Q4_5	800	AY903566 AY903546 AY903545	Chu_Q4 Chu_Q4 Chu_Q4		
ASL/Q4_6	800	AY903550	Chu_Q4	943 ins A	N.s.
ASL/Q4_7	749	AY827606	Chu_Q4		775 E->K
ASL/Q4_8	800	AY903564	Chu_Q4	1538 del T	N.s.
ASL/Q4_9	800	AY903565	Chu_Q4	15 SNPs	790 E->K

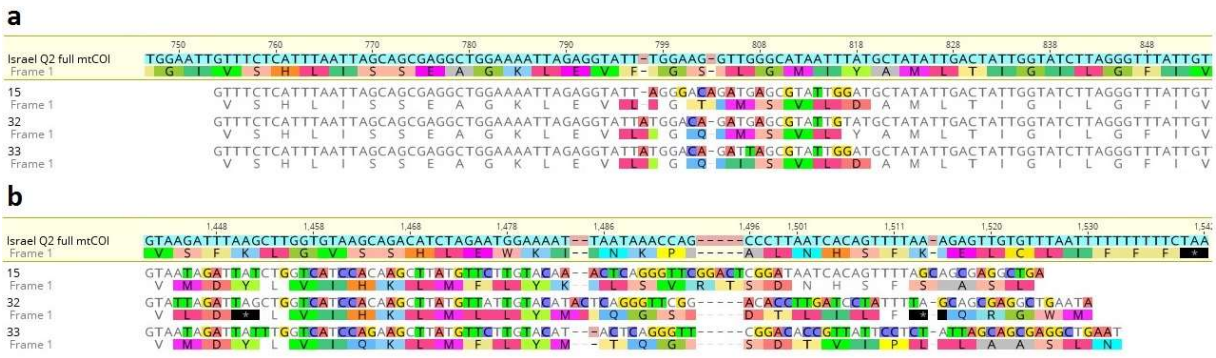
					805 L->V 853 V->G 859 G->S 862 H->P 898 R->P 1027 F->I 1114 V->G 1117 V->G 1129 H->P
ASL/Q4_10	800	AY903552	Chu_Q4	949 ins T	N.s.
ASL/Q4_11	800	AY903529	Chu_Q4	1539 ins T	N.s.
ASL/Q4_12	800	AY903573	Chu_Q4		1003 P->A 1360 S->F 1498 L->F
ASL/Q4_13	800	AY903531	Chu_Q4		
ASL/Q4_14	800	AY903541 AY903540	Chu_Q4 Chu_Q4		
ASL/Q4_15	800	AY903534	Chu_Q4	1538 del T	N.s.
ASL/Q4_16	800	AY903535	Chu_Q4	1083 ins C	N.s.
ASL/Q4_17	740	AY827590	Chu_Q4		
ASL/Q4_18	740	AY827588	Chu_Q4		
ASL/Q4_19	741	FJ766388 FJ766390 FJ766391 FJ766393 FJ766398 FJ766410 FJ766413 FJ766415	Gueguen_ASL Gueguen_ASL Gueguen_ASL Gueguen_ASL Gueguen_ASL Gueguen_ASL Gueguen_ASL Gueguen_ASL		
ASL/Q4_20	741	FJ766389	Gueguen_ASL		1468 H->Y
ASL/Q4_21	741	FJ766435	Gueguen_ASL		
ASL/Q4_22	740	AY827587	Chu_Q4		
ASL/Q4_23	739	AY827582	Chu_Q4	19 SNPs	763 L->C 775 E->K 790 E->A 817 Y->M 1414 L->S 1438 L->W 1486 N->I 1489 K->N
ASL/Q4_24	740	AY827579	Chu_Q4		1444 S->K
ASL/Q4_25	740	AY827589	Chu_Q4		
ASL/Q4_26	709	AY827581	Chu_Q4		1249 F->C 1372 I->N 1438 L->W
ASL/Q4_27	740	AY827580	Chu_Q4		1249 F->C 1405 F->L

					1468 H->L 1474 E->D
ASL/Q4_28	741	FJ766418 FJ766399 FJ766402 FJ766403 FJ766404 FJ766406 FJ766411 FJ766407 FJ766412 FJ766416	Gueguen_ASL Chu_Q4 Gueguen_ASL Chu_Q4 Gueguen_ASL Gueguen_ASL Gueguen_ASL Gueguen_ASL Chu_Q4 Gueguen_ASL		
ASL/Q4_29	741	FJ766433 FJ766437	Gueguen_ASL Gueguen_ASL		
ASL/Q4_30	778	AY057136	Chu_Q4		1519 L->I 1525 L->W 1528 I->S
ASL/Q4_31	741	FJ766383	Gueguen_ASL		
ASL/Q4_32	808	EU760731	Gueguen_ASL	751 del A	N.s.
ASL/Q4_33	783	EU760758	Gueguen_ASL		

Appendix 4: Alignments of eleven MED Q1 haplotypes containing clusters of non-synonymous mutations and INDELs (**a**) at 5' end and (**b**) at 3' end of the partial mtCOI sequence.



Appendix 5: Alignments of three MED Q2 haplotypes containing clusters of non-synonymous mutations and INDELs (**a**) at 5' end and (**b**) at 3' end of the partial mtCOI sequence.

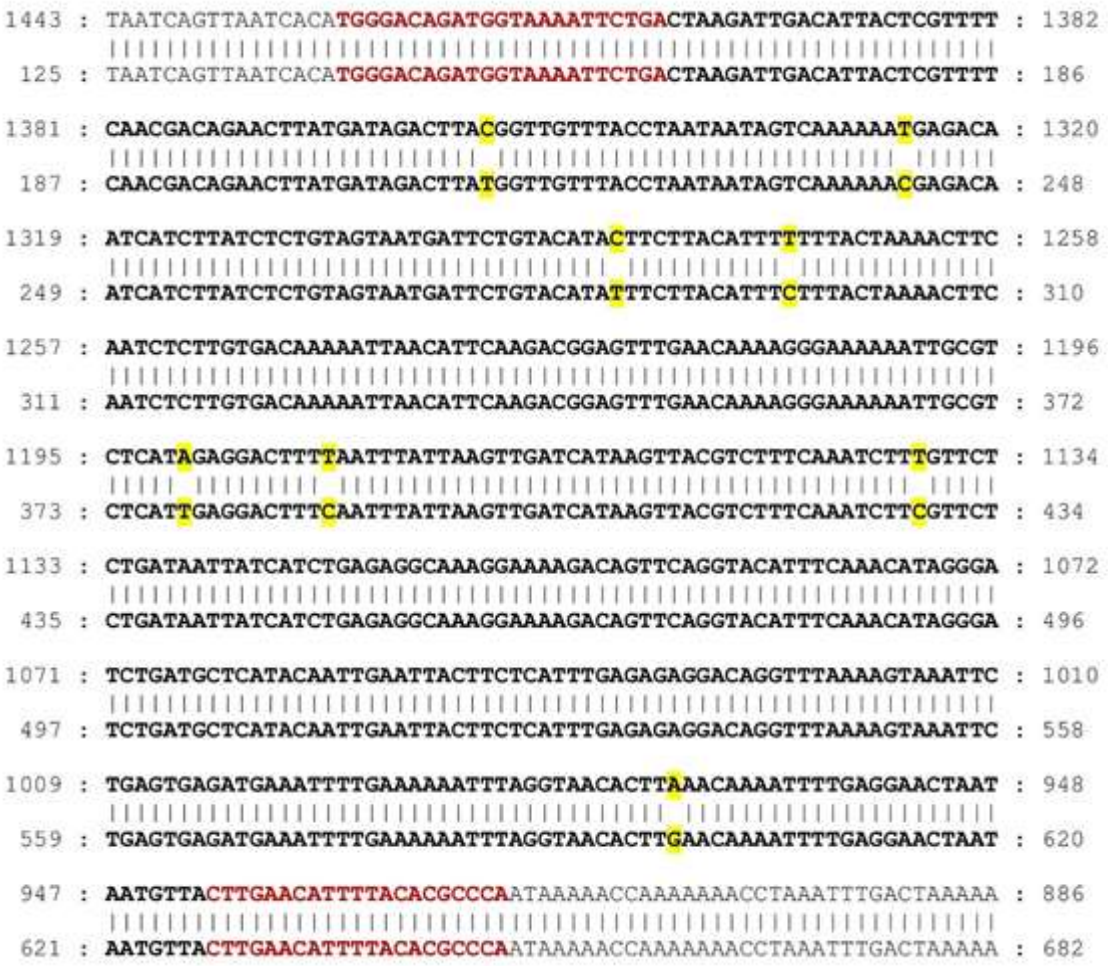


Appendix 6: Candidate nuclear genes encoding mitochondrial proteins that were also found in Spain Q1 and Israel Q2 transcriptome. Grey highlights mark selected candidate genes for which primers were designed.

Category	Protein	Flybase annotation
Complex I	15 kDa	CG11455
Complex I	19 kDa	CG3683
Complex I	23 kDa	CG3944
Complex I	NADH-Q-oxidoreductase, 75 kDa	CG2286
Complex I	B8	CG15434
Complex I	B14	CG7712
Complex I	B14.5A	CG3621
Complex I	B14.5B	CG12400
Complex I	B14.7	CG9350
Complex I	B17	CG13240
Complex I	B17.2	CG3214
Complex I	B18	CG5548
Complex I	B22	CG9306
Complex I	ACP	CG9160
Complex I	MLRQ	CG32230
Complex I	PDSW	CG8844
Complex I	SGDH	CG9762
Complex II	Flavoprotein	CG17246
Complex II	Succinate dehydrogenase, D	CG10219
Complex III	6.4 kDa	CG14482
Complex III	7.2kDa	CG8764
Complex III	Cytochrome CI	CG4769
Complex III	Core protein 2	CG4169
Complex III	Ubiquinone-binding protein QP-C	CG7580
Complex IV	IV	CG10664
Complex IV	Va	CG14724
Complex IV	VIb	CG14235
Complex IV	VIc	CG14028
Complex IV	VIIc	CG2249
ATP synthase	Alpha	CG3612
ATP synthase	Epsilon	CG9032
ATP synthase	D	CG6030
ATP synthase	E	CG3321
ATP synthase	F	CG4692
ATP synthase	Coupling factor 6	CG4412
ATP synthase	Lipid-binding protein	CG1746
Cytochrome c proximal	Cyt-c-p	CG17903
Cytochrome c distal	Cyt-c-d	CG17903
Glutamate carrier 2	DmGC2p	CG12201

Appendix 7: Alignment of the NADH-Q-oxidoreductase, 75kDa subunit gene transcriptomic sequences from Israel Q2 (top) and Spain Q1 (bottom). Highlighted are the amplified region (bold), primer annealing sites (red) and polymorphic nucleotides (yellow). Primers: g17_FWD_2206 and g17_REV_3177.

Appendix 8: Alignment of the Succinate dehydrogenase, subunit D gene transcriptomic sequences from Israel Q2 (top) and Spain Q1 (bottom). Highlighted are the amplified region (bold), primer annealing sites (red) and polymorphic nucleotides (yellow). Primers: g2_FWD and g2_REV.



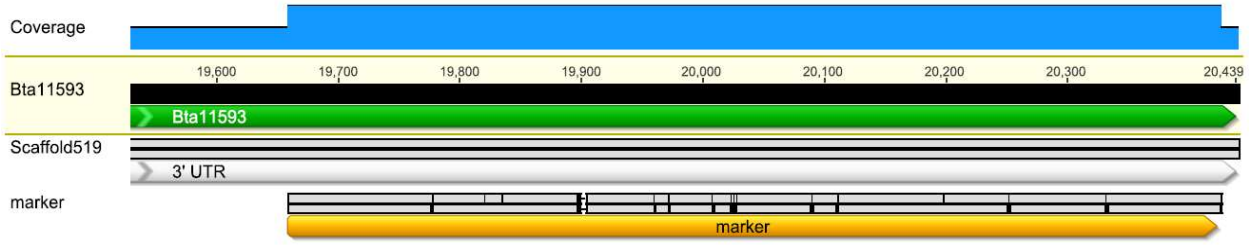
Appendix 9: Alignment of the region 1 of Glutamate carrier 2 gene transcriptomic sequences from Israel Q2 (top) and Spain Q1 (bottom). Highlighted are the amplified region (bold), primer annealing sites (red) and polymorphic nucleotides (yellow). Primers: g3r1_FWD and g3r1_REV2.

```
5812 : CTTTCATAACAAAAAAATTCAACTAGGTATAGCAACCAGATAAGGGCTAGTAAAGCAGGGAA : 5751
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
  60 : CTTTCATAACAAAAAAATTCAACTAGGTATAGCAACCAGATAAGGGCTAGTAAAGCAGGGAA : 121

5750 : ACATACAATTTCGACATTTTGATCATTTTTTTTAAATTTCTTAATGATTCCCAAAGAGATGG : 5689
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
 122 : ACATACAATTTCGACATTTTGATCATTTTTTTTAAATTTCTTAATGATTCCCAAAGAGCTGG : 183

5688 : TGGTATTGCTTGTAAAATTATGATGTTATAGGTTTAATGACAATATATTTCAGGCTTAAATC : 5627
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 184 : TGGTATTGCTTGTAAAATTATGATGTTATAGGTTTAATGACAATATATTTCAGGCTTAAATC : 245

5626 : GATGGCCAGCTCTTAACCTGAACATCTGCAAGCATGTCTAAAAACAGAATATATCCCAGGGC : 5565
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 246 : GATGGCCAGCTCTTAACCTGAACATCTGCAAGCATGTCTAAAAACAGAATATATCCCAGGGC : 307

5564 : TGCAGTTGAATTTACTAAGTTTCTGAGCAACTGCCGTTTTGCTTGCTTTAGATAGTTTCTTG : 5503
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 308 : TGCAGTTGAATTTACTAAGTTTCTGAGCAACTGCCGTTTTGCTTGCTTTAGATAGTTTCTTG : 369

5502 : ACCCTTAGGAACATACTGTCTCAGGGTTTTAGTGCTGGGCACCTTCATTATTTCAGCTGCAT : 5441
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
 370 : ACCCTTAGGAACATACTGTCTCAGGGTTTTAGTGCTGGGCACCTTCATTATTTCGGCTGCAT : 431

5440 : TACATTCGCTTACATTCATGATATTTTAATTATGATGTACGTAGCTCTATAATTAGGGACCA : 5379
       ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
 432 : TACATTCGCTTACATTCATGATATTTTAATTATGATGTATGTAGCTCTATAATTAGGGACCA : 493

5378 : AACTCAAGGGGAAAAGTCTGACAAGAAAACATTCA----CTTCTCTTGACTGAACGGAGAGC : 5321
       ||||||||||||||||||||||||||||||||||||      | ||||||||||||||||||||
 494 : AACTCAAGGGGAAAAGTCTGACAAGAAAACATTCAAGAGATCCTCTTGACTGAACGGAGAGC : 555

5320 : TCAGCTAAAAAATTTTTCTGGTATTTCTCAGCGATGCTGAATTGCTCACACAGGCCTCGGAC : 5259
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 556 : TCAGCTAAAAAATTTTTCTGGTATTTCTCAGCGATGCTGAATTGCTCACACAGGCCTCGGAC : 617

5258 : CACCAACCTAAAGAGTTCAAAATTGGGAAGATACTGTAGAATAATTTATTGTTTTACAAAAA : 5197
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 618 : CACCAACCTAAAGAGTTCAAAATTGGGAAGATACTGTAGAATAATTTATTGTTTTACAAAAA : 679

5196 : AGCATCGAGATAATGAGGGATTTATCCCAAAATCAAATATGATGAAATTTAGGCGACTGACT : 5135
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 680 : AGCATCGAGATAATGAGGGATTTATCCCAAAATCAAATATGATGAAATTTAGGCGACTGACT : 741

5134 : TTTGATGACACAGATTGTTTACATTTCCTCCTCGGTTGGATCAAAAATAGAATTTTATACAA : 5073
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 742 : TTTGATGACACAGATTGTTTACATTTCCTCCTCGGTTGGATCAAAAATAGAATTTTATACAA : 803
```

Appendix 10: Alignment of the region 2 of Glutamate carrier 2 gene transcriptomic sequences from Israel Q2 (top) and Spain MED-Q1 (bottom). Highlighted are the amplified region (bold), primer annealing sites (red) and polymorphic nucleotides (yellow). Primers: g3r2_FWD and g3r2_REV.

```
422  : TTACTGAACCTTGAAGTGTGCCTCGAAAGGATCTGAAGGGGGAGCTGATGTTTGAAATGCGC : 361
5452 : TTACTGAACCTTGAAGTGTGCCTCGAAAGGATCTGAAGGGGGAGCTGATGTTTGAAATGCGC : 5513

360  : CAGGGAGAGCACGCGTAAGGAAAAACGTTTAGCAAGTTAATTCACGCTATTGAATTTTGGAT : 299
5514 : CAGGGAGAGCACGCGTAAGGAAAAACGTTTGGCAAGTTAATTCACGCTATTGAATTTTGGAT : 5575

298  : CGAATAACGATAATGGACAGATTTTGCTTAGACTGGTACGTTTTCAGACTCTCCGGCCCCTT : 237
5576 : CGAATAACAATAATGGTCAGATTTTGCTTAGACTGGTACGTTTTCAGACTCTCCGGCCCCTT : 5637

236  : GACACTCGCGAAAAATCACCAGAGAAGTCGAAGAACAAGACACTAAACTTTCAATGAAACAC : 175
5638 : GACACTCGCGAAAAATCACCAGAGAAGTCGAAGAACTAGACACTAAACTTTCAATGAAACAC : 5699

174  : GAAAAATAGATAACAATGAATGAGGAGATAACCAACAGCGACGATAATTAGAGCGGGGGACA : 113
5700 : GAAAAATAGATAACAATGAATGAGGAGATAACCAACAGCGACGATAATTAGAGCGGGGGACA : 5761

112  : TGGTAAATTAATACAAGGAGGGACGCGAAAGGCGGAAAAGAGTGACAAATAAAATATGATGT : 51
5762 : TGGTCAATTAATACAAGGAGGGACGCGAAAGGCGGAAAAGAGTGACAAATAAAATATGATGT : 5823
```

Appendix 11: Alignment of the GC2, region 1 marker sequence (from Israel Q2) to the 3' UTR region of the GC1-like gene Bta11593 from the MEAM1 genome.



