UNIVERSITY OF PADOVA

Department of Agronomy, Food, Natural Resources, Animals and the Environment

(DAFNAE)

DOCTORATE SCHOOL OF CROP SCIENCE

CURRICULUM CROP PROTECTION – CYCLE XXVI

Symbiotic bacteria inhabiting tephritid flies:

a worldwide specific interaction

Director of the school: Ch.mo Prof. Antonio Berti

Supervisor: Ch.mo Prof. Vincenzo Girolami

PhD Student: Elisabetta Viale

Date of thesis submission

July 31th 2014

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

(Elisabetta Viale/July 31, 2014)

A copy of the thesis will be available at *http://paduaresearch.cab.unipd.it/*

Dichiarazione

Con la presente affermo che questa tesi è frutto del mio lavoro e che, per quanto io ne sia a conoscenza, non contiene materiale precedentemente pubblicato o scritto da un'altra persona né materiale che è stato utilizzato per l'ottenimento di qualunque altro titolo o diploma dell'università o altro istituto di apprendimento, a eccezione del caso in cui ciò venga riconosciuto nel testo.

(Elisabetta Viale/31 Luglio, 2014)

Una copia della tesi sarà disponibile presso *http://paduaresearch.cab.unipd.it/*

Table of contents

Riassunto 9

Summary 11

Chapter I – Introduction 13

Bacterial endosymbionts of insect 14

The Tephritidae 17

Bacterial symbiosis in Tephritid flies 18

Objective and contents of the thesis 19

References 20

Chapter II - Pattern of association between endemic Hawaiian fruit flies (Diptera, Tephritidae) and their symbiotic bacteria: evidence of cospeciation events and proposal of '*Candidatus* Stammerula trupaneae' 27

Abstract 28

Introduction 29

Material and methods 32

Insect sampling 32

Insect host analyses: DNA extraction, amplification and sequencing 32

Phylogenetic analyses of the insects 34

Symbiont analyses: Bacterial DNA amplification and sequencing 35

Symbiont phylogenetic analyses 35

Host-symbiont cophylogeny analyses 35

Evolution of ancestral characters 37

Results 38

Sequence analyses of insect 38

Phylogeny of Hawaiian tephritids 38

Presence of endosymbionts and sequence analyses 39

Phylogeny of symbionts 39
Host-symbiont cophylogenetic analyses 40
Character reconstruction 41
Discussion 42
'Candidatus Stammerula trupaneae' 42
Molecular relationships among Hawaiian Tephritinae symbionts 43
Concordant evolution of Hawaiian Tephritinae and their symbiotic bacteria 43
Conclusion 45
Tables and figures 47
References 61

Chapter III - Genetic variability of '*Candidatus* Erwinia dacicola' in *Bactrocera oleae* populations across their geographical distribution 67
Introduction 68
Material and methods 70
Olive fly sampling and dissection 70
Symbiont data analyses 70
Insect host analyses: amplification and sequencing 71
Insect data analyses 72
Symbiont-host haplotype analysis 73
Results 74
Haplotype analysis and distribution of Ca. E. dacicola 74
Mitochondrial haplotype analysis and population structure of B. oleae 75
Relationships between Ca. Erwinia and B. oleae haplotypes 77
Discussion 79
Genetic variability and pattern of distribution of Ca. E. dacicola 79
Population structure of B. oleae 80
Conclusion 82
Tables and figures 84
References 100

Conclusion 105
References 108
Acknowledgements 110

Dedidated to all those who have worked in the past, who are presently working,
and who will work in the future
toward understanding Hawaii's natural heritage.

Riassunto

Diverse specie di insetto si sono evolute in associazione con i loro batteri simbionti. Questo è il caso di alcuni membri dei Tephritinae, la più specializzata sottofamiglia delle mosche della frutta (Diptera: Tephritidae), che ospitano nell'intestino batteri simbionti coevoluti e trasmessi in maniera verticale, conosciuti come "*Candidatus* Stammerula spp.". Nella mosca dell'olivo, *Bactrocera oleae*, i batteri simbionti sono localizzati nel bulbo esofageo, un diverticolo presente nel capo della mosca, e identificati con il nome di "*Candidatus* Erwinia dacicola".

Questo lavoro, basato su due principali studi, si focalizza su diversi aspetti delle relazioni filogenetiche che intercorrono tra le mosche della frutta e i loro batteri simbionti.

Il primo lavoro studia la presenza di specifici batteri simbionti in 15 delle 25 specie descritte di tefritidi endemici dell'Arcipelago delle Hawaii, uno spettacolare esempio di radiazione adattativa, e le relazioni molecolari che intercorrono con i simbionti delle Tephritinae non Hawaiiani. Inoltre è stata analizzata la concordanza evolutiva tra la filogenesi dell'insetto rispetto a quella del simbionte. Uno specifico simbionte è stato individuato mediante saggi di PCR in tutti gli individui analizzati e nominato "*Candidatus* Stammerula trupaneae", in quanto incluso nel gruppo monofiletico formato da *Ca.* Stammerula spp. La filogenesi dell'insetto ospite è stata ricostruita analizzando due regioni del DNA mitocondriale (16S rDNA e COI-tRNALeu-COII), mentre il gene batterico 16S rRNA è stato utilizzato nell'analisi del simbionte. Le filogenesi dell'ospite e del simbionte sono state quindi comparate e valutate per lo studio del modello di congruenza filogenetica e cospeciazione. La congruenza tra la filogenesi delle Tephritinae Hawaiiane e i loro batteri simbionti suggerisce un ridotto, ma significativo livello di cospeciazione. L'evoluzione dei caratteri ancestrali, basata su tre aspetti dell'insetto quali l'isola di origine, la pianta ospite e il tessuto vegetale attaccato dalla mosca, è stata infine ricostruita sulla base della filogenesi del simbionte ipotizzando la presenza di cospeciazione.

Il secondo studio analizza la variabilità genetica del simbionte della mosca dell'olivo, *Ca.* Erwinia dacicola, insieme al grado di differenziazione genetica di *B.*

oleae, su un ampio raggio della sua distribuzione geografica, comprendendo molte regioni del Mediterraneo e alcuni campionamenti puntiformi in Sud Africa, California e Pakistan. Tre aplotipi batterici, con una significativa distribuzione geografica, sono stati identificati ed è stata esclusa la coesistenza di diversi aplotipi di *Ca.* E. dacicola nella stessa mosca. Nelle popolazioni della mosca dell'olivo raccolte nel Mediterraneo, solo due aplotipi batterici (htA e htB), identificati in precedenza nelle popolazioni Italiane, sono stati trovati, mostrando una significativa distribuzione Est-Ovest. Le popolazioni del Sud Africa e della California sono rappresentate in maniera esclusiva da uno dei due aplotipi, rispettivamente htA e htB. Un nuovo aplotipo (htC) inoltre è stato individuato esclusivamente nelle popolazioni Pakistane.

D'altro lato, un alto grado di variabilita' genetica caratterizzato da una certa differenziazione geografica è stato osservato nelle popolazioni di *B. oleae* analizzate; i nostri risultati mostrano la presenza di 39 aplotipi dell'insetto. Gli aplotipi del simbionte e quelli dell'insetto sono stati quindi confrontati e un'associazione significativa, con una stretta correlazione al territorio, è stata trovata, evidenziando la presenza di una prevalente trasmissione verticale del simbionte durante il ciclo vitale dell'insetto. Inoltre, il fatto che la distribuzione degli aplotipi batterici sia più strettamente correlata al territorio rispetto a quella ritrovata nei numerosi aplotipi dell'insetto ospite, può rappresentare un importante mezzo per ricostruire la dibattuta origine della mosca dell'olivo.

Summary

Several insect lineages have evolved mutualistic associations with their symbiotic bacteria. This is the case of some members of Tephritinae, the most specialized subfamily of fruit flies (Diptera: Tephritidae), harboring co-evolved and vertically transmitted bacterial symbionts in their midgut, known as '*Candidatus* Stammerula spp.'. In the tephritid fly *Bactrocera oleae*, the major olive pest, symbionts are located in the oesophageal bulb, a diverticulum of the fly head, and designated as '*Candidatus* Erwinia dacicola'.

This research, based on two main studies, is focused on different aspects of the relationships between species of the family Tephritidae and their nonculturable symbiotic bacteria.

The first study investigated the presence of specific symbiotic bacteria in 15 of the 25 described endemic tephritids of the Hawaiian Archipelago, which represent a spectacular example of adaptive radiation, and their molecular relationships with symbionts of non-Hawaiian tephritids. Moreover the concordant evolution between host and symbiont phylogenies was tested. A specific symbiont was detected through PCR assays in all endemic individuals analyzed and it was designated as '*Candidatus* Stammerula trupaneae' as it was included in *Ca.* Stammerula spp. monophyletic clade. The phylogeny of the insect host was reconstructed based on two regions of the mitochondrial DNA (16S rDNA and COI-tRNALeu-COII), while the bacterial 16S rRNA was used for the symbiont analysis. Host and symbiont phylogenies were then compared and evaluated for patterns of cophylogeny and strict cospeciation. Topological congruence between Hawaiian Tephritinae and their symbiotic bacteria phylogenies suggests a limited, but significant degree of host-symbiont cospeciation. The character evolution of three host traits, as island location, host lineage, and host tissue attacked, was finally reconstructed based on the symbiont phylogenies under the hypothesis of cospeciation.