Appendix 12: P values from multiple comparison of progeny counts and female proportions from F_1 crosses among Spain Q1, Sudan Q1, Israel Q2 and Uganda ASL. Missing data indicate that no comparison was done due to the absence of females in the offspring. Key: '***' <0.001, '**' <0.01, '*' <0.05, '.' <0.1.

Comparison	P values	
	Progeny counts	Female proportion
Q2xSpQ1 - Q2control	0.9992	0.983
Q2xSuQ1 - Q2control	0.9880	
Q2xASL - Q2control	0.9999	
SpQ1control - Q2control	1.0000	0.998
SpQ1xQ2 - Q2control	0.9805	1.000
SpQ1xSuQ1 - Q2contro	0.6203	
SpQ1xASL - Q2control	1.0000	
SuQ1control - Q2control	1.0000	0.999
SuQ1xQ2 - Q2control	1.0000	
SuQ1xSpQ1 - Q2control	0.8184	
SuQ1xASL - Q2control	0.9530	
ASLcontrol - Q2control	0.7165	0.998
ASLxQ2 - Q2control	1.0000	
ASLxSpQ1 - Q2control	1.0000	
ASLxSuQ1 - Q2control	0.9992	
Q2xSuQ1 - Q2xSpQ1	0.5291	
Q2xASL - Q2xSpQ	0.7291	
SpQ1control - Q2xSpQ1	0.9958	1.000
SpQ1xQ2 - Q2xSpQ1	1.0000	0.979
SpQ1xSuQ1 - Q2xSpQ1	0.0558 .	
SpQ1xASL - Q2xSpQ1	1.0000	
SuQ1control - Q2xSpQ1	1.0000	0.999
SuQ1xQ2 - Q2xSpQ1	1.0000	
SuQ1xSpQ1 - Q2xSpQ1	0.1834	
SuQ1xASL - Q2xSpQ1	0.3073	
ASLcontrol - Q2xSpQ1	1.0000	0.767
ASLxQ2 - Q2xSpQ1	0.9860	
ASLxSpQ1 - Q2xSpQ1	0.9997	
ASLxSuQ1 - Q2xSpQ1	1.0000	
Q2xASL - Q2xSuQ1	1.0000	
SpQ1control - Q2xSuQ1	0.9907	
SpQ1xQ2 - Q2xSuQ1	0.2665	
SpQ1xSuQ1 - Q2xSuQ1	1.0000	
SpQ1xASL - Q2xSuQ1	0.8234	
SuQ1control - Q2xSuQ1	0.6947	
SuQ1xQ2 - Q2xSuQ1	0.7728	
SuQ1xSpQ1 - Q2xSuQ1	1.0000	
SuQ1xASL - Q2xSuQ1	1.0000	
ASLcontrol - Q2xSuQ1	0.0477 *	
ASLxQ2 - Q2xSuQ	0.9992	
ASLxSpQ1 - Q2xSuQ1	0.9941	
ASLxSuQ1 - Q2xSuQ1	0.4345	
SpQ1control - Q2xASL	1.0000	

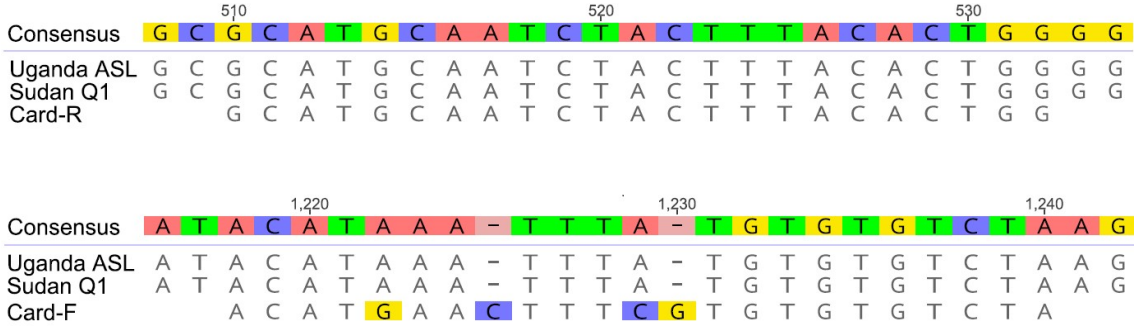
SpQ1xQ2 - Q2xASL	0.3585	
SpQ1xSuQ1 - Q2xASL	0.9525	
SpQ1xASL - Q2xASL	0.9587	
SuQ1control - Q2xASL	0.8729	
SuQ1xQ2 - Q2xASL	0.9349	
SuQ1xSpQ1 - Q2xASL	0.9892	
SuQ1xASL - Q2xASL	0.9998	
ASLcontrol - Q2xASL	0.0318 *	
ASLxQ2 - Q2xASL	1.0000	
ASLxSpQ1 - Q2xASL	1.0000	
ASLxSuQ1 - Q2xASL	0.5619	
SpQ1xQ2 - SpQ1control	0.9397	0.998
SpQ1xSuQ1 - SpQ1control	0.6012	
SpQ1xASL - SpQ1control	1.0000	
SuQ1control - SpQ1control	0.9999	1.000
SuQ1xQ2 - SpQ1control	1.0000	
SuQ1xSpQ1 - SpQ1control	0.8247	
SuQ1xASL - SpQ1control	0.9574	
ASLcontrol - SpQ1control	0.4976	0.912
ASLxQ2 - SpQ1control	1.0000	
ASLxSpQ1 - SpQ1control	1.0000	
ASLxSuQ1 - SpQ1control	0.9942	
SpQ1xSuQ1 - SpQ1xQ2	<0.01 **	
SpQ1xASL - SpQ1xQ2	0.9979	
SuQ1control - SpQ1xQ2	1.0000	0.999
SuQ1xQ2 - SpQ1xQ2	1.0000	
SuQ1xSpQ1 - SpQ1xQ2	0.0622 .	
SuQ1xASL - SpQ1xQ2	0.1067	
ASLcontrol - SpQ1xQ2	1.0000	0.993
ASLxQ2 - SpQ1xQ2	0.8910	
ASLxSpQ1 - SpQ1xQ2	0.9926	
ASLxSuQ1 - SpQ1xQ	1.0000	
SpQ1xASL - SpQ1xSuQ1	0.1231	
SuQ1control - SpQ1xSuQ1	0.0716 .	
SuQ1xQ2 - SpQ1xSuQ1	0.1306	
SuQ1xSpQ1 - SpQ1xSuQ1	1.0000	
SuQ1xASL - SpQ1xSuQ1	1.0000	
ASLcontrol - SpQ1xSuQ1	<0.01 ***	
ASLxQ2 - SpQ1xSuQ1	0.8614	
ASLxSpQ1 – SpxSuQ1	0.7687	
ASLxSuQ1 - SpQ1xSuQ1	0.0167 *	
SuQ1control - SpQ1xASL	1.0000	
SuQ1xQ2 - SpQ1xASL	1.0000	
SuQ1xSpQ1 - SpQ1xASL	0.3873	
SuQ1xASL - SpQ1xASL	0.5868	
ASLcontrol - SpQ1xASL	0.8403	
ASLxQ2 - SpQ1xASL	1.0000	
ASLxSpQ1 - SpQ1xASL	1.0000	
ASLxSuQ1 - SpQ1xASL	1.0000	
SuQ1xQ2 - SuQ1control	1.0000	

SuQ1xSpQ1 – Sucontrol	0.2698	
SuQ1xASL - SuQ1control	0.4303	
ASLcontrol - SuQ1control	0.9734	0.909
ASLxQ2 - SuQ1control	0.9991	
ASLxSpQ1 - SuQ1control	1.0000	
ASLxSuQ1 - SuQ1control	1.0000	
SuQ1xSpQ1 - SuQ1xQ2	0.3521	
SuQ1xASL - SuQ1xQ2	0.5450	
ASLcontrol - SuQ1xQ2	0.9802	
ASLxQ2 - SuQ1xQ2	0.9997	
ASLxSpQ1 - SuQ1xQ2	1.0000	
ASLxSuQ1 - SuQ1xQ2	1.0000	
SuQ1xASL - SuQ1xSpQ1	1.0000	
ASLcontrol - SuQ1xSpQ1	<0.01 **	
ASLxQ2 - SuQ1xSpQ1	0.9477	
ASLxSpQ1 - SuQ1xSpQ1	0.8869	
ASLxSuQ1 - SuQ1xSpQ1	0.1145	
ASLcontrol - SuQ1xASL	<0.01 **	
ASLxQ2 - SuQ1xASL	0.9950	
ASLxSpQ1 - SuQ1xASL	0.9780	
ASLxSuQ1 - SuQ1xASL	0.1903	
ASLxQ2 - ASLcontrol	0.4356	
ASLxSpQ1 - ASLcontrol	0.8537	
ASLxSuQ1 - ASLcontrol	0.9982	
ASLxSpQ1 - ASLxQ2	1.0000	
ASLxSuQ1 - ASLxQ2	0.9813	
ASLxSuQ1 - ASLxSpQ1	0.9998	

Appendix 13: P values from multiple comparison of progeny counts and female proportions fro m F_1 and F_2 crosses between Spain Q1 and Israel Q2 only. Key: '***' <0.001, '**' <0.01, '*' < 0.05, '.' <0.1.

Comparison	P values	
	Progeny counts	**Female proportion**
Q1xQ2 - Q1 control	0.79852	0.9997
Q1xQ2xQ1 - Q1 control	0.34790	0.5896
Q1xQ2xQ2 - Q1 control	0.42764	0.2370
Q2 control - Q1 control	1.00000	0.9997
Q2xQ1 - Q1 control	0.94614	1.0000
Q2xQ1xQ1 - Q1 control	0.65187	0.1932
Q2xQ1xQ2 - Q1 control	0.08323 .	0.5681
Q1xQ2xQ1 - Q1xQ2	0.99984	0.1642
Q1xQ2xQ2 - Q1xQ2	0.99996	0.0349 *
Q2 control - Q1xQ2	0.88726	1.0000
Q2xQ1 - Q1xQ2	0.99999	0.9929
Q2xQ1xQ1 - Q1xQ2	0.04530 *	0.0830 .
Q2xQ1xQ2 - Q1xQ2	0.00203 **	0.4420
Q1xQ2xQ2 - Q1xQ2xQ1	1.00000	0.9923
Q2 control - Q1xQ2xQ1	0.52700	0.3335
Q2xQ1 - Q1xQ2xQ1	0.99563	0.8037
Q2xQ1xQ1 - Q1xQ2xQ1	0.00299 **	0.7256
Q2xQ1xQ2 - Q1xQ2xQ1	< 0.001 ***	0.8850
Q2 control - Q1xQ2xQ2	0.60317	0.1126
Q2xQ1 - Q1xQ2xQ2	0.99800	0.4221
Q2xQ1xQ1 - Q1xQ2xQ2	0.00498 **	0.9107
Q2xQ1xQ2 - Q1xQ2xQ2	< 0.001 ***	0.9499
Q2xQ1 - Q2 control	0.97564	0.9947
Q2xQ1xQ1 - Q2 control	0.65105	0.1086
Q2xQ1xQ2 - Q2 control	0.09079 .	0.4509
Q2xQ1xQ1 - Q2xQ1	0.13110	0.2782
Q2xQ1xQ2 - Q2xQ1	0.00749 **	0.6415
Q2xQ1xQ2 - Q2xQ1xQ1	0.91386	0.9999

Appendix 14: Mapping *Cardinium*-specific primers to contig assembled by mapping Uganda and Sudan HTS reads to a Cardinium 16S sequence from MED Q1 (AB981353). The Card-F primer shows two INDELs and two mismatches with the *Cardinium* contig sequence.

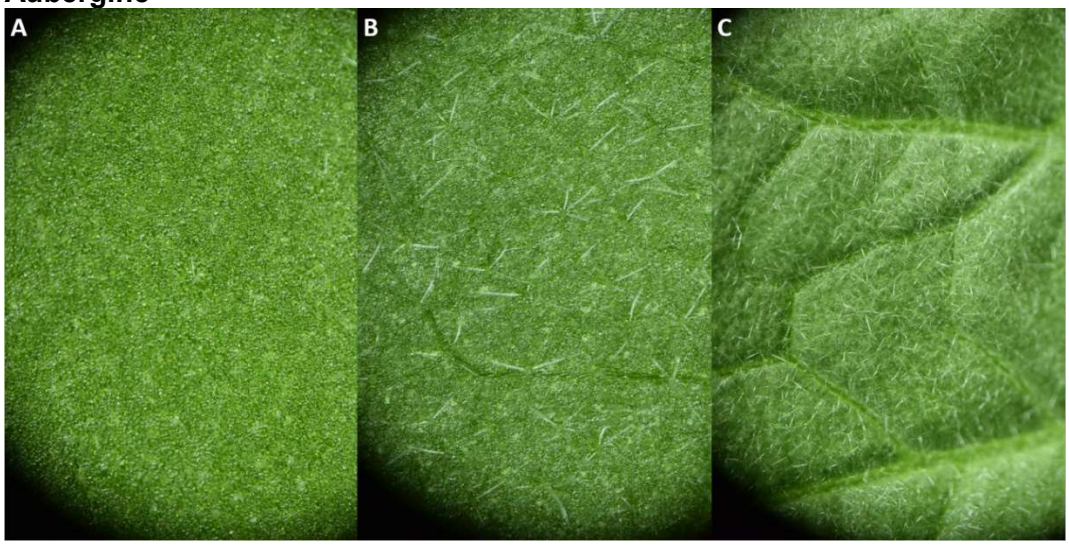


Appendix 15: Averages (Avg.) and standard deviations (St. dev.) of leaf surface area for each leaf stage on 10 plants included in the analysis of oviposition preference for leaves of different age.

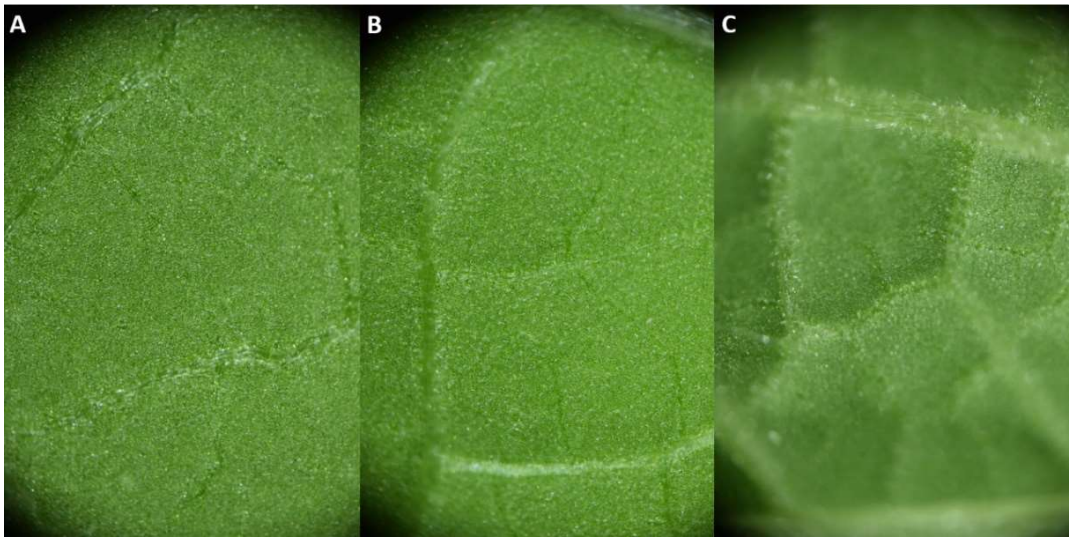
Host	Leaf stage	Avg. area (cm^2)	St. dev.
Aubergine	Mature	30.82	7.22
	Young	66.77	26.27
	Immature	29.93	13.49
Borecole	Mature	17.57	4.89
	Young	26.73	9.58
	Immature	11.59	2.58
Cassava	Mature	11.71	3.91
	Young	11.17	2.59
	Immature	6.17	2.45
Cotton	Mature	23.89	7.95
	Young	41.09	7.68
	Immature	14.37	4.62
Okra	Mature	18.27	14.06
	Young	53.78	22.63
	Immature	21.38	12.50
Pepper	Mature	15.69	6.42
	Young	32.44	12.00
	Immature	23.85	12.99
Squash	Mature	32.73	20.83
	Young	38.85	16.04
	Immature	20.43	15.82
Sweet potato	Mature	22.35	8.79
	Young	27.74	11.40
	Immature	20.73	8.36
Tobacco	Mature	38.38	14.29
	Young	28.98	7.18
	Immature	16.28	2.14
Tomato	Mature	14.46	13.44
	Young	43.70	15.63
	Immature	18.52	7.95

Appendix 16: Photographs of abaxial leaf surface of 13 plants used in our experiments taken at 50x magnification. **A** = mature, **B** = young, **C** = immature leaf.