The second study surveys the genetic variability of the olive fly symbiont, *Ca.* Erwinia dacicola, together with the patterns of genetic differentiation of *B. oleae*, over a large area of its geographical distribution, including most regions of the

Mediterranean area, plus South Africa, California and Pakistan. Three bacterial haplotypes, showing a significative geographic distribution, were identified and the co-existence of different *Ca.* E. dacicola haplotypes in a single fly was never found. Our results reveal the presence of three symbiont haplotypes with a significant phylogeographic distribution related to the territory. In the Mediterranean populations only two bacterial lineages (htA and htB), previously recovered in Italian olive fly populations, have been found, showing a significative East-West genetic differentiation. The South African and Californian olive fly populations were represented only by one of these two lineages, respectively htA and htB. Moreover, a new haplotype (htC) was detected exclusively in the Pakistani population. On the other hand, a high degree of mitochondrial genetic variability with a substantial phylogeographic differentiation has been observed in the *B. oleae* populations analyzed, revealing the presence of 39 insect haplotypes. Symbiont and host haplotypes were then compared and a significant correlation was found suggesting the predominant presence of vertical transmission. Moreover, the bacterial haplotypes distribution seems to be more related to the territory than the numerous insect host haplotypes, representing an useful tool to reconstruct the debated olive fly's historical origin.

Chapter I

Introduction

Bacterial endosymbionts of insect

Mutualistic symbioses between unicellular and multicellular organisms have contributed significantly to the evolution of life on Earth (Margulis and Fester, 1991). Contrary to most prokaryotes, eukaryotes have rather limited metabolic capabilities and, hence, symbiosis has provided an evolutionary strategy for eukaryotes to gain access to a wider range of metabolic resources. One special case of symbiosis is endosymbiosis, in which one partner, generally a prokaryote, is located inside the body of the other. In some cases, the prokaryote is literally sequestered within a eukaryotic cell and cannot be cultured on common laboratory media. Some of the best examples of eukaryotic-bacterial symbioses are mutualisms between insects and proteobacteria. The vast diversity of insect species has been hypothesized to be due to their propensity to associate with beneficial bacteria (Janson et al. 2008).

Insects are the most species-rich group of organisms (Basset et al., 2012), and it has been estimated that at least more than 15–20% of all insects, in several taxonomic orders, live in symbiotic relationships with bacteria (Buchner, 1965) and depend on obligate bacterial mutualists for their viability and reproduction (Buchner, 1965; Wernegreen et al., 2002, Moran et al., 2008). These ancient associations between bacteria and insects, known since last century by Petri (1909), allow hosts to exploit new niches and therefore contribute to host diversification and success (Wernergreen et al., 2002). As for essentially all animals, microbial communities are particularly prominent in the digestive tract, where they may be key mediators of the varied lifestyles of insect hosts (Engel and Moran, 2013).

During the past 20 years, technological advances in molecular phylogenetic characterization have enabled exploration of the world of these uncultured bacteria (Moran and Wernergreen, 2000). Although there are a few evolutionary studies of symbionts in other invertebrate groups (Parecer and Ahmadjian, 2000), most studies have focused on mutualistic and obligate insect–bacteria endosymbiosis, yielding considerable insight into their evolutionary histories and into the specific adaptations of bacteria towards symbiosis, and are the focus of this thesis.

Symbiotic bacteria have diverse ecological and evolutionary effects on hosts,

influencing aspects of ecological interactions from nutrition to defense and affecting reproductive systems, with consequences for population structure, reproductive isolation, and speciation (Moran et al., 2008).

The early establishment of symbiotic associations between insects and bacteria, estimated about 300 MY ago (Moran and Telang, 1998), together with the nutritional enrichment that bacteria offer to insects, could be the key factors in the evolutionary success of this group of organisms (Douglas, 1989; Moran and Bauman, 2000).

Beginning with early descriptive studies based on microscopy (Buchner, 1965), and later molecular studies, symbioses have been categorized based on both apparent evolutionary age and on the extent of obligate codependence between the host and symbiont. Therefore, heritable symbionts of insects can be divided into two intergrading categories: obligate and facultative endosymbionts.

Obligate endosymbionts, or "primary" symbionts, reside in the cytosol of specialized host cells called bacteriocytes, which may form an organ-like structure (bacteriome) in the body cavity of the insects. They are evolutionarily ancient and involve specialization on the part of both host and symbiont (Engel and Moran, 2013). These associations, in which bacteria form persistent infections within host individuals, are obligate for both partners: the bacteria cannot be cultured outside the host, whereas the host needs the bacteria for normal growth and reproduction (Baumann, 2005). Obligate symbionts also show highly specific interactions with host cell populations, both during colonization of progeny and during early development of host individuals (Braendle et al., 2003). Bacteriocyte-associated endosymbionts of insects are vertically transmitted from the mother to the offspring via infection of eggs or embryos, often through complex developmental events that ensure transovariole transfer to the developing egg or embryo (Buchner, 1965; Houk and Griffiths, 1980). The most intensively studied example is the bacteriocyte symbiosis between aphids and the bacterial symbiont, *Buchnera aphidicola* (Munson et al., 1991), originating about 200 million years ago (Baumann, 2005) but several endosymbiotic relationships have also been described in many other insect families, including weevils (Nardon and Grenier, 1988), mealybugs (Tremblay, 1989), whiteflies (Clark et al., 1992), tsetse flies (Aksoy et al., 1995), psyllids (Fukatsu and Nikoh, 1998), leafhoppers (Moran et

al., 2003), carpenter ants (Blochmann, 1892; Schröder et al., 1996), cockroaches (Bandi et al., 1995) and fruit flies (Girolami, 1983; Capuzzo et al., 2005). Genome sequencing reveals that obligate bacterial endosymbionts of insects have among the smallest of known bacterial genomes and have lost many genes that are commonly found in closely related bacteria, in addition to have the fastest in rate of gene sequence evolution (Nakabachi et al., 2006; Shigenobu et al., 2000; Moran, 2006, Moran et al., 2008). Molecular phylogenetic studies based on 16S rDNA sequences revealed phylogenic congruence of these primary endosymbiotic bacteria with their hosts, indicating in most cases host–symbiont cospeciation (Baumann et al., 2000; Moran et al., 2003, Mazzon et al., 2010). All these studies span ancient divergences among taxa, indicating that each symbiosis is the result of a single bacterial infection of the insect ancestor, which was followed by co-evolution of both partners across millions of years (Dale and Moran, 2006). For those reasons, they represent an ideal system in which to study the changes in rates and patterns of bacterial evolution that have occurred during the transition from a free-living to an intracellular lifestyle.

These ancient, obligate symbioses can be contrasted with heritable symbionts not required by the host, mostly facultative endosymbionts, or "secondary" symbionts, that are more sporadically associated with host individuals and vary in tissues occupied, and may reside extracellularly in the body cavity (hemolymph) (Dobson, 1999; Moran et al., 2008). They originate in multiple independent infections and are exemplified by *Wolbachia pipientis*, *Spiroplasma species,* and *Hamiltonella defensa,* which are largely maternally transmitted but undergo occasional horizontal transmission, causing host and symbiont evolution to be decoupled (Baumann et al., 2005). In these cases, the symbionts retain larger and more dynamic genomes (Werren et al., 2008; Degnan et al., 2009) and possess mechanisms for actively invading host tissues and for affecting host biology in a way that promotes the increased frequency of infected hosts in the host population (Werren et al., 2008; Oliver et al., 2010, Engel and Moran, 2013).

The Tephritidae

The Tephritidae family, commonly known as fruit flies, is one of the largest families of Diptera, with about 4800 described species in almost 500 genera (White, 2006). The family is distributed in the temperate, subtropical, and tropical regions of the world, with the greatest diversity of species occurring in the tropics (White, 1988). Within the Tephritidae, approximately 70 species are considered as important agricultural pests while many others may cause minor damage or are potentially harmful (White and Elson-Harris, 1992). Fruits are the main hosts of the most deleterious genera: *Anastrepha, Ceratitis, Bactrocera, Dacus* and *Rhagoletis*, however, host range varies considerably, often among closely related species, including major commercial crops (Norrbom & Kim, 1988; Goeden, 1994, Norrbom, 1999).

Many of these flies are highly polyphagous, utilizing a large variety of fruits or other food sources, such as *Ceratitis capitata* (Widemann), reported on more than 300 hosts (Liquido *et al.,* 1991); in contrast, other species are strictly oligophagous or monophagous, such as *Bactrocera oleae* (Rossi), which breeds only on olive fruits.

Three subfamilies of Tephritidae are recognized, the Dacinae, Trypetinae and Tephritinae (Foote, 1993), but the classification of subfamilies is currently under revision. Therefore there are substantial differences in the feeding behaviors of flies and in their developmental stages. Dacinae and Trypetinae groups use the fleshy fruit of host plants from a wide variety of families as larval food sources.

In the subfamily Tephritinae, the most specialized subfamily of Tephritidae (Korneyev, 1999), the larvae feed on plant tissue, either within capitula (seedhead), shootips meristems, or stem galls mainly of members of the Asteraceae (Headrick and Goeden, 1994). For this reason Tephritinae are considered a non-frugivorous group of tephritids (Zwölfer, 1983; Straw, 1989; Headrick and Goeden, 1998). The subfamily Tephritinae includes about 200 genera with over 1,800 species from all zoogeographical regions (Norrbom et al., 1999) to subarctic and mountain tundra, as well as alpine and arid deserts (Korneyev, 1999).

Bacterial symbiosis in Tephritid flies

The relations between fruit flies and bacteria are of different kinds.
The first hereditary symbiosis has been described by Petri (1909) in the olive fly, in which a bacteria-filled esophageal bulb is present, and the gut symbionts are restricted to the intestinal lumen (Capuzzo et al., 2005). The mother transmits symbiotic bacteria to the new generation, smearing the surface of its eggs with bacteria. Bacteria are extracellular but multiply inside the intestinal caeca at the larval stage (Petri, 1909; Stammer, 1929), presumably in contact with free living intestinal bacteria. In the adult insect's head, obligate symbionts multiply within the oesophageal bulb, detected in all adult tephritid flies (Girolami, 1973). In the olive fly, Capuzzo and colleagues (2005) characterized these hereditary, uncultured specific symbiotic bacteria as '*Candidatus* Erwinia dacicola'. Subsequent recent studies regarding symbiosis in olive fly have confirmed the presence of *Erwinia dacicola* in specimens coming from different geographical areas (Sacchetti et al., 2008; Estes et al., 2009; Kounatidis et al., 2009).

In the Tephritinae, first studies were carried out by Stammer (1929), which recovered the presence of bacterial symbiosis in several genera. The digestive tract of many of these flies has evolved to contain specialized cavities or organs within which bacterial symbionts are hosted (Stammer, 1929; Mazzon et al., 2008; Mazzon et al., 2011). In adults flies, the esophageal bulb appears to be devoid of microorganisms and symbiotic bacteria are located in the gut lumen, outside the peritrophic membrane (a thin chitinous-proteinaceous membrane that separates food from midgut tissue) in contact with midgut epithelial cells (Girolami, 1973, 1983). Larval stages, instead, maintain bacteria in their intestinal ceca similarly to the larvae of B. *oleae.*

As reported by Mazzon (2008), these specific bacteria, known as '*Candidatus* Stammerula spp.', could not be cultured and belong to the family *Enterobacteriaceae* that appear to constitute the largely dominant symbiotic clade in these associations with their populations and functions exhibiting large variations between the different life stages of the hosts.

Objective and contents of the thesis

In the present thesis, based on mitochondrial markers, different aspects of the relationships between species of the family Tephritidae and their non culturable symbiotic bacteria have been studied. The following is a summary of the most important findings of this work based upon two main studies.

The first contribution (Chapter II) extends the studies started in 2008, focused on the bacterial symbioses in fruit flies of the subfamily Tephritinae (Mazzon et al., 2008; 2010), to other non-European tephritids. In particular, this work is focused on specific symbiotic bacteria of tephritids endemic to the Hawaiian Archipelago and on their molecular relationship with symbionts of non-Hawaiian species (*Candidatus* Stammerula spp.). This study enables us to better understand the essential mechanism of symbiosis associated with this group of tephritid flies. Moreover, the high level of endemism of Hawaiian species offers an important opportunity to evaluate the pattern of host-symbiont cospeciation.

The second study (Chapter III) surveys the genetic variability of the olive fly symbiont, *Candidatus* Erwinia dacicola, already analyzed in a previous work (Savio et al., 2012) on Italian olive fly populations, over a large area of *Bactrocera oleae* geographical distribution, including most regions of the Mediterranean area, plus South Africa, California and Pakistan. This approach gives us an interesting insight into the long host-symbiont coevolutionary history providing more information about the debated expansion and colonization history of the species.

References

Aksoy, S., Chen, X., Chow, A., 1995. Mycetome endosymbionts of tsetse flies constitute a distinct lineage related to the Enterobacteriaceae. Insect. Mol. Biol. 4, 15–22.

Basset, Y., Cizek, L., Cuénoud, P., Didham, R. K., Guilhaumon, F., Missa, O., Leponce, M., 2012. Arthropod diversity in a tropical forest. Science. 338, 1481-1484.

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

Blochmann, F., 1892. Über das Vorkommen bakterienähnlicher Gebilde in den Geweben und Eiern verschiedener Insekten. Zbl Bakteriol. 11, 234–240.

Braendle, C., Miura, T., Bickel, R., Shingleton, A.W., Kambhampati, S., Stern, D.L., 2003. Developmental origin and evolution of bacteriocytes in the aphid-*Buchnera* symbiosis. PLoS Biol. 1, E21.

Bandi, C., McCall, J.W., Genchi, C., Corona, S., Venco, L., Sacchi, L., 1999. Effects of tetracycline on the filarial worms *Brugia pahangi* and *Dirofilaria immitis* and their bacterial endosymbionts *Wolbachia.* Int. J. Parasitol. 29, 357–364.

Baumann, P., Moran, N.A., Baumann, L., 2000. The Prokaryotes, A Handbook on the Biology of Bacteria; Ecophysiology, Isolation, Identification, Applications. The Marine Gram-negative Eubacteria: Genus Photobacterium, Beneckea, Alteromonas, Pseudomonas and Alcaligenes, Springer-Verlag, New York, pp. 1302-1330.

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

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

Capuzzo, C., Firrao, G., Mazzon, L., Squartini, A., Girolami, V., 2005. '*Candidatus* Erwinia dacicola', a coevolved symbiotic bacterium of the olive fly *Bactrocera oleae* (Gmelin). Int. J. Syst. Evol. Microbiol. 55, 1641–1647.

Clark, M.A., Baumann, L., Munson, M.A., Baumann, P., Campbell, B.C., Duffus, J.E., Moran, N.A., 1992. The eubacterial endosymbionts of whiteflies (Homoptera:

Aleyrodoidea) constitute a lineage separated from the endosymbionts of aphids and mealybugs. Curr. Microbiol. 25, 119–123.

Dale, C., Moran, N.A., 2006. Molecular mechanisms underlying symbiosis with heritable bacteria. Cell. 126, 453–65.

Degnan, P.H., Yu, Y., Sisneros, N., Wing, R.A., Moran, N.A., 2009. Hamiltonella defensa, genome evolution of protective bacterial endosymbiont from pathogenic ancestors. P. Natl. Acad. Sci. USA. 106, 9063–9068.

Dobson, S.L., Bourtzis, K., Braig, H.R., Jones, B.F., Zhou, W., Rousset, F., O'Neill, S.L., 1999. *Wolbachia* infections are distributed throughout insect somatic and germ line tissues. Insect Biochem. Mol. Biol. 29, 153–60.

Douglas, A.E., 1989. Mycetocyte symbiosis in insects. Biol. Rev. Camb. Philos. Soc. 64, 409–434.

Engel, P., and Moran, N.A., 2013. The gut microbiota of insects–diversity in structure and function. FEMS Microbiol. Rev. *37*, 699-735.

Estes, A.M., Hearn, D.J., Bronstein, J.L., Pierson, E.A., 2009. The olive fly endosymbiont, "*Candidatus* Erwinia dacicola", switches from an intracellular to an extracellular existence during host insect development. Appl. Environ. Microbiol. 75, 7097–7106.

Foote, R.H., Blanc, F.L., Norrbom, A.L., 1993. Handbook of the Fruit Flies (Diptera:Tephritidae) of America North of Mexico.Ithaca, NY/London: Comstock., pp. 571.

Fukatsu, T., and Nikoh, N., 1998. Two intracellular symbiotic bacteria from the mulberry psyllid *Anomoneura mori* (Insecta, Homoptera). Appl. Environ. Microbiol. 64, 3599– 3606.

Girolami, V., 1973. Reperti morfo-istologici sulle batteriosimbiosi del *Dacus oleae* Gmelin e di altri ditteri tripetidi, in natura e negli allevamenti su substrati artificiali. Redia. 54, 269–294.

Girolami, V., 1983. Fruit fly symbiosis and adult survival: general aspects. In: Cavalloro, R., (Eds.), Fruit flies of economic importance, Proceedings of the CEC/IOBC International Symposium, Athens, pp. 74–76.

Goeden R.D., 1994 Analysis of known and new host records for *Paroxyna* from California (Diptera: Tephritidae). Proc. Entomol. Soc. Wash. 96, 281-287.

Headrick, D.H., and Goeden, R.D., 1998. The biology of nonfrugivorous tephritid fruit flies. Annu. Rev. Entomol. 43, 217–241.

Houk, E.J., Griffiths, G.W., 1980. Intracellular symbiotes of the Homoptera. Annu. Rev. Entomol. 25, 161–187.

Janson, E.M., Stireman, J. O., Singer, M.S., Abbot, P., 2008. Phytophagous insect-microbe mutualisms and adaptive evolutionary diversification. Evol. 62, 997- 1012.

Korneyev, V.A. (1999). Phylogeny of the subfamily Tephritinae: relationships of the tribes and subtribes. In: Aluja, M., and Norrbom, A.L. (Eds.), Fruit flies (Tephritidae): phylogeny and evolution of behavior, CRC Press, Boca Raton, FL, pp. 549–580.

Kounatidis, I., Crotti, E., Sapountzis, P., Sacchi, L., Rizzi, A., Chouaia, B., Bandi, C., Alma, A., Daffonchio, D., Mavragani-Tsipidou, P., Bourtzis, K., 2009. *Acetobacter tropicalis* is a major symbiont in the olive fruit fly *Bactrocera oleae*. Appl. Environ. Microbiol. 75, 3281–3288.

Liquido, N.J., Cunningham, R.T., Nakagawa S., 1990. Host plants of Mediterranean fruit fly (Diptera: Tephritidae) on the Island of Hawaii (1949–1985 survey). J. Econ. Entomol. 83, 1863–1878.