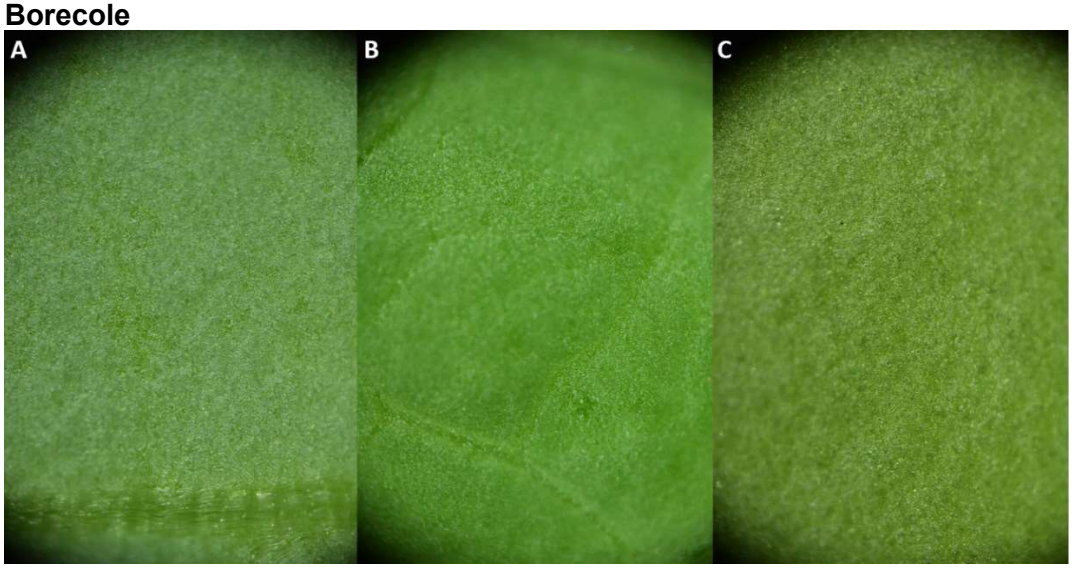
Aubergine



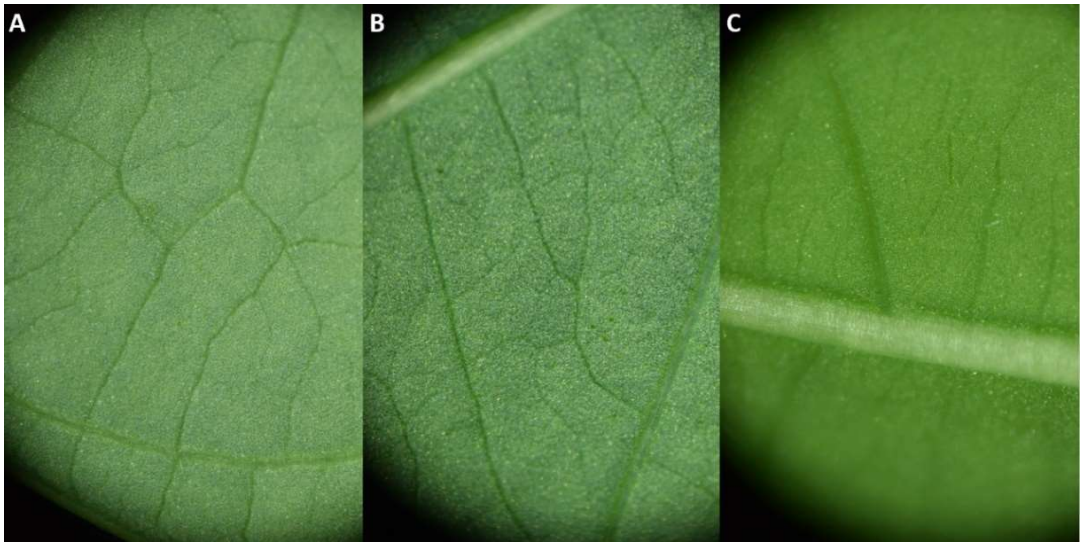
Bean



Borecole

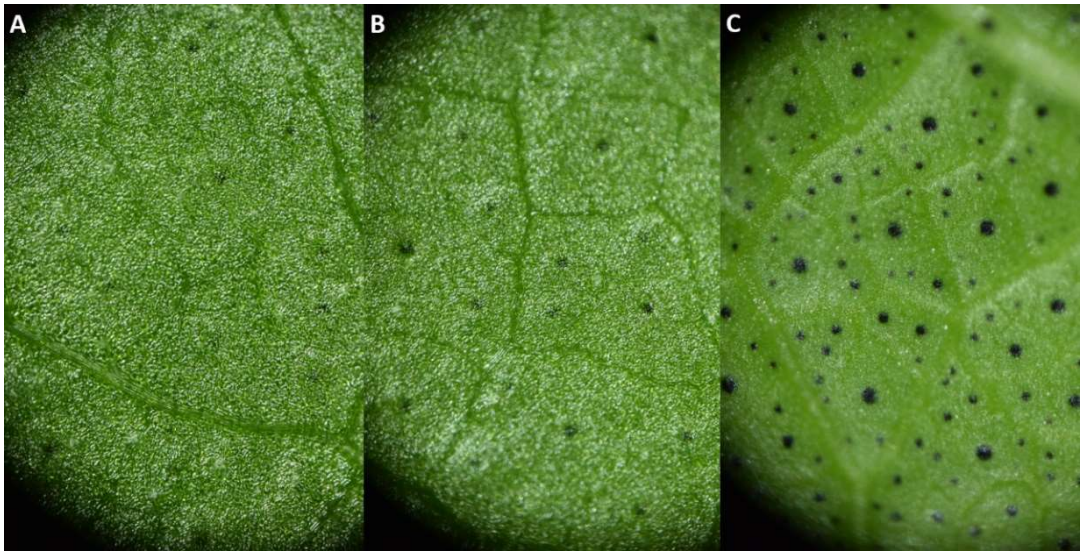


Cassava



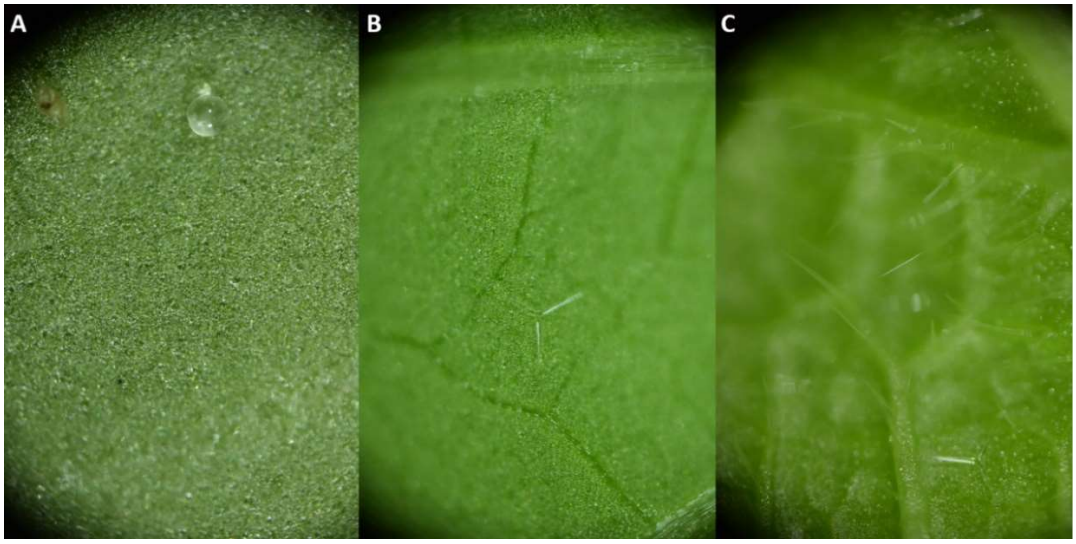
Chard

Cotton

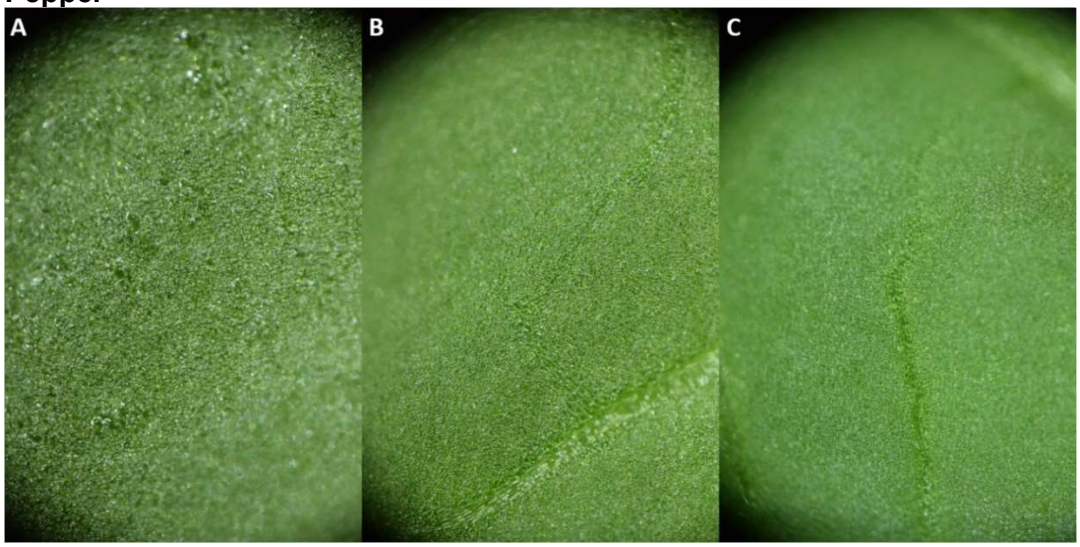


Mint

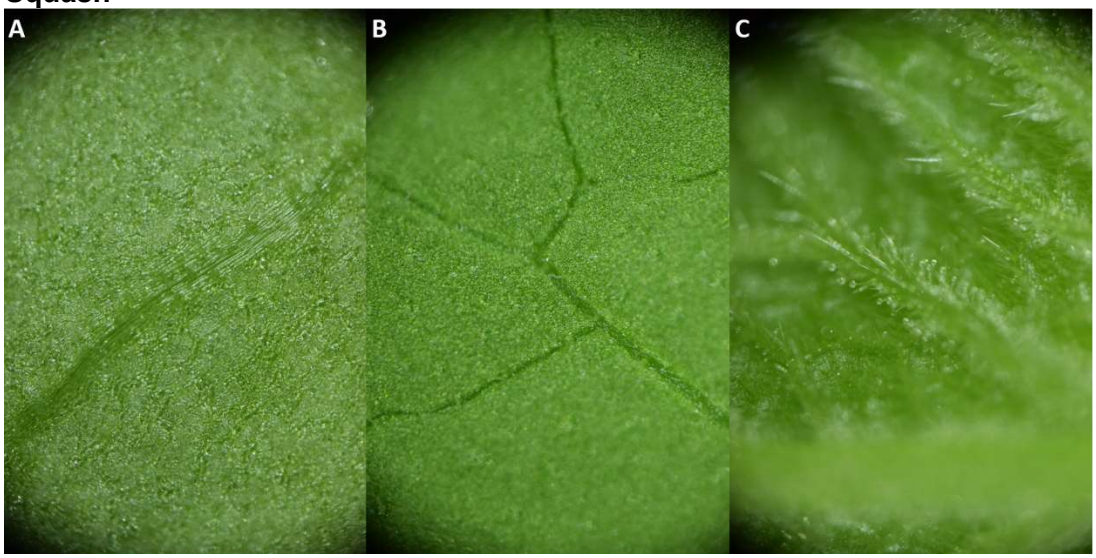
Okra



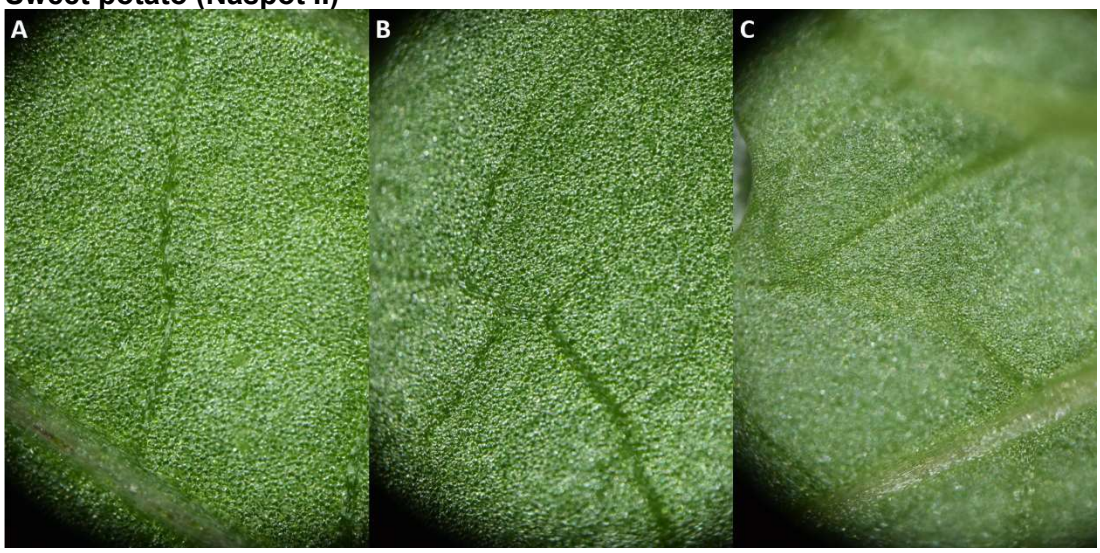
Pepper



Squash

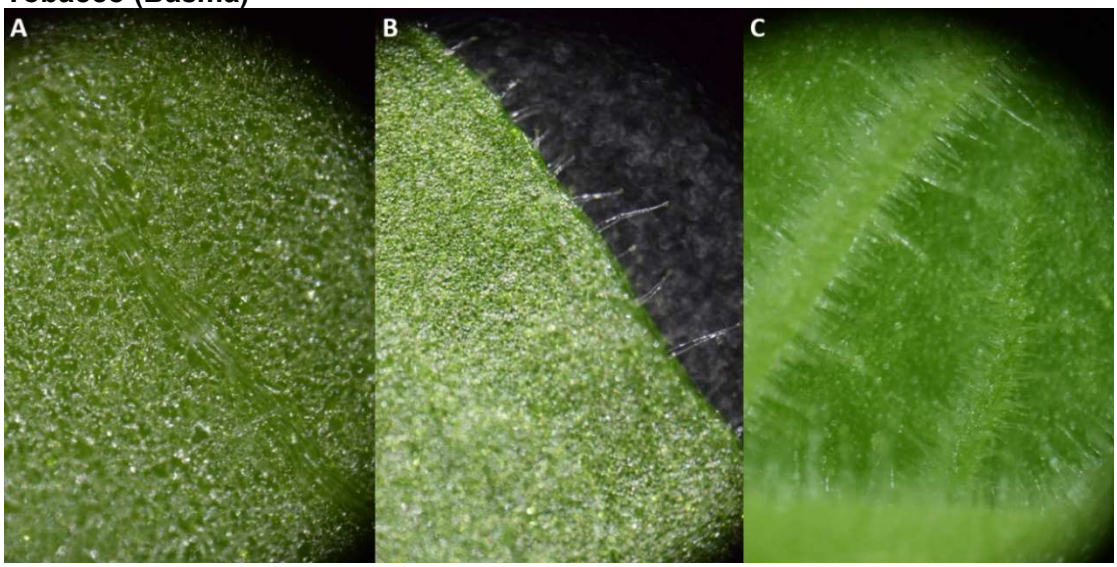


Sweet potato (Naspot II)

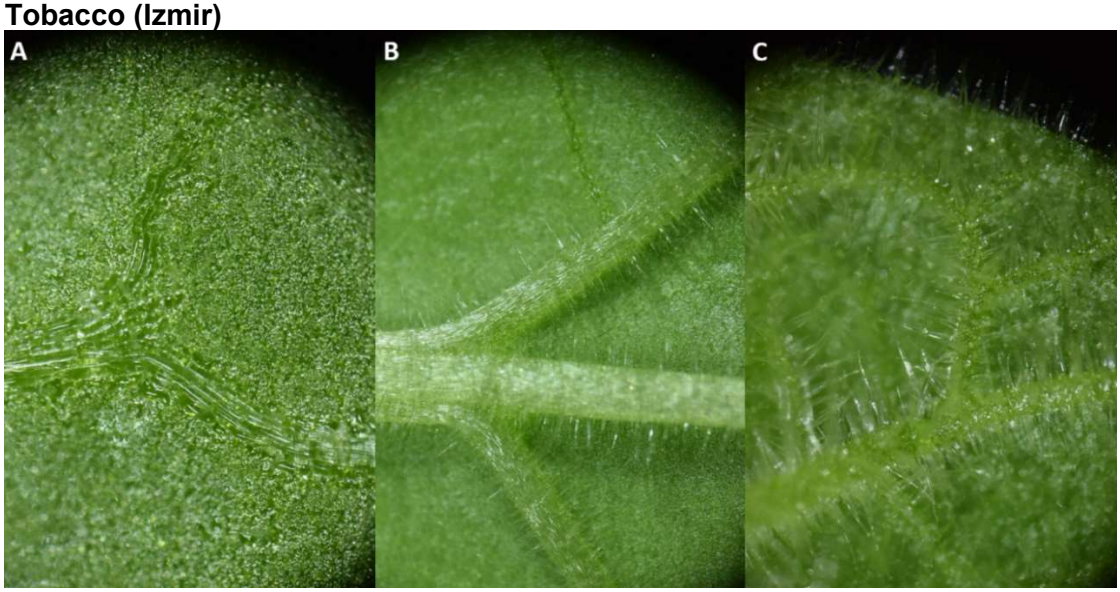


Sweet potato (Beauregard)

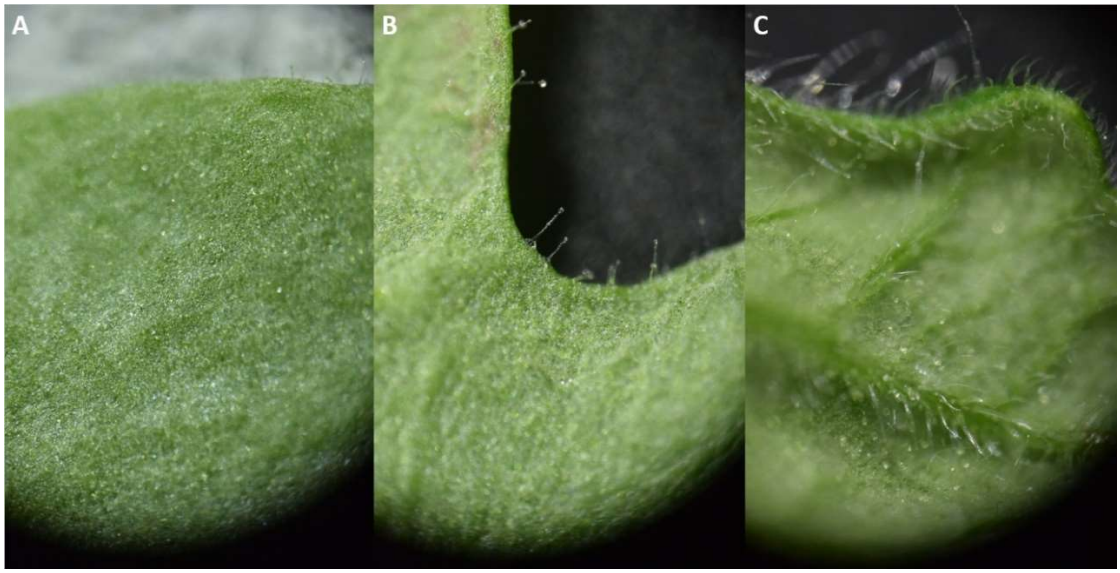
Tobacco (Basma)



Tobacco (Izmir)



Tomato



Appendix 17: P-values from multiple comparisons of oviposition rates. Key: '***' <0.001, '**' <0.01, '*' <0.05, '.' <0.1.