Mazzon, L., Piscedda, A., Simonato, M., Martinez-Sañudo, I., Squartini, A., and Girolami, V., 2008. Presence of specific symbiotic bacteria in flies of the subfamily Tephritinae (Diptera Tephritidae) and their phylogenetic relationships: proposal of '*Candidatus* Stammerula tephritidis.' Int. J. Syst. Evol. Microbiol. 58, 1277–1287.

Mazzon, L., Martinez-Sanudo I., Simonato M., Squartini A., Savio C., Girolami V., 2010. Phylogenetic relationships between flies of the Tephritinae subfamily (Diptera, Tephritidae) and their symbiotic bacteria. Mol. Phylogenet. Evol. 56, 312-326.

Mazzon, L., Martinez-Sañudo, I., Savio, C., Simonato, M., Squartini, A., 2011 - Stammerula and other symbiotic bacteria within the fruit flies inhabiting Asteraceae flowerheads In: Manipulative Tenants Bacteria Associated With Arthropods. Edited

by Zchori-Fein E. and Bourtzis K. CRC Press, Taylor and Francis Group, pp. 1-136. ISBN: 978-1-4398-2749-9.

Margulis, L., and Fester, R., 1991. Symbiosis as a Source of Evolutionary Innovation (Eds). Cambridge, MA: MIT Press.

Moran, N. and Telang, A., 1998. Bacteriocyte-associated symbionts of insects. Biosci. 48, 295–304.

Moran, N.A., and Baumann, P., 2000. Bacterial endosymbionts in animals. Curr. Opin. Microbiol. 3, 270–275.

Moran, N.A., and Wernegreen, J.J., 2000. Lifestyle evolution in symbiotic bacteria: insights from genomics. Trends Ecol. Evol. 15, 321–326.

Moran, N.A., Dale, C., Dunbar, H., Smith, W.A., Ochman, H., 2003. Intracellular symbionts of sharpshooters (Insecta: Hemiptera: Cicadellinae) form a distinct clade with a small genome. Environ. Microbiol. 5, 116–126.

Moran, N.A., Dunbar, H.E., 2006. Sexual acquisition of beneficial symbionts in aphids. Proc. Natl. Acad. Sci. USA. 103, 12803–12806.

Moran, N. A., McCutcheon, J. P., & Nakabachi, A., 2008. Genomics and evolution of heritable bacterial symbionts. Annu. Rev. Genet. 42, 165-190.

Munson, M.A., Baumann, P., Kinsey, M.G., 1991. *Buchnera* gen. nov. and *Buchnera aphidicola* sp. nov., a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids. Int. J. Syst. Bacteriol. 41, 566–568.

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H.E., Moran, N.A., Hattori, M., 2006. The 160-kilobase genome of the bacterial endosymbiont Carsonella. Science. 314, 267-267.

Nardon, P., and Grenier, A.M., 1988. Genetical and biochemical interactions between the host and its endocytobiotes in the weevils *Sitophilus* (Coleoptera, Curculionidae) and other related species. In: Scannerini, S., Smith, D.C., Bonfante-Fasolo, P., and Gianinazzi-Pearson, V. (Eds), Cell to Cell Signals in Plants, Animals and Microbial Symbiosis. NATO ASI Series H17. Berlin: Springer Verlag, pp. 255– 270.

Norrbom, A.L., and Kim K.C., 1988. A list of the reported host plants of the species of *Anastrepha* (Diptera: Tephritidae). U.S. Dep. Agric. Animal Plant Health Insp.

Serv. APHIS 81.

Norrbom, A.L., Carroll, L.E., Thompson, F.C., White, I.M., Freidberg, A., 1999. Status of knowledge. In: Thompson, F.C. (Eds.), Fruit fly expert identification system and systematic information database, Myia 9, ix., pp. 9–48.

Oliver, K.M., Degnan, P.H., Burke, G.R., Moran, N.A., 2010. Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. Annu. Rev. Entomol. 55, 247–266.

Parecer, S., and Ahmadjian, V., 2000. Symbiosis. An Introduction to Biological Associations. Oxford: Oxford University Press.

Petri, L. (1909). Ricerche Sopra i Batteri Intestinali della Mosca Olearia. Roma: Memorie della Regia Stazione di Patologia Vegetale di Roma.

Sacchetti, P., Granchietti, A., Landini, S., Viti, C., Giovannetti, L., and Belcari, A. (2008). Relationships between the olive fly and bacteria. *J. Appl. Entomol.* 132: 682–689.

Schröder, D., Deppisch, H., Obermayer, M., Krohne, G., Stackebrant, E., Hölldobler, E., Gross, R., 1996. Intracellular endosymbiotic bacteria of Camponotus species (carpenter ants): systematics, evolution and ultrastructural characterization. Mol. Microbiol. 21, 479–489.

Shigenobu, S., Watanabe, H., Hattori, M., Ishikawa, H., 2000. Genome sequence of the endocellular bacterial symbionts of aphids *Buchnera* sp. APS. Nature. 407, 81–86.

Stammer, H.J., 1929. Die bakteriensymbiose der trypetiden (Diptera). Zoomorphology. 15, 481–523.

Straw, N.A., 1989. Taxonomy, attack strategies and host relations in flowerhead Tephritidae: a review. Ecol. Entomol. 14, 455–462.

Tremblay, E., 1989. Coccoidea endosymbiosis. In: Schwemmler, W., and Gassner, G. (Eds), Insect Endocytobiosis: Morphology, Physiology, Genetics, Evolution. Boca Raton: CRC Press, pp. 145–173.

Wernegreen, J. J., 2002. Genome evolution in bacterial endosymbionts of insects. Nature Rev. Genet. 3, 850-861.

Werren, J.H., Baldo, L., Clark, M.E., 2008. *Wolbachia*: master manipulators of

invertebrate biology. Nat. Rev. Microbiol. 6, 741–751.

White, I.M., 1988. Handbooks for the identification of british insects. Tephritid flies, Diptera: Tephritidae, Vol. X, Part 5(a). Royal Entomological Society, London, pp. 1–134.

White, I.M., and Elson-Harris, M.M., 1992. Fruit Flies of Economic Significance: Their Identification and Bionomics. Wallingford, USA: CAB International.

White, I.M., 2006. Taxonomy of the Dacina (Diptera: Tephritidae) of Africa and the Middle East. Afr. Entomol. Mem. pp. 152.

Zwölfer, H., 1983. Life systems and strategies of resource exploitation in tephritids. In: Cavalloro, R., (Eds.). Fruit flies of economic importance (Eds.), Proceedings of the CEC/IOBC International Symposium, Athens, pp. 16–30.

Chapter II

Pattern of association between endemic Hawaiian fruit flies (Diptera, Tephritidae) and their symbiotic bacteria: evidence of cospeciation events and proposal of '*Candidatus* Stammerula trupaneae'

Submitted as:

E. Viale[a], I. Martinez-Sañudo[a], J.M. Brown[b], M. Simonato[a], A. Squartini[a], A. Bressan[c], M. Faccoli[a] , L. Mazzon[a]

[a] Department of Agronomy, Food, Natural Resources, Animals and the Environment (DAFNAE), University of Padova - Agripolis, Viale dell'Università, 16, 35020 Legnaro (Padova), Italy

[b] Department of Biology, Noyce Science Center Grinnell College Grinnell, IA, USA

[c] Bayer Crop Science LP, R&D Trait Research, 3500 Paramount Parkway, Morrisville, NC 27560, US.

I contributed to all parts of the experimental work and analysis, and to the paper writing.

Abstract

Several insect lineages have evolved mutualistic association with symbiotic bacteria. This is the case of some species of mealybugs, whiteflies, weevils, tsetse flies, cockroaches, termites, carpenter ants aphids and fruit flies. Some species of Tephritinae, the most specialized subfamily of fruit flies (Diptera: Tephritidae), harbour co-evolved vertically transmitted, bacterial symbionts in their midgut, known as "*Candidatus* Stammerula spp.". The 25 described endemic species of Hawaiian tephritids, plus at least three undescribed species, are taxonomically distributed among three genera: the cosmopolitan genus *Trupanea* (21 described spp.), the endemic genus *Phaeogramma* (2 spp.) and the Nearctic genus *Neotephritis* (2 spp.). We examined the presence of symbiotic bacteria in the endemic tephritids of the Hawaiian Islands, which represent a spectacular example of adaptive radiation, and tested the concordant evolution between host and symbiont phylogenies.

We detected through PCR assays the presence of specific symbiotic bacteria, designated as "*Candidatus* Stammerula trupaneae", from 35 individuals of 15 species,. The phylogeny of the insect host was reconstructed based on two regions of the mitochondrial DNA (16S rDNA and COI-tRNALeu-COII), while the bacterial 16S rRNA was used for the symbiont analysis. Host and symbiont phylogenies were then compared and evaluated for patterns of cophylogeny and strict cospeciation. Topological congruence between Hawaiian Tephritinae and their symbiotic bacteria phylogenies suggests a limited, but significant degree of host-symbiont cospeciation. We also explored the character reconstruction of three host traits, as island location, host lineage, and host tissue attacked, based on the symbiont phylogenies under the hypothesis of cospeciation.