Host	Spain - Israel	Sudan - Israel	Uganda - Israel	Sudan - Spain	Uganda - Spain	Uganda - Sudan
Aubergine	0.266	1.000	0.996	0.248	0.348	0.993
Bean	0.01048 *	0.27728	0.00497 **	0.15501	0.93035	0.06711 .
Borecole	0.9999	0.0739 .	0.0637 .	0.0675 .	0.0585 .	0.9996
Cassava	0.10850	0.34923	0.00973 **	0.81750	0.35823	0.11180
Chard	0.956	0.638	0.524	0.377	0.292	0.997
Cotton	0.0328 *	0.3681	0.1168	0.3510	0.8028	0.8209
Okra	0.98688	0.99952	0.00670 **	0.99587	0.00991 **	0.00746 **
Pepper	0.7841	0.1041	0.0215 *	0.3666	0.0800 .	0.6805
Squash	0.843	0.409	0.992	0.842	0.698	0.290
Sw. potato	1.000	0.918	1.000	0.884	0.998	0.911
Tobacco	0.933	0.824	0.738	0.992	0.421	0.303
Tomato	0.1523	0.8817	0.0969 .	0.0543 .	0.9874	0.0345 *

Appendix 18: P-values from multiple comparisons of mean estimated adult survival times. Key: '***' <0.001, '**' <0.01, '*' <0.05, '.' <0.1.

Host	Spain - Israel	Sudan - Israel	Uganda - Israel	Sudan - Spain	Uganda - Spain	Uganda - Sudan
Aubergine	0.90731	0.07561 .	0.10883	0.00977 **	0.01892 *	1.00000
Bean	0.88889	0.00514 **	< 0.001 ***	0.04422 *	< 0.001 ***	0.00705 **
Borecole	0.0303 *	<0.001 ***	<0.001 ***	0.3745	0.0112 *	0.4788
Cassava	0.963	0.980	<0.001 ***	1.000	<0.001 ***	<0.001 ***
Chard	0.4815	<0.001 ***	<0.001 ***	<0.001 ***	0.0105 *	0.0247 *
Cotton	0.98018	0.00919 **	0.00796 **	0.02456 *	0.02284 *	0.98824
Mint	0.28809	0.07502 .	0.00983 **	0.90062	0.48571	0.89343
Okra	<0.0001 ***	0.00084 ***	<0.0001 ***	<0.0001 ***	<0.0001 ***	0.79152
Pepper	<0.001 ***	0.0382 *	<0.001 ***	0.0702 .	<0.001 ***	<0.001 ***
Squash	0.95232	0.00225 **	0.45768	< 0.001 ***	0.19056	0.18694
Sw. potato	0.0405 *	0.0424 *	<0.001 ***	1.0000	0.3660	0.3706
Tobacco	0.0229 *	0.0121 *	<0.001 ***	<0.001 ***	<0.001 ***	<0.001 ***
Tomato	<0.001 ***	<0.001 ***	<0.001 ***	0.343	0.776	0.896

Appendix 19: P-values from multiple comparisons of F_1 adult progeny counts. Key: '***' <0.001, '**' <0.01, '*' <0.05, '.' <0.1.

Host	Spain - Israel	Sudan - Israel	Uganda - Israel	Sudan - Spain	Uganda - Spain	Uganda - Sudan
Aubergine	<0.001 ***	<0.001 ***	<0.001 ***	0.174	0.565	0.867
Bean	<0.001 ***	<0.001 ***	0.0171 *	0.0617 .	<0.001 ***	<0.001 ***
Borecole	0.9676	0.3769	0.0933 .	0.2127	0.0492 *	0.7285
Cotton	0.26864	0.53781	0.02013 *	0.04521 *	0.45232	0.00432 **
Mint	0.26400	0.31996	0.01518 *	0.02027 *	0.24242	0.00155 **
Okra	< 0.001 ***	< 0.001 ***	< 0.001 ***	< 0.001 ***	< 0.001 ***	0.00597 **
Pepper	< 0.001 ***	0.0152 *	1.0000	0.1749	< 0.001 ***	0.0152 *
Squash	0.0143 *	0.0703 .	<0.001 ***	0.5650	0.1170	0.0125 *
Sw. potato	0.9807	0.0205 *	0.0546 .	0.0342 *	0.0928 .	0.8922
Tobacco	0.04009 *	0.99937	0.25829	0.06278 .	0.00294 **	0.23988
Tomato	0.682	0.322	0.682	0.889	1.000	0.889

Appendix 20: P-values from multiple comparisons of the proportion of females in F_1 progeny. Key: '***' <0.001, '**' <0.01, '*' <0.05, '.' <0.1. Missing data indicate that no comparison was done due to the absence of females in the offspring.

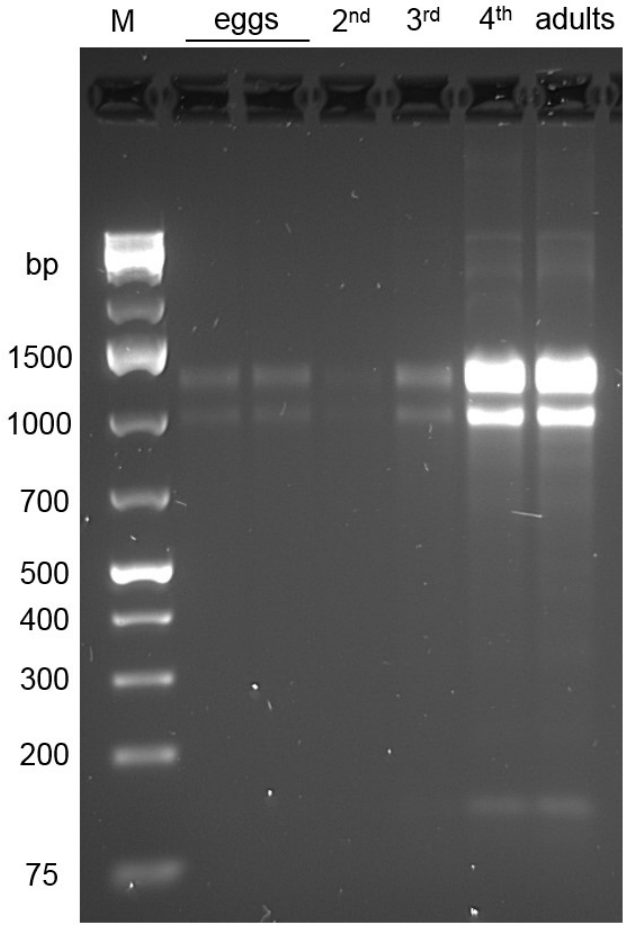
Host	Spain - Israel	Sudan - Israel	Uganda - Israel	Sudan - Spain	Uganda - Spain	Uganda - Sudan
Aubergine	0.6032	0.0526 .	0.6337	<0.001 ***	0.0728 .	0.6126
Bean				0.762	0.881	0.973
Borecole	0.466	0.893	0.380	0.880	0.999	0.812
Cotton	0.15017	0.00517 **	0.17749	< 0.001 ***	1.00000	< 0.001 ***
Mint	0.00115 **	0.57925	0.00383 **	0.07192 .	0.99036	0.14479
Okra		0.955	0.986			0.991
Pepper				0.0003 ***		
Squash	0.822	0.869	0.509	0.998	0.938	0.861
Sw. potato	0.10046	0.00124 **	0.75340	< 0.001 ***	0.56462	< 0.001 ***
Tobacco	0.920	0.630		0.838		
Tomato	0.831		0.831		1.000	

Appendix 21: List of samples of Uganda ASL (UgASL) and Spain Q1 (SpQ1) populations (Pop.) collected from three host plants in year 2017. Each line corresponds to one tube. Lines are ordered by population, plant and life stage. Asterisks mark samples collected for practising purposes only.

Date	Pop.	Host	Stage	Amount	Day	Cage	Note
8.3.	UgASL	okra	eggs	140	6	A	
8.3.	UgASL	okra	eggs	60	6	A	
6.4.	UgASL	okra	eggs	150	7	E	
7.4.	UgASL	okra	eggs	50	8	E	
12.6.	UgASL	okra	eggs	200	1–6	P	
12.6.	UgASL	okra	eggs	200	1–6	P	
14.3.	UgASL	okra	2. instar	100	12	A	
21.3.	UgASL	okra	2. instar	100	15	B	
13.4.	UgASL	okra	2. instar	120	14	E	
21.6.	UgASL	okra	2. instar	200	15	P	
21.3.	UgASL	okra	3. instar	100	15	B	RNA extracted
21.3.	UgASL	okra	3. instar	100	19	A	
24.3.	UgASL	okra	3. instar	100	18	B	
21.4.	UgASL	okra	3. instar	100	22	E	
27.6.	UgASL	okra	3. instar	150	21	P	
27.6.	UgASL	okra	3. instar	210	21	P	
28.3.	UgASL	okra	4. instar	60	26	A	RNA extracted
28.3.	UgASL	okra	4. instar	96	26	A	
31.3.	UgASL	okra	4. instar	100	25	B	
11.4.	UgASL	okra	4. instar	70	36	B	
21.4.–5.5.	UgASL	okra	4. instar	100	22–36	E	
31.3.	UgASL	okra	adults	100	29	A	
7.4.	UgASL	okra	adults	100	36	A	
10.4.	UgASL	okra	adults	100	35	B	
15.5.	UgASL	okra	adults	100	46	E	
15.5.	UgASL	okra	adults	100	46	E	
19.4.	UgASL	pepper	eggs	200	1–6	G	
1.6.	UgASL	pepper	eggs	200	1–8	N	
27.6.	UgASL	pepper	eggs	320	5	T	
28.4.	UgASL	pepper	2. instar	100	11–15	G	
1.6.	UgASL	pepper	2. instar	48	1–8	N	
5.6.	UgASL	pepper	2. instar	51	12	N	
24.5.	UgASL	cotton	eggs	200	1–8	M	
24.5.	UgASL	cotton	eggs	200	1–8	M	
13.6.	UgASL	cotton	eggs	200	2–7	Q	
30.5.	UgASL	cotton	2. instar	100	15–18	M	
30.5.	UgASL	cotton	2. instar	100	15–18	M	
21.6.	UgASL	cotton	2. instar	200	15	Q	
28.6.	UgASL	cotton	2. instar	120	16–22	Q	
6.6.	UgASL	cotton	3. instar	100	21	M	
23.6.	UgASL	cotton	3. instar	120	30–39	M	
28.6.	UgASL	cotton	3. instar	100	16–22	Q	
23.6.	UgASL	cotton	4. instar	100	30–39	M	
23.6.	UgASL	cotton	4. instar	100	30–39	M	
23.6.	UgASL	cotton	adults	100	30–39	M	
23.6.	UgASL	cotton	adults	100	30–39	M	
23.6.	UgASL	cotton	adults	100	30–39	M	
29.3.	SpQ1	okra	eggs	85	6	C	
29.3.	SpQ1	okra	eggs	115	6	C	
7.4.	SpQ1	okra	eggs	100	7	F	

10.4.	SpQ1	okra	eggs	100	10	F	
15.6.	SpQ1	okra	eggs	200	8	R	
5.4.	SpQ1	okra	2. instar	100	13	C	
25.4.	SpQ1	okra	2. instar	100	25	F	
23.5.	SpQ1	okra	2. instar	100	11	J	
23.5.	SpQ1	okra	2. instar	120	11	J	
23.5.	SpQ1	okra	2. instar	120	11	J	
18.4.	SpQ1	okra	3. instar	100	26	C	
25.4.	SpQ1	okra	3. instar	100	25	F	
6.6.	SpQ1	okra	3. instar	100	25	K	
7.6.	SpQ1	okra	3. instar	100	26	J	
12.5.	SpQ1	okra	4. instar	49	42	F	
22.5.	SpQ1	okra	4. instar	37	51	F	
2.6.	SpQ1	okra	4. instar	9	62	F	
6.6.	SpQ1	okra	4. instar	75	25	K	
12.6.	SpQ1	okra	4. instar	30	31	K	
12.5.	SpQ1	okra	adults	100	42	F	
21.6.	SpQ1	okra	adults	100	40	K	
5.4.	SpQ1	pepper	eggs	100	7	D	
5.4.	SpQ1	pepper	eggs	110	7	D	
27.4.	SpQ1	pepper	eggs	200	1–6	H	RNA extracted
27.4.	SpQ1	pepper	eggs	200	1–6	H	
21.6.	SpQ1	pepper	eggs	200	8	S	
22.6.	SpQ1	pepper	eggs	200	9	S	
22.6.	SpQ1	pepper	eggs	230	9	S	
12.4.	SpQ1	pepper	2. instar	110	14	D	
4.5.	SpQ1	pepper	2. instar	100	13–18	H	
4.5.	SpQ1	pepper	2. instar	100	13–18	H	
17.5.	SpQ1	pepper	2. instar	100	20–26	H	
24.5.	SpQ1	pepper	2. instar	100	12	L	
5.6.	SpQ1*	pepper*	2. instar*	98*	24	L	RNA extracted
22.6.	SpQ1	pepper	2. instar	100	9	S	
22.6.	SpQ1	pepper	2. instar	100	9	S	
17.5.	SpQ1	pepper	3. instar	100	20–26	H	
23.5.	SpQ1	pepper	3. instar	110	26–32	H	
5.6.	SpQ1*	pepper*	3. + 4. instar*	75*	24	L	
23.5.	SpQ1	pepper	4. instar	120	26–32	H	
24.5.	SpQ1	pepper	4. instar	100	27–33	H	
4.5.	SpQ1	pepper	adults	70	36	D	
10.5.	SpQ1	pepper	adults	40	42	D	
23.5.	SpQ1	pepper	adults	100	26–32	H	RNA extracted
23.5.	SpQ1	pepper	adults	100	26–32	H	
23.5.	SpQ1	pepper	adults	100	26–32	H	
5.5.	SpQ1	cotton	eggs	220	3–7	I	
8.6.	SpQ1*	Cotton*	eggs*	200*	F_2 gen.	I	RNA extracted
11.5.	SpQ1	cotton	2. instar	110	9–13	I	
22.5.	SpQ1	cotton	3. instar	100	24	I	
25.5.	SpQ1	cotton	3. instar	120	27	I	
25.5.	SpQ1	cotton	4. instar	100	27	I	
2.6.	SpQ1	cotton	4. instar	100	35	I	
2.6.	SpQ1	cotton	adults	100	35	I	
2.6.	SpQ1	cotton	adults	100	35	I	

Appendix 22: Gel electrophoresis of denatured total RNA extracted from whiteflies in five different life stages (eggs, 2nd–4th instar and adults). 10 µl of eluate was loaded into each well. M = molecular marker. The well-defined two bands, the two-fold intensity of upper band compared with the lower one, and low amount of smearing indicate a good integrity of RNA.



www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

Received: 2 May 2018
Accepted: 5 July 2018
Published online: 18 July 2018

An integrative approach to discovering cryptic species within the *Bemisia tabaci* whitefly species complex

Soňa Vyskočilová[1], Wee Tek Tay[2], Sharon van Brunschot[1,3], Susan Seal[1] & John Colvin[1]

Bemisia tabaci **is a cryptic whitefly-species complex that includes some of the most damaging pests and plant-virus vectors of a diverse range of food and fibre crops worldwide. We combine experimental evidence of: (i) differences in reproductive compatibility, (ii) hybrid verification using a specific nuclear DNA marker and hybrid fertility confirmation and (iii) high-throughput sequencing-derived mitogenomes, to show that the "Mediterranean" (MED)** ***B. tabaci*** **comprises at least two distinct biological species; the globally invasive MED from the Mediterranean Basin and the "African silver-leafing" (ASL) from sub-Saharan Africa, which has no associated invasion records. We demonstrate that, contrary to its common name, the "ASL" does not induce squash silver-leafing symptoms and show that species delimitation based on the widely applied 3.5% partial mtCOI gene sequence divergence threshold produces discordant results, depending on the mtCOI region selected. Of the 292 published mtCOI sequences from MED/ASL groups, 158 (54%) are low quality and/or potential pseudogenes. We demonstrate fundamental deficiencies in delimiting cryptic** ***B. tabaci*** **species, based solely on partial sequences of a mitochondrial barcoding gene. We advocate an integrative approach to reveal the true species richness within cryptic species complexes, which is integral to the deployment of effective pest and disease management strategies.**

Accurate species identification underpins our understanding of biodiversity and enables clear scientific communication. In addition, if the organism is a pest having a negative impact on humans or the environment, it is essential for developing targeted and effective pest management strategies. Adequate species identification, however, is often difficult, especially when the organism belongs to a complex of cryptic species that cannot be distinguished by morphology alone. In such cases, it is logical to study populations of such organisms using the biological species concept. This concept defines a species as a group of individuals that can breed together (panmixis), but cannot with other groups, *i.e.* the group is isolated reproductively and genetically from other groups[1,2]. Evidence of both reproductive and genetic isolation is necessary for practical and accurate species identification while providing the opportunity to develop species-specific molecular markers. Ideally, an integrated approach involving the study of multiple factors (reproductive compatibility, morphology, ethology, ecology and molecular markers) is required to solve the problem[3,4].

A popular and widely used technique for species assignment is DNA barcoding[5]. The process involves sequencing of a standardised region in the organism's DNA and comparing it to an established 'DNA Barcode' database to enable species identity confirmation. For the Insecta, the commonly used region for DNA barcoding is a partial sequence of the mitochondrial cytochrome oxidase I (mtCOI) gene. Although this approach is attractive, simple and relatively fast, it has some fundamental problems, including: (i) the overlap between intra- and inter-species sequence divergences[6,7], (ii) the large number of misidentified and erroneous sequences in online databases[8–10], (iii) the arbitrary approach to selecting a 'preferred' barcoding gene region, *i.e.* either the 5′ region of mtCOI[5] or its 3′ region[11], and (iv) difficulties differentiating nuclear mitochondrial DNA segments (NUMTs) from genes in mitochondrial DNA (mtDNA)[12].

[1]Natural Resources Institute, University of Greenwich, Central Avenue, Chatham Maritime, ME4 4TB, United Kingdom. [2]CSIRO Black Mountain Laboratories, Clunies Ross Street, ACT 2601, Canberra, Australia. [3]School of Biological Sciences, The University of Queensland, St Lucia, Queensland, 4072, Australia. Correspondence and requests for materials should be addressed to S.V. (email: s.vyskocilova@gmail.com)

The practice of submitting DNA sequences generated for barcoding purposes to public data repositories such as GenBank has made them readily available for further analyses including phylogenetic analyses. That is the case for the *Bemisia tabaci* cryptic species complex of whiteflies (Hemiptera: Aleyrodidae), some of whose members are regarded to be amongst the worst invasive pests and plant-virus vectors[13,14]. A considerable effort, therefore, has already been invested in attempting to resolve the systematics of this species group. The former system of classifying *B. tabaci* populations into biotypes and host races based on various biological and biochemical markers has been superseded by molecular analyses based on the 657 bp partial sequence of the 3′ end of the mtCOI gene[11,15–17]. In addition, a genetic distance threshold of 3.5% was identified based on a gap in the distribution of pairwise sequence divergences amongst unique mtCOI partial sequences of *B. tabaci*[16]. In subsequent analyses, patterns of clusters of putative species were recognised which could be defined by sequence divergence equal to or higher than 3.5%. Mating experiments among some of these phylogenetic species have reported either complete or partial reproductive isolation[18–21].

At least 39 putative species have now been proposed in Bayesian and maximum likelihood phylogenetic analyses of the mtCOI sequenced from *B. tabaci* populations collected worldwide[7,16,22–29]. Of these, two species are especially of significant economic importance as highly invasive pests: Mediterranean (MED) and Middle East-Asia Minor 1 (MEAM1)[14]. These two species form the "Africa/Middle East/Asia Minor" clade together with the non-invasive Indian Ocean (IO) species, which was found to have asymmetric mating interactions with MEAM1 while showing 7–8% divergence at the mtCOI level[30]. Until recently, a Middle East-Asia Minor 2 (MEAM2) species also formed part of the clade, but its species status was based exclusively on pseudogene sequences originating from MEAM1 nuclear DNA[12]. A "MEAM2" population was recently reported from Uganda[31], however, this is a new putative species that was named on the basis of 97.1% similarity (*i.e.* within the 3.5% species delimitation boundary) to a published MEAM2 pseudogene sequence and should not be confused with the MEAM2 pseudogene clade[12].

The mtCOI partial sequence of MED corresponded to a syntype of *Aleurodes tabaci* originally collected by Gennadius in 1889, which indicates that MED represents the original *B. tabaci*[32]. MEAM1 was controversially named *B. argentifolii*[33] due to its ability to induce silver-leafing symptoms when feeding on squash[34]. However, this trait was later also observed in IO[35], and reportedly also in the "African silver-leafing" group of MED[14–16].

MED was proposed to be a single species after grouping multiple biotypes (Q, J, L and "sub-Saharan African silver-leafing"), based on the 3.5% species boundary threshold[16]. An increased availability of new mtCOI sequences, however, led to the distinction of two subclades that were called Q1 and Q2[36], or MedBasin1 and MedBasin2[37], followed by further subdivision into four subclades (Q1, Q2, Q3, and African silver-leafing (ASL)[38]) or five subclades called Q1–Q5[39], without an indication of how the group names corresponded to those previously published. Microsatellite analyses of sympatric field populations suggested that Q1 and Q2 were reproductively compatible[40–42], while evidence of gene flow was not detected between Q1 and ASL in the field[43], suggesting that the ASL population is a reproductively isolated entity. However, the model-based Bayesian methods for detecting hybrids that were implemented in the microsatellite field studies are biased towards the detection of F_1 hybrids[44]. To ascertain reproductive compatibility between populations it is important to test the fertility of F_1 hybrids.

In this study, we tested the hypothesis that the MED putative species (as described in[16,17,38,39]) constitutes a single biological species. We performed two-generational reciprocal crossing experiments in laboratory conditions associated with nuclear-gene based molecular markers to verify hybrid progeny. We then contrasted our findings with the mtCOI-based classification of MED and compared the sequence divergence among two different barcoding regions and 15 mitochondrial genes. Full mitogenomes were assembled from high-throughput sequencing (HTS) data, which enabled distinction between true mtDNA sequences and potential nuclear pseudogenes in the MED mtCOI datasets.

Results

F_1 and F_2 reciprocal crossing experiments. The lack of females in the F_1 generation from reciprocal crosses revealed that the ASL population from Uganda was incompatible with both Q1 from Spain and Q2 from Israel. Unlike the control crosses that produced 31.4–41.8% females in the progeny, the reciprocal crosses involving ASL resulted in male offspring only (Table 1). Reciprocal crosses between Q1 and Q2, however, resulted in female and male offspring in both directions of the cross. Mean counts of F_1 adults in these successful crosses, as well as proportions of females, were not significantly different from the controls (Table 1, Supplementary Table 1).

Fertility of the F_1 hybrid females resulting from crosses between Q1 and Q2 was further investigated by backcrossing them with males of both parental types. In all four combinations, female offspring were produced (Table 1). The F_2 females were present in 33 out of 35 replicates of all four combinations; the only two exceptions with purely male offspring occurred in the ♀ (♀Q2 × ♂Q1) × ♂Q2 cross.

There were, however, significant differences in the F_2 offspring counts, depending on the direction of the cross (Table 1, Supplementary Table 2). The offspring of (♀Q2 × ♂Q1) females backcrossed with Q2 males was 6–6.2 times smaller than the one produced by (♀Q1 × ♂Q2) females backcrossed to males of either parental type, which was a highly significant difference ($P < 0.001$). The offspring of (♀Q2 × ♂Q1) females crossed with Q1 males was 3.5–3.6 times smaller than that of (♀Q1 × ♂Q2) females backcrossed to males of either parental type, which was also significantly different ($P < 0.05$).

The proportion of females also differed between F_1 and F_2 generations and between the directions of the cross. In controls and F_1 reciprocal crosses of Q1 and Q2, the female percentage ranged from 36.1% to 45.2%. However, backcrossing the hybrid F_1 females resulted in female-biased progeny with 59.5–93.8% F_2 females. Within F_2 crosses, there was a stronger bias (albeit non-significant) toward female offspring produced by hybrid females with a Q2 maternal origin (83.8% and 93.8% of F_2 females) compared to their counterparts with a Q1 maternal background (59.5% and 65.3%).

Cross (1 × 3)	n	Mean no. of progeny	Mean no. of females	Mean % females
Controls				
♀Q1 × ♂Q1	9	$21.8 \pm 4.5^{ad,AC}$	9.1 ± 2.5	$41.8 \pm 8.2^{ab,AB}$
♀Q2 × ♂Q2	7	$22.6 \pm 5.3^{ad,AC}$	8.1 ± 2.6	$36.1 \pm 8.8^{ab,AB}$
♀ASL × ♂ASL	10	42.7 ± 8.2^{d}	13.4 ± 3.5	31.4 ± 5.2^{a}
Reciprocal (F_1)				
♀Q1 × ♂ASL	12	25.9 ± 4.6^{bd}	0.0 ± 0.0	0.0 ± 0.0
♀ASL × ♂Q1	5	22.6 ± 6.3^{ad}	0.0 ± 0.0	0.0 ± 0.0
♀Q2 × ♂ASL	13	17.4 ± 3.1^{abc}	0.0 ± 0.0	0.0 ± 0.0
♀ASL × ♂Q2	7	20.1 ± 4.8^{ad}	0.0 ± 0.0	0.0 ± 0.0
♀Q1 × ♂Q2	7	$36.6 \pm 8.4^{cd,BC}$	13.4 ± 4.2	$36.7 \pm 7.0^{ac,A}$
♀Q2 × ♂Q1	6	$32.8 \pm 8.2^{bd,BC}$	14.8 ± 4.9	$45.2 \pm 8.2^{ab,AB}$
Reciprocal (F_2)				
♀(♀Q1 × ♂Q2) × ♂Q1	12	$42.2 \pm 7.4^{d,C}$	25.1 ± 5.8	$59.5 \pm 5.1^{bc,AB}$
♀(♀Q1 × ♂Q2) × ♂Q2	11	$41.1 \pm 7.5^{d,C}$	26.8 ± 6.4	$65.3 \pm 5.2^{b,B}$
♀(♀Q2 × ♂Q1) × ♂Q1	7	$11.7 \pm 2.9^{ab,AB}$	9.6 ± 3.0	$83.8 \pm 9.6^{bc,AB}$
♀(♀Q2 × ♂Q1) × ♂Q2	5	$6.8 \pm 2.1^{a,A}$	6.0 ± 2.3	$93.8 \pm 9.9^{ab,AB}$

Table 1. Means and standard errors from reciprocal crossing experiments among the Spanish (Q1), Israeli (Q2) and Ugandan (ASL) populations of the MED putative species (*sensu* Dinsdale *et al.*[16]). Different superscript letters indicate statistically significant differences ($P < 0.05$) between crosses (Tukey's test). The lowercase superscript letters relate to the multiple comparison of all 13 types of crosses, whereas the capital letters relate to a separate comparison including only crosses between Q1 and Q2.

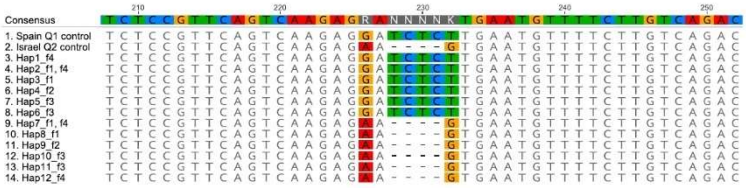


Figure 1. Molecular proof of parental origin of hybrid F_1 females. Spain Q1 control and Israel Q2 control show population-specific polymorphisms: 4 bp TCTC INDEL and SNPs at positions 226 (Q1: 'G', Q2: 'A') and 232 (Q1: 'T', Q2: 'G'). Unique haplotypes of the *GC1* marker (Hap1–Hap12) from four F_1 females produced in reciprocal crosses between ♀Q1 × ♂Q2 (f1, f2) and ♀Q2 × ♂Q1 (f3, f4) were aligned against *GC1* sequences from females produced in control crosses (Q1 control, Q2 control). The presence of both Spain Q1- and Israel Q2-specific sequences in all four females confirm that the F_1 females were hybrids between these two MED populations.

F_1 hybrid verification by nuclear *GC1* marker. The parental origin of F_1 females from the Q1 and Q2 cross was analysed by Sanger sequencing of the *GC1* nuclear DNA marker. Its sequence enabled the Q1 and Q2 populations to be distinguished based on a 4 bp insertion/deletion (INDEL) and two single nucleotide polymorphisms (SNPs) (Fig. 1). Analysis of trace files from direct Sanger sequencing of the *GC1* PCR product revealed heterozygous sequences in the F_1 females, compared to the homozygous ones from control individuals. The heterozygosity was demonstrated as mixed traces downstream from the INDEL site.

The mixed sequences of *GC1* marker from four F_1 females (f1 and f2 from ♀Q1 × ♂Q2; f3 and f4 from ♀Q2 × ♂Q1) were separated by molecular cloning and Sanger sequencing. The final sequence alignment of unique *GC1* haplotypes showed that each of the four F_1 females contained sequences specific for both Q1 and Q2 populations (Fig. 1), confirming their hybrid status.

Pseudogene identification. We applied four criteria in analysing published partial mtCOI sequences of five MED groups in order to detect potential pseudogene sequences. The presence of INDELs (criterion i) and significant amino acid changes (criterion iv) occurred in all five groups (Table 2). No premature stop codons (criterion ii) were detected, with the exception of clusters of non-synonymous mutations (criterion iii) found in Q1 and Q2 haplotypes.

Within the Q1 dataset, 67 (31.6%) of the 212 unique haplotypes failed criterion (i), out of which 12 overlapped with criterion (iii) (Supplementary Table 3). A particularly common INDEL was deletion of "A" at position 750 or 751 that occurred in 17 and 25 haplotypes, respectively. Criterion (iii) excluded 25 Q1 haplotypes, out of which two contained 40–41 SNPs within 757–769 bp (5.18–5.33% of the sequence), 11 contained 9–116 bp long stretches of non-synonymous mutations at the beginning or the end of the sequence (Supplementary Figure 1) and another 12 were identical to accessions from a different *Bemisia* species or chloroplast genome sequences in

MED group	Total haplotypes	(i) INDEL	(ii) STOP	(iii) Polymorphisms	(iv) AA change[a]	Potential NUMTs	
						Total	%
Q1[b,c]	212	67	0	25	47	116	54.7%
Q2[b,c]	37	9	0	4	10	20	54.1%
Q3[c]	2	0	0	0	1	1	50.0%
Q3[b] (=Q5[c])	8	1	0	0	3	4	50.0%
ASL[b] (=Q4[c])	33	7	0	0	10	17	51.5%

Table 2. Summary of analysis of partial mtCOI sequences from GenBank against HTS-derived mitogenomes from this study. Numbers of haplotypes failing each criterion are shown in columns. [a]AA = amino acid. Only sequences without INDELs were analysed at the amino acid level. Only amino acid changes in sites conserved across all reference *Bemisia* species were considered. [b]Naming system of Gueguen *et al.*[38]. [c]Naming system of Chu *et al.*[39].

BLAST search. Criterion (iv) failed 47 (22.2%) Q1 haplotypes based on the alignment with a reference set of COI protein sequences, which revealed amino acid changes in positions that are fully conserved across the *B. tabaci* complex and the *Bemisia* "JpI" species.

The Q2 dataset also included haplotypes with INDELs (9/37, 24.3%), failing criterion (i). Four Q2 haplotypes were detected by criterion (iii). Three of them contained clusters with high density of non-synonymous mutations (20–21 bp on 5′ end and 87–89 bp at 3′ end; Supplementary Figure 2), including multiple INDELs. The fourth haplotype contained a high number of SNPs throughout the sequence (40 SNPs in 836 bp, 4.78% of the sequence). In addition, ten haplotypes (27.0%) showed non-synonymous substitutions resulting in amino acid changes in positions that were conserved across the reference set of protein sequences, failing criterion (iv).

The two haplotypes from Croatia, called Q3 in Chu *et al.*[39], were 99.2–99.5% identical to the partial mtCOI of Israel Q2 and one of them contained a non-synonymous mutation in a conserved site. The group of haplotypes called Q3[38] or Q5[39], all from Burkina Faso, contained one haplotype with an INDEL and three haplotypes with one significant amino acid change. Finally, within the ASL[38] or Q4[39] group, 7/33 (21.2%) haplotypes contained INDELs and additional 10/33 (30.3%) haplotypes showed amino acid changes in conserved sites.

mtCOI nucleotide divergence within the Africa/Middle East/Asia Minor clade. A sliding window analysis was performed to investigate and compare the level of nucleotide divergence in a 657 bp window across the full length of mtCOI gene (1,532 bp). Levels of nucleotide diversity in the 5′ and 3′ barcoding regions were compared.

Pairwise comparisons among the putative species from Africa/Middle East/Asia Minor clade[16] showed that the distribution of sequence divergence is not uniform across the gene length (Fig. 2). In comparisons Q1 + Q2 vs. ASL and MED vs. IO, the nucleotide divergence was higher towards the 5′ end of the mtCOI gene compared to the 3′ end. In contrast, in MEAM1 vs. all other samples the divergence rose even higher at the 3′ end compared to the 5′ end. In the 3′ region, therefore, the observed genetic distances between MEAM1 and MED or IO are higher than in the 5′ region.

The divergence between ASL and Q1 + Q2 groups ranged from 2.48 to 4.77% (Supplementary Table 4), depending on the chosen region of mtCOI gene. Thus, the divergence between the ASL population and other MEDs was either 1.02% below or 1.27% above the species delimitation boundary of 3.5% for the barcoding region selected for the *B. tabaci* species complex[16]. In contrast, the divergence between Q1 and Q2 populations showed a smaller variation in sequence divergence (0.61–1.38% depending on the region of mtCOI).

Mitogenome divergence and phylogeny within members of *B. tabaci* species complex. The numbers of mapped Illumina HiSeq reads to the reference MED mitogenome JQ906700[45] were 513,612 (Spain Q1), 249,142 (Israel Q2) and 149,343 (Uganda ASL), with mean base coverage 4606.5 ± 3502.6, 2231.5 ± 1576, and 1351 ± 666 reads, respectively.

The nucleotide divergence of mitochondrial DNA was compared amongst our populations and eight other species/populations of the *B. tabaci* species complex for three regions: (i) the 3′ mtCOI barcoding region (Fig. 3a), (ii) the 5′ mtCOI barcoding region (Fig. 3b) and (iii) the concatenation of 13 protein-coding sequences and 2 rRNA genes in the mitogenome (Fig. 3c).

The divergence within Q1 populations from Spain, China and Burkina Faso and Q2 from Israel remained low in all three regions, ranging from 0.15 to 1.37%. However, the divergence between the ASL population from Uganda and MED populations was below 3.5% only at the 3′ mtCOI barcoding region (range 1.98–2.74%), while at the 5′ the divergence rose to 3.81–3.96%. The divergence among the Q1, Q2, and ASL populations across 15 mitochondrial genes was more closely reflected in the 3′ region for Q1 and Q2 populations, while for the ASL population it was the 5′ region.

The relationships between MED populations, Uganda ASL population and six other species of *B. tabaci* complex were reconstructed in a phylogenetic analysis based on the concatenations of 15 mitochondrial genes (Fig. 4). The topologies of the resulting trees were identical in all three cases with or without data partitioning, however, the branch support was the highest after partitioning the data into 15 individual genes. The substitution models used for each partition were TPM2u + F + G4 (*atp6*), K3Pu + F + I (*atp8*), HKY + F + G4 (*cox1* and *cox2*), K3Pu + F + I + G4 (*cox3*), TIM + F + I + G4 (*cytb*), TPM3 + F + I (*nd1* and *rrnL*), TPM2u + F + I + G4 (*nd2*),

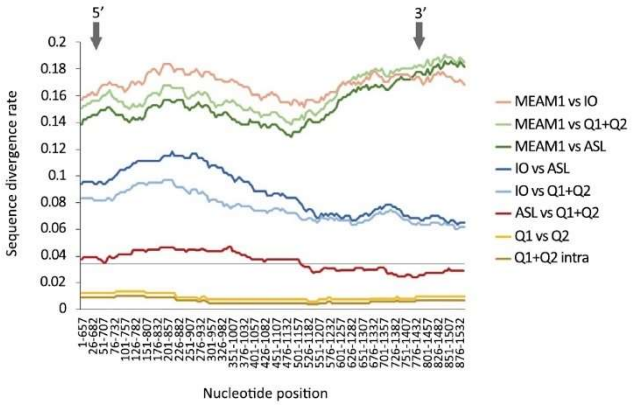


Figure 2. Pairwise sequence divergence across the full length of mtCOI gene (1,542 bp) from whiteflies of the Africa/Middle East/Asia Minor clade. The grey horizontal line marks the commonly applied 3.5% species boundary threshold[16]. The Q1 consists of two sequences (Spain Q1 and Burkina Faso Q1). Grey arrows indicate the position of 5′ and 3′ barcoding regions widely used to assist with species identification across taxa.