Introduction

Adaptive radiations, are among the most spectacular processes in organismal evolution. This evolutionary process is amplified for island taxa, which display increased speciation rates and elevated levels of morphological diversity (Schluter, 2000; Jordan et al., 2003). The Hawaiian Islands are well known for having experienced radiations of plants and animals (Rubinoff and Schmitz, 2010; Roderick and Gillespie, 1998; Gillespie, 2004; Baldwin, 2006; Jordan et al., 2003; Goodmanet al., 2013). Over 89% of Hawaiian flowering plants (Wagneret al., 1999) and 66% of insect species (Zimmerman, 1981; Ziegler, 2002) are endemic to the archipelago, each arising from one or more independent colonization events. Due to their extreme isolation and unique geological history, along with the progressive island formation, the Hawaiian Islands provide an ideal setting for evolutionary studies (Roy et al., 2013; Roderick and Gillespie, 1998).

The Hawaiian Archipelago is currently composed of eight high islands arranged sequentially by age from the oldest, Kauai (5.1 My), located in the northwest corner of the Hawaiian chain, to the youngest, Hawaii (0.43 My), situated in the far south-east (Carson and Clague, 1995; Clague and Dalrymple, 1987) (Fig.1). The low eroded atolls located to the northwest of the main islands, known as the ‘‘leeward islands’’, range from about 7 to 28 My (Clague and Dalrymple, 1987; Fleischer, et al., 1998). Each island is currently isolated from the others by open ocean, although some of them have been connected during historical periods of lower sea levels. In particular, the islands comprising the Maui Nui complex (Molokai, Lanai, Maui, Kahoolawe) were connected during much of their histories (Price and Elliott-Fisk, 2004).

Several native Hawaiian insect genera have evolved following exponential speciation patterns (Ziegler, 2002). The Hawaiian Diptera, with 24 families containing over 10% of the known endemic taxa of the Hawaiian Islands (Nitta and O’Grady, 2008), offer an excellent opportunity for studies of evolution and speciation. Eight of these families contain radiations with more than ten species (Goodman et al., 2013). One of these families, the Tephritidae, commonly known as fruit flies, are primarily non-frugivorous species and predominantly infest Asteraceae, although some of them

have become important agricultural pests. The Tephritinae is considered the most specialized subfamily of tephritids containing over 200 genera and 1850 species from all zoogeographical regions (Foote et al., 1993; Norrbom, 1999; Korneyev, 1999).

Hawaiian endemic Tephritinae represent a good example of adaptive radiation, comprising 25 described species (Hardy and Delfinado, 1980) plus at least three undescribed species (Brown, unpublished data) that occupy a wide variety of habitats. Twenty of these species have a distribution restricted to a single island, while five are found on more than one island (multi-island species) (Messing et al. in prep). These tephritid species are taxonomically distributed among three genera: the cosmopolitan genus *Trupanea* (21 described spp.), the endemic genus *Phaeogramma* (2 spp.) and the Nearctic genus *Neotephritis* (2 spp.) (Hardy and Delfinado, 1980)

Endemic Hawaiian tephritids species feed as larvae on plant tissue of Asteraceae; most of them are seed feeders living in the flowerheads of native plants while a few are stem miners and gall formers. Host plants are represented by the endemic *Bidens, Lipochaeta,* and *Artemisia* spp. and the endemic Hawaiian silversword alliance, considered one of the best examples of adaptive radiation in plants (Carr 1985; Robichaux et al. 1990; Baldwin and Robichaux, 1995; Baldwin, 1997). The alliance, comprising more than 50 perennial species among the genus *Argyroxiphium, Dubautia,* and *Wilkesia* (Carr, 1985; Carr, 1998; Caraway et al., 2001) grows in a wide range of environments throughout the island chain, from rainforests and wet summit bogs to alpine and desert habitats including exposed lava and dry scrub (Caraway et al., 2001; Barrier et al., 1999).

We previously reported (Mazzon et al., 2008) the presence of a specific symbiotic bacterium, designated as *'Candidatus* Stammerula', in European representatives of seven genera of the subfamily Tephritinae (*Tephritis*, *Acanthiophilus*, *Capitites*, *Trupanea*, *Sphenella*, *Campiglossa* and *Oxyna*). Among these, the specific symbiotic bacteria of *Tephritis* spp. were designated as '*Candidatus* Stammerula tephritidis'. Stammerula bacteria are located in the first tract of the midgut, in contact with the epithelium but outside the peritrophic membrane. Symbionts are non-culturable and vertically transmitted during host reproduction (Stammer, 1929; Girolami, 1983; Mazzon et al. 2008). A further study (Mazzon et al. 2010) on the phylogenetic

relationships between the Palearctic Tephritinae and their symbiotic bacteria, showed a robust history of tandem diversification indicating that, despite their extracellular location, vertical transmission is the primary basis of the Tephritinae-symbiont concordant phylogeny.

The endemic Hawaiian Tephritinae provide an exceptional opportunity for investigating co-evolutionary hypothesis between host and their symbionts, as they are apparently a radiation from a single ancestor. We tested this hypothesis using a molecular phylogeny of 15 of the tephritid species based on two regions of the mitochondrial DNA (16S rDNA and COI-tRNALeu-COII) and a phylogeny of their corresponding symbionts, based on the bacterial 16S rRNA.

This approach aims at: (i) verifying the presence of specific symbiotic bacteria and studying their molecular relationship with symbionts of non-Hawaiian tephritids (*Candidatus* Stammerula); (ii) comparing endemic Hawaiian host flies with their symbiotic bacteria phylogenies and evaluating the extent of host-symbiont cospeciation.

Materials and methods

Insect sampling

A total of 35 individuals of 15 endemic species of Hawaiian tephritids were collected between August 1994 and October 2011 from across 5 of the main islands (Oahu, Maui, Kauai, Hawaii, Molokai) (Fig. 1 and Tab.1). For some specimens, flower heads, inspected for the presence of larvae or pupae, were kept in the laboratory for some days to allow larvae to maturate and pupate. Emerged adults were transferred into net cages and reared for a week in the laboratory before analysis. On the basis of our previous experience with symbionts of Tephritinae, rearing conditions do not alter the sequencing results (Mazzon et al., 2011). The remaining fly specimens came from the collection of the J. Brown, Biology Department, Grinnell College, AI, USA or were provided preserved in absolute alcohol by Prof. R. Messing (Department of Entomology, University of Hawaii, Hawaii, USA). All the adult samples were identified on a morphological basis using appropriate keys (Hardy and Delfinado, 1980).

A total of 13 Hawaiian endemic species of the *Trupanea* genus and two species of the endemic genus *Phaeogramma* were obtained and preserved in 100% ethanol at -80°C until DNA analysis. For the bacterial symbiont analysis, flies were dissected to extract the midguts hosting symbionts, following the procedure described in (Mazzon et al., 2008). The rest of the abdomen was used for the insect host analysis.

Insect host analyses: DNA extraction, amplification and sequencing

For DNA extractions, the abdomens of the flies were macerated in a 1.5 ml microcentrifuge tube with a micro pestle, and nucleic acids were extracted using the salting-out protocol (Patwary et al. 1994) or the Qiagen DNeasy kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. When possible, two or three samples for each species, host plant and location were processed. In order to assess the quality of the DNA, aliquots from extracted samples were separated in a 1% agarose gel and viewed under UV after staining with SYBR Safe DNA gel stain (Invitrogen).

Two regions of the mitochondrial DNA of the insect host, a fragment of 16S rRNA and a fragment including the 3' region of cytochrome oxidase sub-unit I (COI), tRNA-Leu and the 5' region of cytochrome oxidase sub-unit II (COII), were amplified. The 16S rRNA gene fragment of the insect host was amplified using the following two pairs of primers: DFl2 (forward: 5'-GATTTATAGGGTCTTCTCGTC-3') and DR (revers: 5'-GATGTACCGGAAGGTGTATCT-3') (Mazzon et al. 2010) andN1-J12261m (forward: 5'-TACTTCGTAAGAAATTGTTTGAGC-3') and LRN13398 (reverse: 5'-CGCCTGTTTAACAAAAACAT-3') (Simon, 1994; Simon et al., 2006). Additionally, the fragment of mitochondrial COI-tRNALeu-COII was amplified using two pairs of primers (Simon, 1994; Simon et al., 2006): C1-J-2195 (forward: 5'-TTGATTTTTTGGTCATCCAGAAGT-3'), TKN3796 (reverse: 5'-ACTATAAAATGGTTTAAGAG-3'). For the specimens with high DNA degradation, shorter segments were amplified using another pair of primers: C1-J-2183 (forward: 5'-CAACATTTATTTTGATTTTTTGG-3') and TL2-N-3014 (reverse: 5'-TCCATTGCACTAATCTGCCATATTA-3').

All PCR reactions were performed in a 20 µl volume containing 4 µl PCR of 5x colorless GoTaq Flexi Buffer (Promega), 2.5 mM MgCl, 100 mM dNTPs, 0.5 µM of each primer, 1U of GoTaq Flexi DNA polymerase (Promega) and 2 µl of extracted DNA. All target regions were amplified using an Eppendorf Mastercycler Gradient (Applied Biosystems, Foster City, CA, USA). The standard thermal profile for the amplification of the 16s rRNA gene was: hold for 5 min at 96 °C, 35 cycles of 96 °C for 50s, an annealing step ranging between 52 °C and 60 °C for 50s, 72 °C for 1 min, and extension for 5 min at 72°C. For the COI-tRNA-Leu-COII amplification, the following thermocycling profile was used: hold for 5 min at 96 °C, 35 cycles of 96 °C for 1 min, an annealing step ranging between 50 °C and 60 °C for 1 min, 72 °C for 2 min and extension for 5 min at 72°C. The amplified products were examined by gel electrophoresis on 1% agarose gel. PCR-amplified products were purified with a mix of Exonuclease and Antarctic Phosphatase (New England Biolabs) and sequenced at the BMR Genomics service at Padova, Italy. All host sequences obtained in this study have been deposited in GenBank under the accession numbers listed in Table 1.