TPM3 + F + G4 (*nd3*), TIM3 + F + I + G4 (*nd4*), TPM3u + F + I (*nd4l*), TIM3 + F + I + G4 (*nd5*), TPM2 + F + I (*nd6*), HKY + F + G4 (*rrnS*). The UltraFast bootstrap values for all branches were 80–100%, which unlike the normal bootstrap corresponds roughly to 80–100% probability that the clades are true[46]. Uganda ASL was placed outside the Q1 and Q2 clade with 100% bootstrap support, between MED and MEAM1 species.

Biological and molecular assessment of the African silver-leafing population. The ability of the Ugandan "ASL" population to induce squash silver-leafing was tested in a bioassay with MEAM1 as a positive control. Squash plants infested with MEAM1 whiteflies developed silver-leafing symptoms within two weeks after whitefly infestation (Fig. 5a). The "ASL" population failed to induce the symptoms (Fig. 5b), even five weeks post-infestation.

The partial mtCOI sequences of our Uganda "ASL" population were compared with 27 published sequences of "sub-Saharan African silver-leafing"[15], "okra biotype"[47] and "Ug4"[48] (Supplementary Table 5). After checking the sequence quality against the four criteria, 14 haplotypes were discarded as potential pseudogenes or errors. The remaining 13 haplotypes shared 98.87–100% similarity with Uganda "ASL".

Discussion

We demonstrate an integrative approach to resolving the systematics of a group of cryptic species, using the Mediterranean clade of the *B. tabaci* species complex as an example. The laboratory crossing experiments confirmed and expanded on the evidence of reproductive isolation between the sub-Saharan "ASL" population and the Q1 and Q2 MED whiteflies from the Mediterranean Basin[43]. Our backcrossing experiments confirmed the viability and fertility of Q1 × Q2 hybrids. A comparison of GenBank accessions of partial mtCOI sequences from different MED subclades with HTS-derived full mtCOI genes unravelled the scale of potential pseudogenes and/or low-quality sequences included in previous phylogenetic analyses. We also show that the "ASL" group is genetically distinct from the Q1 and Q2 MED groups, which are more closely related to one another. We also demonstrate that "ASL" does not induce silver-leafing in squash and so should be renamed.

The reproductive compatibility between Q1 and Q2 was previously suggested based on field population genetics studies[40–42]. Our study expanded on this evidence by laboratory crossing experiments in which the progeny counts and sex ratios could be observed and compared, as well as the viability and fertility of the hybrid progeny. In all four types of backcrossing, the F_1 hybrids of Spain Q1 and Israel Q2 were fertile, regardless of which population was the source of maternal and paternal genetic background. In addition, for the first time, the parental origin of hybrid F_1 females was verified by a molecular test based on a specifically selected nuclear marker.

Sequences of the cloned inserts of the *GC1* marker showed that the number of unique haplotypes present in three out of the four F_1 females (f1, f3 and f4) tested, exceeded the theoretically possible number of alleles per individual. For example, two Q1-specific (Hap2 and Hap3) and two Q2-specific haplotypes (Hap7 and Hap8) were detected in the hybrid female f1, while in reality she would have received only one allele from each parent. The 12 *GC1* haplotypes differed from one another by 1–4 SNPs that occurred outside the diagnostic region shown in Fig. 1. These SNPs were probably errors introduced during the PCR amplification step by DreamTaq polymerase. The error rate of *Taq* DNA polymerases is $1–20 \times 10^{-5}$ [49–51], which is 6- to 50-fold higher than more precise DNA polymerases such as *Pfu* or Phusion Hot Start[52]. Despite these introduced errors, the cloned sequences proved unequivocally that each hybrid female contained a copy from both parental types as shown by the INDEL region. This method of checking for genuine hybrid female progeny could be improved by using a high-fidelity DNA polymerase and a reduced number of PCR cycles for amplification of the marker sequence prior to cloning.

a: 3' mtCOI

		1	2	3	4	5	6	7	8	9	10	11
1	Asia I	-										
2	Australia I	10.81	-									
3	*B. emiliae* (Asia II-7)	12.33	13.85	-								
4	New World	14.16	16.59	16.13	-							
5	Indian Ocean	13.24	15.53	14.92	15.68	-						
6	MEAM1	14.16	14.31	15.22	14.46	7.00	-					
7	Uganda ASL*	14.61	15.83	14.31	15.22	6.54	4.72	-				
8	Israel MED Q2*	14.00	15.98	14.76	14.61	6.09	4.41	1.98	-			
9	Burkina MED Q1	14.76	16.44	14.92	14.76	6.24	5.02	2.74	1.07	-		
10	China MED Q1	14.76	16.44	15.22	15.07	6.54	5.02	2.74	1.07	0.30	-	
11	Spain MED Q1*	14.61	16.29	15.07	14.92	6.39	4.87	2.59	0.91	0.15	0.15	-

b: 5' mtCOI

		1	2	3	4	5	6	7	8	9	10	11
1	Asia I	-										
2	Australia I	12.48	-									
3	*B. emiliae* (Asia II-7)	13.24	14.46	-								
4	New World	14.61	15.53	15.22	-							
5	Indian Ocean	15.07	14.46	15.53	14.76	-						
6	MEAM1	14.46	14.31	15.53	14.00	9.44	-					
7	Uganda ASL*	13.39	13.09	14.61	14.92	8.98	8.07	-				
8	Israel MED Q2*	14.00	13.85	14.31	13.85	8.07	8.22	3.81	-			
9	Burkina MED Q1	14.16	14.00	15.07	13.85	7.76	7.91	3.81	1.22	-		
10	China MED Q1	14.00	14.00	15.22	13.70	8.22	8.07	3.96	1.37	0.46	-	
11	Spain MED Q1*	14.16	14.00	15.07	13.55	7.76	7.61	3.81	1.22	0.30	0.46	-

c: Concatenations

		1	2	3	4	5	6	7	8	9	10	11
1	Asia I	-										
2	Australia I	14.00	-									
3	*B. emiliae* (Asia II-7)	16.45	16.72	-								
4	New World	19.96	20.3	19.88	-							
5	Indian Ocean	19.83	20.37	19.75	19.67	-						
6	MEAM1	19.77	19.91	19.47	19.33	8.74	-					
7	Uganda ASL*	19.72	20.23	19.36	19.34	8.62	7.73	-				
8	Israel MED Q2*	19.84	20.38	19.57	19.44	8.80	7.79	3.91	-			
9	Burkina MED Q1	19.98	20.43	19.66	19.34	8.62	7.88	3.91	1.09	-		
10	China MED Q1	20.02	20.44	19.72	19.40	8.71	7.95	3.94	1.07	0.37	-	
11	Spain MED Q1*	19.97	20.39	19.65	19.36	8.66	7.91	3.93	1.06	0.35	0.15	-

Figure 3. mtDNA sequence divergences amongst 11 species and populations of the *B. tabaci* complex. The regions compared were (**a**) the 3′ barcoding region of mtCOI used by the whitefly community (657 bp), (**b**) 5′ barcoding region of mtCOI used in the Barcode of Life (657 bp) and (**c**) the concatenations of 13 protein coding sequences and two rRNA genes in mtDNA (12,595 bp). The dark grey boxes highlight the divergences between Uganda ASL and MED populations, while divergences within Q1 and Q2 groups of MED are highlighted with light grey. Names with asterisks mark samples generated in this study.

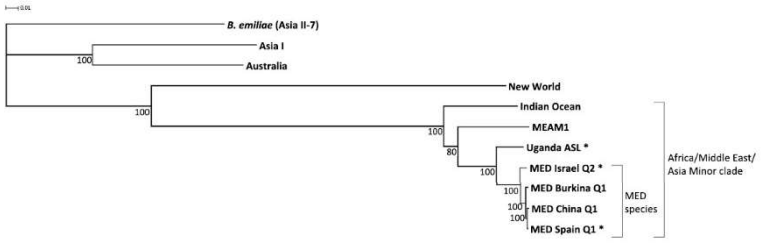


Figure 4. Unrooted maximum-likelihood phylogenetic tree inferred from 12,595 bp concatenations of 15 mitochondrial genes. The values below nodes show statistical support of Ultrafast Bootstrap. Names with asterisks mark samples generated in this study.

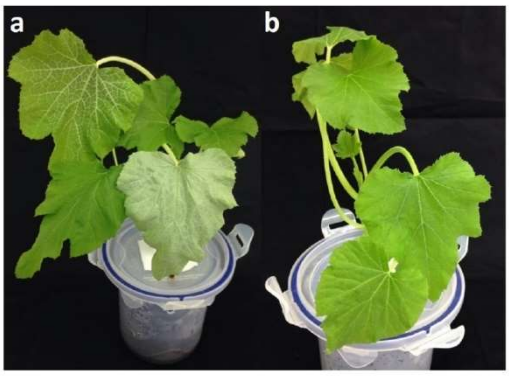


Figure 5. Squash plants (*Cucurbita pepo* 'All green bush') after feeding by: (**a**) MEAM1, and (**b**) Uganda "ASL". Photographs were taken two weeks after infestation. The plant infested with MEAM1 showed leaf silvering symptoms, while the plant fed on by "ASL" population from Uganda did not show the symptoms. (Photograph by S. Vyskocilova).

Alternatively, the hybrids could be detected using microsatellite markers[40–42,53] or evidence of hybridisation between species could also be demonstrated by using genome-wide SNPs[54].

From the biological species concept definition, the Q1 and Q2 subclades belong to the same species. As the mtCOI of the original specimen of *B. tabaci* (Gennadius) from Greece corresponds to MED Q1[32], we conclude that the Q1 and Q2 both belong to *B. tabaci sensu stricto*. These two groups are differentiated in allopatry, but retained the ability to interbreed when occurring sympatrically in the field. The gene flow between Q1 and Q2 MED populations was suggested to be asymmetric, because the nuclear background associated with Q1 individuals introgressed into, or completely replaced, the nuclear background of individuals with Q2 mitochondrial type, but not *vice versa*[41]. However, a different population genetic study based on microsatellites showed evidence for both directions of hybridisation[55]. Samples from France and Canary Islands had "Western Mediterranean" nuclear background and Q2 mitochondrial type, while sample from Greece had Q1-type mitochondria and "Eastern Mediterranean" DNA introgressed into its nuclear genome[55]. The interaction also did not appear asymmetric in Terraz *et al.*[40], as the sympatric populations in France and Spain had homogeneous nuclear backgrounds regardless of the mitochondrial haplotype. In our study, no significant differences were observed between F_1 progeny from the two directions of the cross. In the F_2 generation, however, the fertility was unequal between hybrid females with different parental origin. Albeit the nuclear DNA of our samples was not studied, F_2 females with Q1 mitochondria were more fecund than their Q2 counterparts, which could make hybrids with Q1 mitochondrial type more prevalent and thus more likely to be detected in the field.

We hypothesise that these mating interactions are influenced by secondary endosymbiotic bacteria. Some endosymbionts residing within insects are known to induce cytoplasmic incompatibility[56,57] or affect the host's fitness[58,59]. Cytoplasmic incompatibility could explain the mechanism behind the lack of gene flow between Q1 females and Q2 males in Gauthier *et al.*[41]. Alternatively, the asymmetry could be caused by the difference in fecundity between hybrids which emerged from the opposite directions of the same cross. The effect of endosymbiotic bacteria on the reproductive compatibility and fitness of MED whiteflies remains to be established.

The analysis of 292 unique mtCOI haplotypes deposited in GenBank as assigned to MED subclades Q1–Q5 and ASL[36–39], revealed the extent of NUMTs or erroneous sequences included in previous molecular studies. In all five subclades, a half or more (50–55%) of the unique haplotypes contained mutations that are unlikely to occur in a functional protein-coding sequence. It is important to note that the FaBox tool assigned identical sequences with different lengths as unique haplotypes. Thus, the percentages in Table 2 probably do not reflect the frequency with which pseudogenes are picked up by the primers. The percentages are overestimated or underestimated, depending on whether it was pseudogenes/erroneous sequences, or mtDNA sequences predominantly deposited with unequal lengths.

The inclusion of NUMT sequences in phylogenetic analyses can overestimate the species richness[60] and confound phylogenetic analyses[61,62], as well as create confusion and inconclusive results in the research community. Pseudogene sequences have recently been shown to cause an overestimation of species richness within the Africa/Middle East/Asia Minor clade (the case of MEAM2[12]). This indicates that unrecognised NUMT sequences might commonly occur across multiple clades of the *B. tabaci* species complex and interfere with efforts to resolve its systematics. It is therefore important to recognise this issue and to analyse critically the quality of putative mitochondrial sequences amplified with universal primers. The pseudogene sequences are not immediately obvious if they do not contain stop codons or significant amino-acid changes within the analysed partial sequence. We emphasise the utility of HTS and acquiring mitochondrial sequences by assembly from a large number of reads. Such sequences can be then used as a reference for all the Sanger-sequenced marker sequences (e.g. from a field collection), to identify sequencing errors and potential pseudogenes and exclude them from downstream analyses.

The sub-Saharan ASL population is reproductively isolated from Q1 and Q2 groups as demonstrated by our laboratory crossing experiments and the field microsatellite study in Burkina Faso, where Q1 and ASL *B. tabaci* occurred in sympatry on the same host plants[43]. In addition, our phylogenetic tree shows the genetic distinctness of the ASL population from the MED Q1 and Q2 populations with a high statistical support. The tree also supports the consensus that Indian Ocean occupies a basal phylogenetic position within the Africa/Middle East/Asia Minor clade[7,14,16].

The "Sub-Saharan Africa silver-leafing" clade was originally placed separately from the "Mediterranean invasive" clade in the first global Bayesian phylogenetic study by Boykin *et al.*[15]. These two clades, however, were later merged to form one "Mediterranean" putative species[16]. This placement was in accordance with the 3.5% species-level boundary in the 3′ barcoding region of mtCOI. We show here, however, that the identity values vary considerably across the full length of mtCOI gene. Results from the sliding window analysis demonstrate that the sequence divergence between ASL and Q1 + Q2 populations in a 657 bp long window ranged from 2.48 to 4.77% (a difference of 2.29%), depending on the region chosen for barcoding. Such variability is not apparent between Q1 and Q2, divergence of which remains low across the full gene length (0.61 to 1.38%, difference of 0.77%). A similar pattern was observed in the comparison at the mitogenome scale. Interestingly, both mtCOI barcoding regions quite accurately reflected the divergence across 15 mitochondrial genes, but each for a different group (3′ for Q1 and Q2; 5′ for ASL).

Our study also shows that the "African silver-leafing" population from Uganda does not induce silver-leafing. None of the plants infested with the ASL population developed silver-leafing symptoms five weeks after whitefly infestation, while the positive control with MEAM1 in our study and the one of Yokomi *et al.*[63] showed symptoms within two weeks. This is not surprising, as the original paper by Sseruwagi *et al.*[48] clearly stated that only groups Ug6 and Ug7 (corresponding to MEAM1 and IO, respectively) induced squash silver-leafing, while Ug4 (similar to "okra biotype" from Ivory Coast) did not. The "silver-leafing" group of MED was linked to the former J biotype[14] reported in Nigeria, Ghana, Cameroon, Ivory Coast and Zimbabwe[64]. However, De la Rúa *et al.*[64] did not study the capacity to induce silver-leafing and in earlier studies the J biotype failed to induce this phytotoxic disorder[65,66]. Despite all this evidence, multiple subsequent studies about African silver-leafing populations of MED referred either to Sseruwagi *et al.*[48] or De la Rúa *et al.*[64], or to other papers that have not demonstrated the silver-leafing capacity[14–16]. It is possible that the results of Sseruwagi *et al.*[48] were misinterpreted due to the mixed "Ug4/Ug6" population reported to induce silver-leafing, and/or a confusion caused by the naming system in the study, and this misinterpretation was then cited in subsequent literature.

We propose that our "ASL" population from Uganda belongs to a non-MED species within the Africa/Middle East/Asia Minor clade. The 98.87–100% identity among 3′ mtCOI sequence across samples of "African silver-leafing"[15], "okra biotype"[47,67] and "Ug4"[48] indicates that these groups might belong to the same species, but it remains to be confirmed. A list of these potential synonyms of the "ASL" species from previous literature is summarised in Supplementary Table 6. Similarly, the "Q3" population from Croatia might actually be part of Q2 together with our population from Israel based on their 99.2–99.5% identity. A mating study and ideally also a genome-wide SNP analysis would be needed to confirm these two results.

It is possible that "ASL" has already been described and was later synonymised with *B. tabaci* (Russell 1957). The synonymised species collected in sub-Saharan Africa include *B. gossypiperda var. mosaicivectura* Ghesquière (Congo, 1934), *B. vayssierei* Frappa (Madagascar, 1939), *B. goldingi* Corbett (Nigeria, 1935), *B. nigeriensis* Corbett (Nigeria, 1935) and *B. rhodesiaensis* Corbett (Rhodesia = Zimbabwe, 1936)[68]. DNA sequencing from museum syntypes of these species will be attempted in the near future to establish whether any of the above names apply to ASL/okra/Ug4. If not, a formal description of this new species will be required to accompany a new species name.

In conclusion, we gathered multiple lines of evidence that the MED species (*sensu* Dinsdale *et al.*[16]) comprises at least two biological species. One is the original *B. tabaci* (Gennadius), which includes the MED populations Q1 and Q2 from the Mediterranean Basin; the other is the sub-Saharan "MED ASL" species, awaiting a new binomial name. This division is based on the biological species concept, as "ASL" and MED are reproductively isolated both in the laboratory and in the field[43]. The reproductive compatibility of Q1 and Q2 was demonstrated in both directions, the fertility of F_1 hybrids was verified and for the first time the combined genetic material from both parents in the F_1 hybrids was confirmed using a specific nuclear DNA marker. Furthermore, by using HTS-derived mitogenomes, we were able to compare the genetic divergence among the MED populations and seven other *Bemisia* species, using true mitochondrial sequences (*i.e.* no pseudogenes) and multiple genes. This approach contrasts previous studies inferring phylogenetic relationships and sequence divergence among MED populations from 3′ partial mtCOI sequences, out of which 54% in the public databases contain errors and potentially are NUMTs.