Phylogenetic analyses of the insects

Sequences generated were inspected and aligned using MEGA 5.2.2. (Tamura et al., 2007). Identical insect species sequences, harbouring the same symbiont, were excluded from the analysis. Phylogenetic analyses of the individual 16S rRNA gene were conducted on the conserved sequence blocks filtered by Gblocks (Castresana, 2000) in order to remove ambiguous portions from the alignment.

Sequences of six Italian tephritids from Mazzon et al. (2010) were incorporated into the analyses as outgroups. Additionally, two Californian *Trupanea* spp. sequences were added to the 16S rRNA data (Tab. 2).

Phylogenetics using data combined from different gene regions generally are better resolved than those based on individual partitions, as long as they are congruent (Baker and DeSalle, 1997; Wiens, 1998, Jordan et al., 2003). A partition homogeneity test was performed for the 16S and COI-tRNALeu-COII gene trees using Paup*4b10 (Swofford, 2003). The test confirmed that these regions contained homogeneous signals ($p = 0.26$), allowing data to be pooled for further analyses.

Phylogenetic relationships were estimated, in separate and combined data sets, using two methods: approximate maximum-likelihood (ML) and Bayesian inference (BI) analysis. The most appropriate substitution model for each data set was determined by the Akaike Information Criterion (AIC) (Posada and Buckley, 2004) using the program jModeltest v2.1.4 (Posada, 2008). The approximate ML analysis was performed using PHYML v2.4.4 software (Guindon and Gascuel, 2003) with neighbour-joining starting trees. Branch support in the ML trees was tested by means of 1000 bootstrap replicates. We also used a Bayesian Markov Chain Monte Carlo (MCMC) analysis implemented in MrBayes v3.2 (Ronquist and Huelsenbeck, 2003) Partitions were assigned according to the gene region. Posterior probabilities (Pp) were calculated using four chains for 10 million generations in two independent runs, sampling every 100 generations and convergence was maximized by ensuring the average standard deviation of split frequencies fell below 0.001. Trees were sampled every 500 generations. The program Tracer v.1.5 (Drummond and Rambaut, 2007) was used to check for stationarity, and the burn-in value for obtaining a 50% majority rule consensus tree was set to ignore the first 25% of trees, to include trees only after

stationarity was reached. The 50% majority rule consensus trees were viewed with the program FigTree v.1.3.1 (Rambaut and Drummond, 2008).

Symbiont analyses: Bacterial DNA amplification and sequencing

The bacterial DNA was analyzed by amplifying a fragment of 16S rRNA gene using the universal primers: fD1 (forward: 5'-AGAGTTTGATCCTGGCTCAG-3'), rP1 (reverse: 5'-ACGGTTACCTTGTTACGACTT-3') (Weisburg et al. 1991) and specific primers designed for this study 10F (forward: 5'-AGTTTGATCATGGCTCAGATTG-3'), 510R (reverse: 5'-CTACGCATTTCACCGCTACACC-3') and 1507R (reverse: 5'-TACCTTGTTACGACTTCACCCCAG-3').

PCR reactions were performed with the same volumes used for the insect host DNA amplification. The cycling program for the bacterial 16S rRNA gene included an initial denaturation at 95 °C for 2 min, 35 cycles at 96 °C for 30s, 56 °C for 30s, 72 °C for 90s and a final extension at 72 °C for 10 min.

PCR amplified products were purified and sequenced as performed for the insect host. GenBank accession numbers of symbiont sequences are listed in Table 1.

Symbiont phylogenetic analyses

Bacterial 16S rRNA gene sequences were obtained and aligned following the same method described above for the insect host analysis. For phylogenetic analysis, identical haplotypes were merged in a single representative haplotype. The analyses were run including in the dataset 21 non-Hawaiian symbiont sequences from GenBank, as outgroups, reported in Tab. 2.

Bacterial phylogenies were reconstructed using both approximate maximum-likelihood (ML) and Bayesian inference (BI) analysis. The AIC, implemented in jModeltest, v2.1.4 (Posada, 2008) was used to estimate the best-fit model of substitution.

Host-symbiont cophylogeny analyses

To examine the congruence between the insect host and its symbiotic bacteria

phylogenies two methods of coevolutionary analyses were applied: a tree-based method, implemented in the program Jane 4.0 (Conow et al., 2010) and distance-based method, Parafit (Legendre et al., 2002) implemented in CopyCat (Meier-Kolthoff et al., 2007). Cophylogenetic analyses were performed using a reduced data set including *Trupanea amoena* and Hawaiian insects sequences with their corresponding symbionts.

The software Jane 4.0 uses a polynomial time dynamic programming algorithm in conjunction with a genetic algorithm to compare the two tree topologies by optimally mapping the parasite tree onto the host tree using different event costs to find very good, and often optimal, solutions to reconcile the two phylogenetic trees (Conow et al., 2010; Mendlová et al., 2012). Moreover, Jane 4.0 supports multihost parasites and multiparasite hosts. Two commonly used models with different event cost schemes were evaluated (Conow et al., 2010; Koch et al., 2013; Mendlová et al., 2012). Setting 1 assuming no cost for cospeciation and cost = 1 for all other events (cospeciation = 0, duplication = 1, duplication & host switch = 2, loss = 1, failure to diverge = 1). Setting 2 assuming cost = 1 for all events (cospeciation = 1, duplication = 1, duplication & host switch = 2, loss = 1, failure to diverge = 1). The analysis was performed with 500 generations and a population size of 100 as parameters of the genetic algorithm. Moreover, significant matching of host and bacterial phylogenies was assessed by computing the costs of 500 replicates with random tip mapping and comparing the resulting costs to the cost of the original associations.

The distance-based method ParaFit (Legendre, 2001; 2002) was used to test the null hypothesis of random association between host and parasite data sets. Parafit software is a useful tool because it can accommodate poorly resolved topologies, multiple hosts per symbiont lineage and, as in our case, multiple symbionts per host lineage (Light and Hafner, 2007). Distance matrices for host and bacteria were derived from Kimura 2-parameter estimates of pairwise genetic distances, using Mega 5.2.2. The program DistPCoA (Legendre and Anderson, 1998) was used to convert distance matrices into principal coordinate matrices allowing the software Parafit to calculate the probability of host-symbiont coevolution. Tests of random association were performed with 9999

permutations globally across both matrices and for each individual host-parasite association.

Evolution of ancestral characters

We reconstructed the evolution of host characters (island location, host lineage, and host tissue attacked) on the symbiont tree using stochastic character mapping (Huelsenbeck et al., 2003), as implemented in the R program *phytools* v. 0.4-05 (Revell, 2012). This approach produces a posterior distribution of character states at each node under the model, calculated from 200 simulations using a model of symmetrical rates of change between states.

Results

Sequence analyses of insect

Sequence data for a total of 35 Hawaiian Tephritinae specimens was collated, comprising all the 13 collected species of the genus *Trupanea* and the 2 species of the genus *Phaeogramma.* From the mtDNA amplification and sequencing, portions of 988 bp and 1443 bp on average for the 16S rRNA and COI-tRNALeu-COII gene respectively, were obtained. The final alignment of 16S rRNA data set, previously filtered with Gblocks, was composed of 685 bp. The partitioned COI-tRNALeu-COII data set, including outgroups, consisted of 1350 nucleotides. From the combination of the two datasets, the resulting concatenated data was 1972 bp long.

Phylogeny of Hawaiian tephritids

Analyses were performed on both individual gene fragments and the combined data set, using a general time-reversible model with gamma distribution and number of invariant sites (GTR+I+G) as the best fit-evolutionary model determined by jModeltest. The likelihood score and the I and G parameters were: -InL = 1924.09, I=0.5630 and G= 0.7160 for the 16s rRNA gene and -InL = 5606.82, I= 0.4640 and G= 0.8650for the COI-tRNALeu-COII gene. In the combined analysis the likelihood score of –InL was 7364.88 and parameters of I and G were 0.5320 and 0.8120 respectively. The trees inferred separately from each gene portion (available as supplementary data, Fig. S1 and Fig.S2) were roughly concordant among ML and BI analyses. Compared to the COI-tRNALeu-COII tree, the phylogeny obtained for the 16S rRNA gene appears to be less informative as it displayed poor resolution especially at deeper nodes. In the 16S gene tree, 14 nodes in the ML and 13 nodes in the BI analysis were recognized as statistically supported, while in the COI-tRNALeu-COII tree 17 nodes for both ML and Bayesian inference had statistical support.

Similarly to the COI-tRNALeu-COII phylogeny, the partitioned concatenated analyses produced the same topology in both methods of reconstruction (ML and BI), with high branch support values (Fig. 2). Twenty nodes with high statistical support were recognized in the ML tree, and 17 nodes in the Bayesian inference.

The phylogeny of the combined data provides evidence for the existence of a monophyletic well-supported clade (Bs/Pp=96/100), corresponding to the endemic Hawaiian taxa. The clade composed of Hawaiian insect hosts shows in turn a clear subdivision into 8 distinct clades, supported by both methods, although there was less support for the relationships among them (Fig. 2). These clades are also found in a phylogenetic analysis using more fly taxa and an additional nuclear gene (Brown et al., in preparation.)