The previous inclusion of the "ASL" population in the MED species demonstrates a fundamental problem with species delimitation based exclusively on an arbitrarily chosen barcoding sequence in mitochondrial DNA, because the results varied depending on the selected region of comparison. Furthermore, we showed a case of serial miscitation that led to the perpetuation of the erroneous assignment of the diagnostic trait of squash silver-leafing to the "ASL" species. Overall, we advocate the importance of an integrative approach to cryptic species identification and delimitation, most importantly including the biological species concept and using multiple genes or genomic data to infer *B. tabaci* phylogeny. With the increasing ease with which genome-wide SNPs can be obtained, combining nucleotide polymorphisms from both mitochondrial DNA and nuclear DNA genomes (e.g.[54]) will offer greater power for inference of *B. tabaci* cryptic species phylogenetic relationships. We also suggest that understanding the species diversity in pest-species complexes is the first step towards designing novel and targeted management strategies. It shall also bring consistency and clarity to communication within the scientific community, as well as to the wider public.

Methods

Insect rearing. *B. tabaci* colonies were reared on aubergine (*Solanum melongena* 'Black Beauty'). Plants were grown from seeds in a whitefly-free room at 28 ± 2 °C, 50–60% relative humidity (r.h.) and a 14:10 L:D photoperiod. Whitefly colonies were maintained in rectangular 45 × 44 × 44 cm cages (BugDorm, US) at 28 ± 2°C, 30% r.h. and 14:10 L:D. Colonies were established from field populations collected in Spain (in 2013 from *Cucurbita* sp.; MED, Q1 group), Israel (in 2003 from *Gossypium hirsutum*; MED, Q2 group) and Uganda (in 2013 from *Abelmoschus esculentus*; MED, ASL group). The purity of core whitefly colonies was monitored periodically by sequencing of the partial mtCOI gene (detailed below).

DNA extraction, molecular species identification and endosymbiont screening. Genomic DNA was extracted from adult whiteflies stored in 90% ethanol using the Chelex method[69]. At least five individual adults from each core colony were used for the species identity verification. Whiteflies were homogenised individually using sterile pestles in 50 µl of 10% Chelex® 100 Resin solution (Bio-Rad). The mixture was incubated at 56 °C for 20 min, subsequently at 95 °C for 5 min and centrifuged for 5 min at 13,500 rcf. The supernatant was used as a template for amplification by polymerase chain reaction (PCR).

Whitefly identification was based on the 3′ barcoding region of mtCOI gene. The segment was amplified and sequenced with primers 2195Bt and COI2-BtSh2[31]. PCRs (at 52 °C annealing temperature) were carried out using DreamTaq DNA polymerase (Thermo Scientific, UK) following the manufacturer's instructions. The 867 bp amplicons were purified with reSource (Source Bioscience, UK) or GeneJET (Thermo Scientific, UK) PCR purification kits prior to Sanger sequencing by Source Bioscience (Nottingham, UK).

All DNA sequence analyses were carried out in Geneious version 10.0.8[70], unless stated otherwise. The mtCOI sequences were trimmed to 657 bp, corresponding to nucleotide positions 782–1,439 of the complete mtCOI gene of MED[45]. Sequence-based comparisons were done using two reference datasets: (i) consensus sequences for 24 putative species of *B. tabaci*[16] to verify MED species, and (ii) MED mtCOI haplotypes[38] to place our populations in a recognised naming framework. In addition, the partial mtCOI sequences of Uganda ASL sample were compared to published sequences of "African silver-leafing"[15], "okra biotype"[47,67] and "Ug4"[48] groups.

The presence/absence of the primary bacterial endosymbiont (*Portiera*) and five secondary endosymbionts (*Arsenophonus*, *Cardinium*, *Hamiltonella*, *Rickettsia* and *Wolbachia*) was tested by conventional PCR. Genus-specific primers targeting the 23 rDNA (*Hamiltonella*) or 16 S rDNA genes (the remaining bacteria) from Ghosh *et al.*[71] were used. Total genomic DNA (extracted as above) from ten individual females per each colony was used as a template. Positive and negative controls were included in screening for each endosymbiont. PCR products were visualised by agarose gel electrophoresis, with results summarised in Supplementary Table 7.

F_1 and F_2 reciprocal crossing experiments. Reproductive compatibility among the Q1, Q2 and ASL whitefly populations was determined by reciprocal crossing experiments. *B. tabaci* species produce males and females from unfertilised and fertilised eggs, respectively[72]. Gene flow between two populations, therefore, can be shown by the presence of female offspring.

All crosses were carried out using newly emerged virgin adults, isolated by excision as 4th instar nymphs from a leaf and placed individually into glass tubes. Emerged adults were sexed visually on daily basis using a binocular stereomicroscope prior to mating experiments.

Control crosses were carried out with one female and three males from the same population with 8.7 replications on average. For the reciprocal crosses, the female and males originated from different populations and were used (average n = 8.3). The four adults were released onto aubergine plants in three to seven true-leaf stage, rooted in soil and enclosed in Lock&Lock whitefly-proof cages[73] with additional side openings in the upper container covered by 160 µm nylon mesh. At least 24 h before introducing the adults, all leaves, except a fully expanded one, were removed to facilitate the contact between mating partners. Survival of the parental adults was monitored periodically and deceased males were replaced by new males from the respective colony. In the case of female death, the replicate was discarded. The parents were collected after seven days and stored in 90% ethanol at −20°C. All emerged adults of the F_1 generation were collected and sexed. F_1 females were stored for subsequent molecular analysis to verify their genetic make-up (detailed below).

The fertility of F_1 hybrids was tested in a separate set of reciprocal crossing experiments. A female and three males per replicate (average n = 8.8) were used as above, but the F_1 hybrid female nymphs were harvested from a plant onto which 20 + 20 virgin females and males from populations of interest were given the opportunity to reproduce for 14 days. Virgin hybrid females from both directions of the cross were then backcrossed with males from either parental population, resulting in four types of crosses to produce F_2 generation. The parents were collected after seven days and progeny collected and sexed until all emerged.

Statistical analyses were performed using R[74]. Counts of offspring generated in crossing experiments were analysed by a generalised linear model with negative binomial error distribution and a log link using the MASS library[75]. For the proportion of female progeny, a generalized linear model with quasibinomial error distribution and logit link of the proportional data was used. Multiple comparisons of offspring counts and female proportions were performed by Tukey's test[76] using the multcomp package[77] and significant differences were demonstrated by compact letter display. Reciprocal crosses that produced only male offspring were omitted from the multiple comparison. A separate analysis was also carried out with results from the F_1 and F_2 generations relevant only to the reproductively compatible populations.

Molecular verification of F_1 hybrid females. The parentage of female progeny was studied using a nuclear DNA marker. Mitochondrial DNA marker was not used due to its maternal mode of inheritance, rendering it ineffective for discerning the paternal origin of F_1 females. We targeted nuclear genes encoding mitochondrial proteins, because of their potentially faster mutation rate[78]. Candidate genes were extracted from

transcriptome sequences of adults from Spain MED Q1 and Israel MED Q2 populations (unpublished data) based on *Drosophila melanogaster* homologues (www.flybase.org) and searched for variable regions. The marker chosen for this study was a segment of 3′ untranslated region of the predicted Glutamate Carrier 1 (*GC1*) gene (annotated as Bta11593 in the genome of MEAM1[79]; 97% identity). Primers were developed in Primer 3 software[80] (*GC1* primers; FWD3 5′-TGTTTTGTATTTTGATCTATTCA-3′, REV3 5′-CGAGGAGGAAATGTAAACAA-3′, annealing temperature 52 °C; expected amplicon size 770 bp).

The amplicon was sequenced from representative F_1 females produced in reciprocal and control crosses. Trace files were analysed using the Pregap and Gap4 programs within the Staden molecular analysis software[81]. The *GC1* amplicons from F_1 females were cloned into pGEM®-T Vector System I (Promega) with a 2:1 insert:vector ratio. Colonies of transformant *E. coli* cells JM109 were selected randomly and individually lysed in 200 µl sterile water by brief vortexing, then used as a template for PCR with vector-specific T7 and SP6 primers. Selected amplified inserts were sequenced at the ACRF Biomolecular Resource Facility at the Australian National University in Canberra and analysed as described above. The insert sequences were collapsed into unique haplotypes in FaBox DNA Collapser version 1.41[82] and mapped to reference sequences from control crosses.

High-throughput genome sequencing, quality control and mitogenome assembly. Genomic DNA from single male whiteflies representing each pure, inbred laboratory colony, was isolated using the silica spin-column method[83] and quantified using a Qubit® 2.0 fluorometer (Invitrogen). Genomic DNA (_30 ng in 100 µl of TE buffer) was sheared for 10 min using an ultrasonic cleaner (VGT-1620QTD) and size-selected (300–500 bp) using Blue pippin (Sage Science). Sequencing libraries were prepared separately using each size-selected DNA pool using the NEBNext® Ultra DNA Library Prep Kit for Illumina® with NEBNext Multiplex Oligos for Illumina (New England BioLabs). Individual libraries were checked by capillary microchip electrophoresis (MultiNA, Shimadzu), pooled in equimolar amounts, and then purified using Agencourt AMPure XP Beads (Beckman Coulter). The multiplexed library was sequenced on a single lane of an Illumina HiSeq. 4000 platform (Novogene Bioinformatics Institute, Beijing, China) and 150 bp paired-end reads were generated.

Quality control was carried out in FastQC version 0.11.5 before and after trimming the reads with Skewer 0.2.2 version for Linux[84]. The reads trimmed with a quality threshold of 40 and minimum length of 18 bp were used in subsequent analyses. Draft mitogenomes of the Q1, Q2 and ASL males were assembled by iterative reference-guided assembly. The trimmed reads were mapped to the reference MED mitogenome JQ906700[45] using Geneious 10.0.8 mapper set to medium-low sensitivity with up to 5 iterations. The draft mitogenomes were annotated in MITOS2[84] and manually checked for start and stop codons in all protein-coding genes by translating into amino acid sequences using the invertebrate mtDNA genetic code. As the sequence was assembled from short reads, the accuracy of the low complexity intergenic region between *cox3* and *tRNA-Ile* represents only an estimate. Nevertheless, the lengths of these AT-rich regions among our mitogenomes (973 bp in Q1 and Q2, 971 bp in ASL) were similar to the one in the reference MED mitogenome JQ906700 (972 bp), which was obtained by long PCR and Sanger sequencing[45].

Pseudogene identification. Published partial mtCOI sequences of MED groups Q1, Q2, Q3, Q4, Q5 and ASL[36–39] were downloaded from GenBank (accessed 14th November 2017) and collapsed into unique haplotypes in FaBox (Supplementary Table 3). The haplotypes were then mapped to the full mtCOI genes extracted from the mitogenomes of Q1, Q2 and ASL from this study.

The criteria used to categorise a partial mtCOI sequence as a potential pseudogene included either: (i) the presence of INDELs, (ii) the presence of premature stop codons inside the protein-coding sequence, (iii) anomalous polymorphisms (*i.e.* clusters of non-synonymous mutations or haplotypes with an outlying sequence divergence compared to the intra-group divergence, and/or (iv) non-synonymous mutations resulting in amino-acid substitutions in positions conserved across other *Bemisia* species. For the fourth criterion, only sequences that passed the first two criteria were analysed. Such DNA sequences were translated and aligned by Geneious alignment tool using Blosum62 matrix to a reference set of mtCOI amino acid sequences translated from 12 mitogenomes: (1–3) populations from this study, (4) China MED Q1 (JQ906700)[45], (5) Burkina Faso MED Q1 (KY951447), (6) Peru MEAM1 (KY951450), (7) Australia I (KY951541), (8) Indian Ocean (KY951448)[12], (9) *B. emiliae* (former Asia II-7; KX714967), (10) *Bemisia* "JpL" (KX714968)[85], (11) Asia I (KJ778614)[86] and (12) New World (AY521259)[87]. The likelihood of the non-synonymous substitutions occurring in MED haplotypes was estimated by comparing the intra- and inter-species variability in respective triplet positions. Sequences with high nucleotide divergence from the group were compared against the NCBI database using BLASTn[88] to investigate the origin of the sequence.

Sliding window analysis. Intra- and inter-specific sequence divergence levels within the Africa/Middle East/Asia Minor clade[16], excluding MEAM2 as a pseudogene artefact[12], were compared across the full length of mtCOI gene (1,542 bp) in a sliding window analysis in DnaSP version 5.10.01[89]. The size of commonly used partial mtCOI sequence (657 bp) was used for the window size, sliding along the sequence in 5 bp steps while plotting the divergence (K-JC total) between the compared samples. Sequences included in the sliding window analysis were the three populations from this study, Burkina Faso MED Q1 (KY951447), Peru MEAM1 (KY951450) and Indian Ocean (KY951448)[12].

Mitogenome phylogeny and divergence. Concatenations of 13 protein-coding genes and two rRNA genes were generated from 11 mitogenomes listed above, excluding *Bemisia* "JpL" due to incomplete sequence data. The genes were orientated in 5′−3′ direction, individually aligned by MUSCLE alignment tool in Geneious and trimmed to the same length prior to being manually concatenated into 12,595 bp sequences. The multiple sequence alignment in fasta format was submitted to IQ-tree web server[90] along with a manually created nexus

file partitioning the data[91]. Three different partition schemes were analysed: (i) no data partitioning, (ii) partitioning the protein-coding genes into $1^{st} + 2^{nd}$ codon position, 3^{rd} codon position, and rRNA genes grouped in the third partition; and (iii) partitioning into 15 individual genes. No substitution saturation in the protein coding sequences was detected using DAMBE6[92], indicating that all three codon positions could be used in the phylogenetic analysis. The best-fitting substitution model for each partition was identified using ModelFinder[93]. Ultrafast Bootstrap[94] with 1,000 replications was performed and the tree was visualised in Dendroscope 3[95].

Squash silver-leafing bioassay. The capacity of Uganda ASL to induce the squash silver-leafing symptoms was tested by releasing 30–60 adults on squash plants *Cucurbita pepo* 'All Green Bush', individually enclosed in Lock&Lock cages in five replications. As a positive control, two replicates with 60 MEAM1 adults from Peru were set up using the same protocol. After two weeks the adults were collected and plants were visually assessed for leaf silvering symptoms, for up to five weeks post-infestation.

Data availability. Draft mitogenomes assembled in this study are available under GenBank accession numbers MH205752 (Spain Q1), MH205753 (Israel Q2) and MH205754 (Uganda "ASL"). The partial 3′ mtCOI gene sequences (657 bp) of the whitefly colonies used in this study correspond to positions 782–1,439 of the draft mitogenomes.

Haplotypes of the *GC1* marker acquired from molecular cloning can be found under accessions MH205738–MH205749. Accessions for the *GC1* marker from Spain Q1 and Israel Q2 controls are MH205750 and MH205751, respectively.

References

1. Mayr, E. *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*. (Harvard University Press, 1942).
2. Mayr, E. *Principles of Systematic Zoology*. (McGraw-Hill, 1969).
3. Smith, M. A. *et al*. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proc. Natl. Acad. Sci. USA* **105**, 12359–12364 (2008).
4. Leaché, A. D. *et al*. Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). *Proc. Natl. Acad. Sci. USA* **106**, 12418–12423 (2009).
5. Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. Biological identifications through DNA barcodes. *Proc R Soc Lond B* **270**, 313–321 (2003).
6. Meier, R., Shiyang, K., Vaidya, G. & Ng, P. K. L. DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. *Syst. Biol.* **55**, 715–728 (2006).
7. Lee, W., Park, J., Lee, G. S., Lee, S. & Akimoto, S. Taxonomic status of the *Bemisia tabaci* complex (Hemiptera: Aleyrodidae) and reassessment of the number of its constituent species. *PLoS One* **8**, e63817 (2013).
8. Vilgalys, R. Taxonomic misidentification in public DNA databases. *New Phytol.* **160**, 4–5 (2003).
9. Harris, D. J. Can you bank on GenBank? *Trends Ecol. Evol.* **18**, 317–319 (2003).
10. Balakirev, E. S., Saveliev, P. A. & Ayala, F. J. Complete mitochondrial genomes of the Cherskii's sculpin *Cottus czerskii* and Siberian taimen *Hucho taimen* reveal GenBank entry errors: Incorrect species identification and recombinant mitochondrial genome. *Evol. Bioinforma. Online* **13** (2017).
11. Frohlich, D. R., Torres-Jerez, I., Bedford, I. D., Markham, P. G. & Brown, J. K. A phylogeographical analysis of the *Bemisia tabaci* species complex based on mitochondrial DNA markers. *Mol. Ecol.* **8**, 1683–1691 (1999).
12. Tay, W. T. *et al*. The trouble with MEAM2: Implications of pseudogenes on species delimitation in the globally invasive *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species complex. *Genome Biol. Evol.* **9**, 2732–2738 (2017).
13. Seal, S. E., vandenBosch, F. & Jeger, M. J. Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control. *Crit. Rev. Plant Sci.* **25**, 23–46 (2006).
14. De Barro, P. J., Liu, S. S., Boykin, L. M. & Dinsdale, A. B. *Bemisia tabaci*: A statement of species status. *Annu. Rev. Entomol.* **56**, 1–19 (2011).
15. Boykin, L. M. *et al*. Global relationships of *Bemisia tabaci* (Hemiptera: Aleyrodidae) revealed using Bayesian analysis of mitochondrial COI DNA sequences. *Mol. Phylogenet. Evol.* **44**, 1306–1319 (2007).
16. Dinsdale, A., Cook, L., Riginos, C., Buckley, Y. M. & De Barro, P. J. Refined global analysis of *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. *Ann. Entomol. Soc. Am.* **103**, 196–208 (2010).
17. Boykin, L. M., Armstrong, K. F., Kubatko, L. & De Barro, P. J. Species delimitation and global biosecurity. *Evol. Bioinforma. Online* **8**, 1–37 (2012).
18. Maruthi, M. N. *et al*. Reproductive incompatibility and cytochrome oxidase I gene sequence variability amongst host-adapted and geographically separate *Bemesia tabaci* populations (Hemiptera: Aleyrodidae). *Syst. Entomol.* **29**, 560–568 (2004).
19. Xu, J., De Barro, P. J. & Liu, S. S. Reproductive incompatibility among genetic groups of *Bemisia tabaci* supports the proposition that the whitefly is a cryptic species complex. *Bull. Entomol. Res.* **100**, 359–366 (2010).
20. Sun, D. B., Xu, J., Luan, J. B. & Liu, S. S. Reproductive incompatibility between the B and Q biotypes of the whitefly *Bemisia tabaci* in China: Genetic and behavioural evidence. *Bull. Entomol. Res.* **101**, 211–220 (2011).
21. Liu, S. S., Colvin, J. & De Barro, P. J. Species concepts as applied to the whitefly *Bemisia tabaci* systematics: How many species are there? *J. Integr. Agric.* **11**, 176–186 (2012).
22. Hu, J. *et al*. An extensive field survey combined with a phylogenetic analysis reveals rapid and widespread invasion of two alien whiteflies in China. *PLoS One* **6**, e16061 (2011).
23. Alemandri, V. *et al*. Species within the *Bemisia tabaci* (Hemiptera: Aleyrodidae) complex in soybean and bean crops in Argentina. *J. Econ. Entomol.* **105**, 48–53 (2012).
24. Parrella, G., Scassillo, L. & Giorgini, M. Evidence for a new genetic variant in the *Bemisia tabaci* species complex and the prevalence of the biotype Q in southern Italy. *J. Pest Sci.* **85**, 227–238 (2012).
25. Chowda-Reddy, R. V. *et al*. *Bemisia tabaci* phylogenetic groups in India and the relative transmission efficacy of Tomato leaf curl Bangalore virus by an indigenous and an exotic population. *J. Integr. Agric.* **11**, 235–248 (2012).
26. Firdaus, S. *et al*. The *Bemisia tabaci* species complex: Additions from different parts of the world. *Insect Sci.* **20**, 723–733 (2013).
27. Esterhuizen, L. L. *et al*. Genetic identification of members of the *Bemisia tabaci* cryptic species complex from South Africa reveals native and introduced haplotypes. *J. Appl. Entomol.* **137**, 122–135 (2013).
28. Hu, J. *et al*. Members of *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species and the status of two invasive alien species in the Yunnan Province (China). *J. Insect Sci. Online* **14**, 281 (2014).
29. Hu, J. *et al*. New putative cryptic species detection and genetic network analysis of *Bemisia tabaci* (Hempitera: Aleyrodidae) in China based on mitochondrial COI sequences. *Mitochondrial DNA Part DNA Mapp. Seq. Anal.* **29**, 474–484 (2017).