Presence of endosymbionts and sequence analyses

We successfully sequenced 1305 aligned bp of the mitochondrial 16S rRNA symbionts from all of the midgut contents of the 35samples (15 species). BLAST analysis revealed 98-99% identity of all symbiont sequences analyzed to *Ca. Stammerula sp.* recovered in the non-Hawaiian (from Italy) *Trupanea amoena* and *Trupanea stellata*.

In this study 15 endemic Hawaiian species were analysed. A specific and unique bacterial symbiont was identified in 9 species (*P. "newsp.", P. vittipennis, T. artemisiae, T. arboreae, T. cratericola, T. denotata, T. perkinsi, T. "waikamoi"* and *T. "pohakuloa*"). In addition, in 3 multi-island species (*T. dubautia, T. joycei* and *T. swezeyi*), collected in more than one island, a specific symbiont haplotype was found in every island population (Fig. 4/B). Moreover, two species, *T. beardsleyi* and *T. crassipes*, both collected on a *Dubautia* spp. in close proximity on Maui, shared identical symbiont sequences. In contrast, *T.pantosticta* haplotypes from Hawaii harboured different bacterial sequences from single samples from different locations and fly host species (*Dubatia scabra* and *D. ciliata*).

Phylogeny of symbionts

Phylogenetic analyses were performed using ML and BI methods both based on the HKI+I+G model of evolution indicated by jModeltest, with a likelihood score of -InL= 2953.09 and parameters of I= 0.7580 and G=0.6620. Tree topologies inferred by the two approaches were similar, showing good level of support at deeper nodes, but low support level for more recent divergence events. Twenty-six nodes in the ML and 28

nodes in the BI analysis were recognized as statistically supported.

The phylogenetic trees, including the Hawaiian and European species used in our previous work, showed the symbiotic bacteria of all the endemic Hawaiian tephritids included in a well-supported clade (Bs/Pp=100/100), represented by *Candidatus* Stammerula sp. (Mazzon et al., 2008). Within this clade, in turn, the Hawaiian symbionts are grouped in a monophyletic statistically supported clade (Bs/Pp=75/100) with symbionts of the European *T. stellata, T. amoena* and *Capitites ramulosa* (*Ca.* Stammerula sp.) supported as sister to them (Bs/Pp=99/100)(Fig. 3). Within the Hawaiian clade, the phylogenetic relationships remain poorly resolved due to the lack of sequence divergence. However, 3 interspecific subclades (III, IV and V) were supported by at least one of the two clustering methods used (ML and BI) (Fig. 3), in each case in species inhabiting hosts that are also grouped in the fly phylogeny (Fig. 2). Moreover, 2 intraspecific subclades (I and II) including symbionts of *T.joycei* and *Phaeogramma* spp. were recognized as statistically supported.

Host-symbiont cophylogenetic analyses

To assess the coevolution between hosts and symbionts, we compared the 25 insect Hawaiian sequences and their corresponding symbiotic bacteria. In three cases, there were multiple hosts (species or haplotypes) for a symbiont lineage (Fig. 4). Both tools used for determining the history of the cophylogeny between symbionts and their hosts suggests a significant cophylogeny between the host and their bacteria.

In Jane 4.0 analysis, two common models of event cost settings were used. Four failure-to-diverge and 12 co-speciation events were detected for both the first (0, 1, 1, 2, 1) and the second (1, 1, 1, 2, 1) settings. Although the tanglegram (Fig. 4) does not show full congruence between the host and parasite topologies, the reconciliation analysis indicates that there are more cospeciation events than expected by chance ($P<0.05$) thus indicating a significant, if not perfect, congruence between the symbiont and host tree. Examples of possible reconstructions are shown in Fig. S3 and S4.

ParaFit analysis compares patristic distance between hosts and their corresponding symbionts, to test the global fit between the two data sets. Furthermore, the method assesses if each individual host-symbiont association (link) significantly contributes to

the global fit, to evaluate which ones have a structuring effect. The global test indicated a significant congruence between insect hosts and their symbiotic bacteria (P = 0.001). However, the test of individual links showed that not all host-symbiont associations significantly contributed to this global fit. In addition to *T. amoena,* 4 out of 25 individual links were significant ($P < 0.05$) suggesting their structuring role in the global congruence (Tab. 3).

Character reconstruction

Character reconstruction of host traits (island location, host lineage, and host tissue attacked) on the symbiont phylogeny should reflect host evolution, under the hypothesis of cospeciation. Island (or island complex, in the case of Maui Nui islands) demonstrates highly uncertain ancestral states and many island shifts (Fig. 5/A), which reflects the role that island barriers play in species formation in the flies (Brown et al., 2006; Brown et al. in preparation). In contrast, host plant lineage and tissue attacked are both more conservative and show more certain ancestral state reconstruction (Figs. 5/B and 5/C), which mirrors the results obtained when using a more complete fly phylogeny (Brown et al., in preparation).

Discussion

The present study is the first to confirm the presence of symbiotic bacteria in non-frugivourus fruit flies endemic to the Hawaiian Islands. Moreover, we examined whether cospeciation has occurred between this host/symbiont association. To accomplish this goal, we have reconstructed molecular phylogenies of both host insects and their symbionts. We also explored the insect evolution of three ancestral characters as island location, host lineage, and host tissue attacked, based on the symbiont phylogenies, which should reflect host evolution, under the hypothesis of cospeciation.

As in Brown et al. (2006), our results support the monophyly of all the Hawaiian tephritids, including the cosmopolitan genus *Trupanea* and the endemic genus *Phaeogramma* (Fig. 2). This suggests that adaptive radiation of this group likely occurred within the archipelago from a single colonizing ancestor. In addition, the presence of symbiotic bacteria was detected in all endemic Hawaiian Tephritinae tested confirming previous works (Mazzon et al, 2008) that has regarded this association as an obligate symbiosis. This has allowed us to test the proposed model of speciation in a relatively simple and recent radiation.

'Candidatus Stammerula trupaneae'

The detected symbionts were included in the monophyletic clade represented by '*Candidatus* Stammerula' described by Mazzon et al. (2008) as symbiont of some members of Tephritinae subfamily. Furthermore, these Hawaiian symbionts were grouped in a well-supported subclade with *Stammerula* of non-Hawaiian *Trupanea* spp. (*T. stellata* and *T. amoena*) (Fig. 3). On the basis of these results, we propose the designation of a candidate species '*Candidatus* Stammerula trupaneae' to include symbionts of the *Trupanea* and *Phaeogramma* genera.

These bacteria are characterized by straight rod-shaped cells of 2–3 mm (sometimes giving origin to long chains). They are negative for Gram staining and unculturable on microbiological media. They are located exclusively in association with its host species within the following structures: extraperitrophic space of the midgut.

Molecular relationships among Hawaiian symbionts

Our results showed the symbiotic bacteria of all the endemic Hawaiian tephritids are included in a monophyletic and well-supported group. However, the phylogenetic relationships within the well-supported Hawaiian symbionts clade remain poorly resolved (Fig. 3). The lack of resolution in the internal nodes suggests a rapid and recent radiation of the fly hosts and/or slow evolution of this gene in the symbiont. Indeed, statistically supported values were obtained by both Maximum likelihood and Bayesian analyses only in symbiont haplotypes of the same insect host species. One exception is represented by *Phaeogramma* species, although they are themselves closely related. Symbionts of the three multi-island species (*T. joycei*, *T. dubautia*, and *T. swezeyi*), sampled in different islands, show a genetic diversity related to the island of origin, as observed in the insect phylogeny. For example, in clade I, the two symbiont haplotypes of *T. joycei* from Maui and Molokai mirror the origin of the fly host. Moreover, within the clade IV, the three well-differentiated haplotypes of *T. dubautiae* related to the three different islands (Maui, Oahu and Kauai). Thus, the bacterial haplotypes seem to be correlated to the hosts and in turn by their distribution. The conservatism of plant host lineage and plant host tissue evolution, reconstructed on the phylogeny of the symbiont, also suggests tight fly-symbiont coevolution, as these characters are conservative over the fly lineage (Brown et al., in preparation).

Concordant evolution of Hawaiian Tephritinae and their symbiotic bacteria

Symbiont evolution is greatly influenced by host evolution, and close congruence of host and symbiont phylogenies can indicate cospeciation (Fahrenholz, 1913). In this study, both cophylogenetic analyses suggested a high degree of congruence, even if not perfect, between endemic Hawaiian Tephritid hosts and symbionts. The tree-based reconciliation analysis (Jane 4) indicated that statistically more cospeciation events have occurred than expected by chance suggesting a limited but significant degree of host-symbiont cospeciation (Fig. 4). The distance-based method (ParaFit) software reveals a global significant non-random association pattern between host and symbiont datasets, although the test of individual links indicates that some symbiotic bacterial species did not cospeciate with their hosts. The lack of strict congruence of this host-

symbiont association could be the result of a mix of coevolutionary independent events such as host switching, duplications and sorting events.

Multiple Hawaiian hosts were found to harbor the same symbiotic bacteria. Identical symbiont sequences were identified, respectively, in different haplotypes of *T. arboreae* (Hawaii) and *T. denotata* (Maui). This could be explained by the lower evolutionary rates exhibited by symbionts respect to their hosts (Law and Lewis 1983; Law 1985). Moreover, two phylogenetically distant insect species, *T. beardsleyi* and *T. crassipes*, both collected on Maui, share the same symbiont. Since they feed as larvae on seedheads of the same *Dubautia* hosts, the presence of this identical symbiont could result from a horizontal transfer. Indeed as described by Girolami (1973) and Mazzon et al. (2008), the extracellular condition of these symbionts in larvae, adults and eggs are important triggers for contacts with the outer environment. For these reasons, symbiotic bacteria could be susceptible to accidental horizontal transfers, invasion, and replacement by foreign bacteria. Nevertheless, host-symbiont co-speciation suggests that vertical transmission appears to be the primary force shaping the topological congruence between insect and bacterial phylogenies.