30. Delatte, H. *et al.* Microsatellites reveal extensive geographical, ecological and genetic contacts between invasive and indigenous whitefly biotypes in an insular environment. *Genet. Res.* **87**, 109–124 (2006).
31. Mugerwa, H. *et al.* African ancestry of New World, *Bemisia tabaci*-whitefly species. *Sci. Rep.* **8**, 2734 (2018).
32. Tay, W. T., Evans, G. A., Boykin, L. M. & De Barro, P. J. Will the real *Bemisia tabaci* please stand up? *PLoS One* **7**, e50550 (2012).
33. Bellows, T. S., Perring, T. M., Gill, R. J. & Headrick, D. H. Description of a species of *Bemisia* (Homoptera: Aleyrodidae). *Ann. Entomol. Soc. Am.* **87**, 195–206 (1994).
34. Costa, H. S. & Brown, J. K. Variation in biological characteristics and esterase patterns among populations of *Bemisia tabaci*, and the association of one population with silverleaf symptom induction. *Entomol. Exp. Appl.* **61**, 221–219 (1991).
35. Delatte, H. *et al.* A new silverleaf-inducing biotype Ms of *Bemisia tabaci* (Hemiptera: Aleyrodidae) indigenous to the islands of the south-west Indian Ocean. *Bull. Entomol. Res.* **95**, 29–35 (2005).
36. Chu, D. *et al.* Genetic differentiation of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) biotype Q based on mitochondrial DNA markers. *Insect Sci.* **15**, 115–123 (2008).
37. Ahmed, M. Z. *et al.* Genetic distinctions among the Mediterranean and Chinese populations of *Bemisia tabaci* Q biotype and their endosymbiont *Wolbachia* populations. *J. Appl. Entomol.* **133**, 733–741 (2009).
38. Gueguen, G. *et al.* Endosymbiont metacommunities, mtDNA diversity and the evolution of the *Bemisia tabaci* (Hemiptera: Aleyrodidae) species complex. *Mol. Ecol.* **19**, 4365–4376 (2010).
39. Chu, D. *et al.* Use of mitochondrial cytochrome oxidase I polymerase chain reaction-restriction fragment length polymorphism for identifying subclades of *Bemisia tabaci* Mediterranean group. *J. Econ. Entomol.* **105**, 242–251 (2012).
40. Terraz, G., Gueguen, G., Arnó, J., Fleury, F. & Mouton, L. Nuclear and cytoplasmic differentiation among Mediterranean populations of *Bemisia tabaci*: Testing the biological relevance of cytotypes. *Pest Manag. Sci.* **70**, 1503–1513 (2014).
41. Gauthier, N. *et al.* Genetic structure of *Bemisia tabaci* MED populations from home-range countries, inferred by nuclear and cytoplasmic markers: Impact on the distribution of the insecticide resistance genes. *Pest Manag. Sci.* **70**, 1477–1491 (2014).
42. Thierry, M. *et al.* Mitochondrial, nuclear, and endosymbiotic diversity of two recently introduced populations of the invasive *Bemisia tabaci* MED species in La Réunion. *Insect Conserv. Divers.* **8**, 71–80 (2015).
43. Mouton, L. *et al.* Detection of genetically isolated entities within the Mediterranean species of *Bemisia tabaci*: New insights into the systematics of this worldwide pest. *Pest Manag. Sci.* **71**, 452–458 (2015).
44. Vähä, J. P. & Primmer, C. R. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Mol. Ecol.* **15**, 63–72 (2006).
45. Wang, H. L. *et al.* The characteristics and expression profiles of the mitochondrial genome for the Mediterranean species of the *Bemisia tabaci* complex. *BMC Genomics* **14**, 401 (2013).
46. Minh, B. Q., Nguyen, M. A. T. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**, 1188–1195 (2013).
47. Omondi, B. A., Sseruwagi, P., Obeng-Ofori, D., Danquah, E. Y. & Kyerematen, R. A. Mating interactions between okra and cassava biotypes of *Bemisia tabaci* (Homoptera: Aleyrodidae) on eggplant. *Int. J. Trop. Insect Sci.* **25**, 159–167 (2005).
48. Sseruwagi, P. *et al.* Genetic diversity of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) populations and presence of the B biotype and a non-B biotype that can induce silverleaf symptoms in squash, in Uganda. *Ann. Appl. Biol.* **147**, 253–265 (2005).
49. Tindall, K. R. & Kunkel, T. A. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry (Mosc.)* **27**, 6008–6013 (1988).
50. Keohavong, P. & Thilly, W. G. Fidelity of DNA polymerases in DNA amplification. *Proc. Natl. Acad. Sci. USA* **86**, 9253–9257 (1989).
51. Cline, J., Braman, J. C. & Hogrefe, H. H. PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.* **24**, 3546–3551 (1996).
52. McInerney, P., Adams, P. & Hadi, M. Z. Error rate comparison during polymerase chain reaction by DNA polymerase. *Mol. Biol. Int.* **2014**, 1–8 (2014).
53. McKenzie, C. L. *et al.* Distribution of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes in North America after the Q invasion. *J. Econ. Entomol.* **105**, 753–766 (2012).
54. Elfekih, S. *et al.* Genome-wide analyses of the *Bemisia tabaci* species complex reveal contrasting patterns of admixture and complex demographic histories. *PLoS One* **13**, e0190555 (2018).
55. Hadjistylli, M., Roderick, G. K. & Brown, J. K. Global population structure of a worldwide pest and virus vector: Genetic diversity and population history of the *Bemisia tabaci* sibling species group. *PLoS One* **11**, e0165105 (2016).
56. Duron, O. *et al.* The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol.* **6**, 27 (2008).
57. Werren, J. H., Baldo, L. & Clark, M. E. *Wolbachia*: Master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* **6**, 741–751 (2008).
58. Oliver, K. M., Russell, J. A., Moran, N. A. & Hunter, M. S. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc. Natl. Acad. Sci. USA* **100**, 1803–1807 (2003).
59. Himler, A. G. *et al.* Rapid spread of a bacterial symbiont in an invasive whitefly is driven by fitness benefits and female bias. *Science* **332**, 254–256 (2011).
60. Song, H., Buhay, J. E., Whiting, M. F. & Crandall, K. A. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc. Natl. Acad. Sci. USA* **105**, 13486–13491 (2008).
61. van der Kuyl, A. C., Kuiken, C. L., Dekker, J. T., Perizonius, W. R. K. & Goudsmit, J. Nuclear counterparts of the cytoplasmic mitochondrial 12S rRNA gene: A problem of ancient DNA and molecular phylogenies. *J. Mol. Evol.* **40**, 652–657 (1995).
62. Thalmann, O., Hebler, J., Poinar, H. N., Pääbo, S. & Vigilant, L. Unreliable mtDNA data due to nuclear insertions: A cautionary tale from analysis of humans and other great apes. *Mol. Ecol.* **13**, 321–335 (2004).
63. Yokomi, R. K., Hoelmer, K. A. & Osborne, L. S. Relationships between the sweetpotato whitefly and the squash silverleaf disorder. *Phytopathology* **80**, 895–900 (1990).
64. De la Rúa, P., Simón, B., Cifuentes, D., Martinez-Mora, C. & Cenis, J. L. New insights into the mitochondrial phylogeny of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) in the Mediterranean Basin. *J. Zool. Syst. Evol. Res.* **44**, 25–33 (2006).
65. Bedford, I. D., Briddon, R. W., Brown, J. K., Rosell, R. C. & Markham, P. G. Geminivirus transmission and biological characterisation of *Bemisia tabaci* (Gennadius) biotypes from different geographic regions. *Ann. Appl. Biol.* **125**, 311–325 (1994).
66. Brown, J. K. *et al.* Characterization and distribution of esterase electromorphs in the whitefly, *Bemisia tabaci* (Genn.) (Homoptera: Aleyrodidae). *Biochem. Genet.* **33**, 205–214 (1995).
67. Burban, C., Fishpool, L. D. C., Fauquet, C., Fargette, D. & Thouvenel, J. C. Host-associated biotypes within West African populations of the whitefly *Bemisia tabaci* (Genn.), (Hom., Aleyrodidae). *J. Appl. Entomol.* **113**, 416–423 (1992).
68. Martin, J. H. & Mound, L. A. An annotated check list of the world's whiteflies (Insecta: Hemiptera: Aleyrodidae). *Zootaxa* **1492**, 1–84 (2007).
69. Walsh, P. S., Metzger, D. A. & Higuchi, R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* **10**, 506–513 (1991).
70. Kearse, M. *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinforma. Oxf. Engl.* **28**, 1647–1649 (2012).
71. Ghosh, S., Bouvaine, S. & Maruthi, M. N. Prevalence and genetic diversity of endosymbiotic bacteria infecting cassava whiteflies in Africa. *BMC Microbiol.* **15** (2015).

72. Byrne, F. J. & Devonshire, A. L. Biochemical evidence of haplodiploidy in the whitefly *Bemisia tabaci*. *Biochem. Genet.* **34**, 93–107 (1996).
73. Wang, P., Sun, D. B., Qiu, B. L. & Liu, S. S. The presence of six cryptic species of the whitefly *Bemisia tabaci* complex in China as revealed by crossing experiments. *Insect Sci.* **18**, 67–77 (2011).
74. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2013).
75. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S*. (Springer-Verlag, 2002).
76. Tukey, J. W. Comparing individual means in the analysis of variance. *Biometrics* **5**, 99–114 (1949).
77. Hothorn, T., Bretz, F. & Westfall, P. Simultaneous inference in general parametric models. *Biometrical Journal* **50**, 346–363 (2008).
78. Gershoni, M., Templeton, A. R. & Mishmar, D. Mitochondrial bioenergetics as a major motive force of speciation. *BioEssays* **31**, 642–650 (2009).
79. Chen, W. *et al*. The draft genome of whitefly *Bemisia tabaci* MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. *BMC Biol.* **14**, 110 (2016).
80. Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol. Clifton NJ* **132**, 365–386 (2000).
81. Bonfield, J. K., Smith, K. F. & Staden, R. A new DNA sequence assembly program. *Nucleic Acids Res.* **23**, 4992–4999 (1995).
82. Villesen, P. FaBox: An online toolbox for fasta sequences. *Mol. Ecol. Notes* **7**, 965–968 (2007).
83. Thangaraj, S. R., McCulloch, G. A., Subbarayalu, M., Subramaniam, C. & Walter, G. H. Development of microsatellite markers and a preliminary assessment of population structuring in the rice weevil, *Sitophilus oryzae* (L.). *J. Stored Prod. Res.* **66**, 12–17 (2016).
84. Jiang, H., Lei, R., Ding, S. W. & Zhu, S. Skewer: A fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).
85. Tay, W. T. *et al*. Novel molecular approach to define pest species status and tritrophic interactions from historical *Bemisia* specimens. *Sci. Rep.* **7**, 429 (2017).
86. Tay, W. T., Elfekih, S., Court, L. N., Gordon, K. H. J. & De Barro, P. J. Complete mitochondrial DNA genome of *Bemisia tabaci* cryptic pest species complex Asia I (Hemiptera: Aleyrodidae). *Mitochondrial DNA Part DNA Mapp. Seq. Anal.* **27**, 972–973 (2016).
87. Thao, M. L., Baumann, L. & Baumann, P. Organization of the mitochondrial genomes of whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha). *BMC Evol. Biol.* **4**, 25 (2004).
88. Chen, Y., Ye, W., Zhang, Y. & Xu, Y. High speed BLASTN: An accelerated MegaBLAST search tool. *Nucleic Acids Res.* **43**, 7762–7768 (2015).
89. Librado, P. & Rozas, J. DnaSPv5: A software for comprehensive analysis of DNA polymorphism data. *Bioinforma. Oxf. Engl.* **25**, 1451–1452 (2009).
90. Trifinopoulos, J., Nguyen, L. T., von Haeseler, A. & Minh, B. Q. W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* **44**, W232–W235 (2016).
91. Chernomor, O., von Haeseler, A. & Minh, B. Q. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst. Biol.* **65**, 997–1008 (2016).
92. Xia, X. DAMBE6: New tools for microbial genomics, phylogenetics, and molecular evolution. *J. Hered.* **108**, 431–437 (2017).
93. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
94. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
95. Huson, D. H. & Scornavacca, C. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* **61**, 1061–1067 (2012).

Acknowledgements

We acknowledge the financial support of SV's PhD scholarship and research by University of Greenwich and Bill & Melinda Gates Foundation (African Cassava Whitefly Project, OPP1058938). The high-throughput sequencing was supported by the Cotton Research and Development Corporation (CRDC) Australia under grant UQ1305. SV's visiting PhD student placement was hosted by CSIRO Health & Biosecurity (R-8681-1). We thank Jesús Navas-Castillo, Shai Morin and Habibu Mugerwa for collecting the whitefly samples for core colonies, Mitulkumar Patel and Nikunj Naik for extracting candidate genes from MED transcriptomes, Stephen Young for help with statistical analysis, Paul Visendi Muhindira for help with quality control of Illumina HiSeq reads, Karl Gordon, Andrew Polaszek and David Ouvrard for a helpful discussion, and Bill James and Quentin Bédrune for technical support.

Author Contributions

J.C., S.S., S.V. and W.T.T. conceived the ideas; S.V. and S.v.B. did laboratory work; S.V., W.T.T. and S.v.B. analysed DNA sequence data; S.V. and W.T.T. prepared the figures; all authors contributed to writing the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-29305-w.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2. Journal of Pest Science (in print)

Relative polyphagy of "Mediterranean" cryptic *Bemisia tabaci* whitefly species and global pest status implications

Soňa Vyskočilová[1], Susan Seal[1], John Colvin[1]

[1] Natural Resources Institute, University of Greenwich, Central Avenue, Chatham Maritime, ME4 4TB, Kent, United Kingdom.

Abstract

The *Bemisia tabaci* whitefly was previously considered a single, highly polyphagous species, but is now accepted as a group of cryptic biological species. We investigated the host plant relations of two sister species, the "Mediterranean" (MED) composed of the Q1 and Q2 mitochondrial groups and the "ASL" species (formerly considered a MED group), to discover whether polyphagy was related to the global pest status. We compared their performance by measuring the oviposition rate, survival, fecundity and proportion of female offspring on 13 host plants from nine families. In addition, oviposition preference was compared among leaves of different ages. Significant ($P < 0.05$) differences were found between populations in all parameters, particularly in adult survival and fecundity. Females preferred strongly to oviposit on the oldest true leaf on tobacco and tomato and on young leaves of sweet potato and pepper. The greatest differences in fecundity occurred on bean, okra, squash, pepper and tobacco. Hosts favourable for all four populations were cotton and sweet potato; no offspring were produced on cassava, chard or tomato. Host ranges of Q1 populations from Europe and sub-Saharan Africa differed despite their close genetic relatedness at the mitogenome level. Discrepancies between the parental and offspring fitness were observed. Our findings show that (1) the species have differing but overlapping host plant ranges and (2) the Q1 is the most polyphagous and can utilise tobacco, which predisposes it to evolving resistance to neonicotinoids. Our findings contribute to the understanding of ecology of this pest species complex and aid the development of efficient pest control strategies.

Keywords

Mitochondrial phylogenetic groups, invasiveness, survival analysis, optimal oviposition theory, optimal foraging theory, bacterial endosymbionts

Key message

- The "Mediterranean" (MED) putative species of the *Bemisia tabaci* complex is a globally invasive pest and plant-virus vector.
- MED was recently discovered to be more than one cryptic *B. tabaci* species; MED groups Q1, Q2 form one species and African silver-leafing (ASL) is the other.
- Host-plant range differences between these cryptic species and how this may affect their relative pest status remained unknown.
- We found significant differences in host use between these sister species, which has implications for their control.

Author contributions

SV, SS and JC conceived the research and contributed ideas. SV conducted experiments, analysed the data and wrote the manuscript. All authors read and approved the manuscript.