Conclusion

In conclusion, our study revealed the presence of symbiotic bacteria in all endemic Hawaiian non-frugivourus fruit flies under the designation "*Candidatus* Stammerula trupaneae". The symbiont and their host phylogenies exhibit some congruence and are statistically cospeciating. However, strict cospeciation is not occurring. Future phylogenetic studies, including a larger sample host size, in particular multi-island and oligophagous species, could increase our knowledge of the host-symbiont coevolutionary events. Moreover the genetic characterization of these symbionts could be a useful tool to better understand the mechanism responsible for Hawaiian tephritid flies radiation.

Acknowledgments

We would like to thank Prof Russel H. Messing (Kauai Agricultural Research Center, Department of Entomology, University of Hawaii, Hawaii, USA) for valuable field assistance in Kauai and Forest Starr and Kim Starr for providing most of the tephritids species endemic to Hawaii analyzed from their collection. The Department of Agronomy, Food, Natural Resources, Animals and the Environment, University of Padova in collaboration with Alberto Bressan from the Department of Plant and Environmental Protection Sciences, College of Tropical Agriculture and Human Resources**,** University of Hawaii, provided financial support for this research.

Tab. 1 Collection details for specimens examined and host plant associations. Fh f: Flowerhead feeder; St f: Shoot tip feeder; Sg m: Stem gall maker. S: single-island and M: multi-island species. Quotation marks indicate non-official name of undescribed species.

Taxon and Feeding mode	Origin	Host plant	Coll. date	GeneBank Accession No		
				Insect:		Symbiont:
				COItRNACIII	16S rRNA	16S rRNA
Trupanea						
arboreae (M-Fh f)	Ha: Mauna Kea State Park	*D. linearis*	Nov/06	KM079219	KM079196	KM079241
-	Ha: Mauna Loa Trail	*D. ciliata*	Feb/03	KM079220	-	-
-	Ha: Kuhuku, Havo	*Arg. kaunse*	Aug/11	-	-	-
artemisiae (M-Fh f)	Ma: Gully west of Kapal.	*Art. mauiensis*	Aug/11	KM079221	KM079197	KM079242
-	Ma: Gully west of Kapal.	*Art. mauiensis*	Aug/11	-	-	-
beardsleyi (S-Fh f)	Ma: Behind Kapalaoa Cabin	*Dubautia*	Aug/11	KM079238	KM079198	KM079243
-	Ma: Behind Kapalaoa Cabin	*Dubautia*	Aug/11	-	-	-
crassipes (M-Fh f)	Ma: South Slope, Hale	*D. menziesii*	Jun/11	KM079222	KM079199	KM079244
-	Ma: South Slope, Hale	*D. menziesii*	Jun/11	KM079223	KM079200	-
cratericola (M-Fh f)	Ma: Holua Hilton	*D. menziesii*	Jul/11	KM079224	KM079201	KM079245
-	Ma: Holua Hilton	*D. menziesii*	Jul/11	-	-	-
denotata (S-St f)	Ma: W slope, Pohak. Gulch	*D. platyphylla*	Aug/94	KM079225	KM079202	KM079246
-	Ma: W slope, Pohak. Gulch	*D. platyphylla*	Aug/94	-	KM079203	-
dubautiae (M-Fh f)	Ma: Pu'u Kukui	*D. laxa*	Aug/11	KM079227	KM079205	KM079248
-	Ka: Alakai swamp trail	*D. laxa*	Oct/11	KM079226	KM079204	KM079247
-	Ka: Alakai swamp trail	*D. laxa*	Oct/11	-	-	-
-	Oa: Ewa Forest Reserve	*D. laxa*	Oct/06	KM079228	KM079206	KM079249
-	Oa: Ewa Forest Reserve	*D. laxa*	Oct/06	-	-	-
joycei (M-Sg m)	Ma: Puu kukui	*D. laxa*	May/02	KM079230	KM079208	KM079251
-	Ma: Puu kukui	*D. laxa*	May/02	-	-	-
-	Mo: Pu'u Kolekole	*D. laxa*	May/98	KM079231	KM079209	KM079252
-	Mo: Pu'u Kolekole	*D. laxa*	May/98	-	-	-
-	Ka: Alakai swampt trail	*D. paleata*	Jun/98	KM079229	KM079207	KM079250
pantosticta (M-St f)	Ha: Mauna Loa Trail	*D. ciliolata*	Jun/98	KM079232	KM079210	KM079253
-	Ha: mile 29 Saddle road	*D. scabra*	Jun/98	-	KM079211	KM079254
perkinsi (S-Fh f)	Ka: Kahili Ridge Trail	*B. forbesii*	Feb/02	KM079233	KM079212	KM079255
"pohakuloa" (S-Fh f)	Ha: Mauna Kea SP	*D. linearis*	Oct/11	KM079234	KM079213	KM079256
swezeyi (M-St f)	Ka: Alakai Swamp Trail	*D. paleata*	Jun/98	KM079235	KM079214	KM079257
-	Ka: Alakai Swamp Trail	*D. paleata*	Jun/02	-	-	-
-	Ma: Pu'u Kukui	*A. grayanum*	May/02	KM079236	KM079215	KM079258
-	Ma: Pu'u Kukui	*Arg. grayanum*	May/02	-	-	-
"waikamoi" (S-Sg m)	Ma: Waik. TNC, Ko'olau Gap	*D. reticulata*	Apr/98	KM079237	KM079216	KM079259
Phaeogramma						
"new sp." (M-Sg m)	Ka: Black Pipe Trail, Koke'e	*B. forbesii*	Jun/98	KM079217	KM079194	KM079239
vittipennis (M-Sg m)	Mo: O'oa	*B. menziesii*	Jun/98	KM079218	KM079195	KM079240
-	Mo: O'oa	*B. menziesii*	Jun/98	-	-	-

Tab. 2 List of the sequences retrieved from GeneBank used in insect and symbiont analyses. GeneBank Accession Nos for insect host (COI-tRNA-COII and 16S rRNA) and symbiont (16S rRNA) sequences and corresponding references.

Taxon	**GeneBank Accession No**		**Reference**
Insect	**COI-tRNA-COII**	**16S rRNA**	
Acanthiophilus helianthi	GQ175795	GQ175832	Mazzon *et al.* (2010)
Capitites ramulosa	GQ175803	GQ175840	Mazzon *et al.* (2010)
Tephritis arnicae	GQ175793	GQ175830	Mazzon *et al.* (2010)
Trupanea amoena	GQ175796	GQ175833	Mazzon *et al.* (2010)
Trupanea stellata	GQ175797	GQ175834	Mazzon *et al.* (2010)
Trupanea actinoba	-	DQ471411.1	Han *et al*. (2006)
Trupanea conjunta	-	DQ471413.1	Han *et al.* (2006)
Outgroup			
Bactrocera oleae	GQ175825	GQ175862	Mazzon *et al.* (2010)
Bacteria	**16S rRNA**		
Erwinia dacicola of *Bactrocera oleae*	AJ586620		Capuzzo *et al*. (2005)
Erwinia persicina	AM184098		Randazzo *et al.* (2009)
Stammerula sp. of *Capitites ramulosa*	EF469628		Mazzon *et al*. (2008)
Stammerula sp. of *Trupanea amoena*	EF469626		Mazzon *et al*. (2008)
Stammerula sp. of *Trupanea stellata*	EF469627		Mazzon *et al*. (2008)
Stammerula sp. of *Campiglossa doronici*	EF469636		Mazzon *et al*. (2008)
Stammerula sp. of *Sphenella marginata*	EF469629		Mazzon *et al*. (2008)
Stammerula sp. of *Acanthiophilus helianthi*	EF469625		Mazzon *et al*. (2008)
Stammerula sp. of *Oxyna flavipennis*	EF469630		Mazzon *et al*. (2008)
Stammerula tephritidis of *T. arnicae*	EF469616		Mazzon *et al*. (2008)
Stammerula tephritidis of *T. bardanae*	EF469617		Mazzon *et al*. (2008)
Stammerula tephritidis of *T. cometa*	EF469615		Mazzon *et al*. (2008)
Stammerula tephritidis of *T. divisa*	EF469619		Mazzon *et al*. (2008)
Stammerula tephritidis of *T. matricariae*	EF469623		Mazzon *et al*. (2008)
Symb. of *Campiglossa guttella*	EF469637		Mazzon *et al*. (2008)
Symb. of *Dioxyna bidentis*	EF469631		Mazzon *et al*. (2008)
Symb. of *Noeeta bisetosa*	EF469632		Mazzon *et al*. (2008)
Symb. of *Noeeta pupillata* _A	EF469633		Mazzon *et al*. (2008)
Symb. of *Noeeta pupillata*_C	EF469635		Mazzon *et al*. (2008)
Outgroups			
Escherichia coli	AY616658		Siddiqui *et al.* (2006)
Klebsiella oxytoca	DQ444288		Hao *et al.* (2008)

Tab. 3 Results from ParaFit analysis. The test was used to assess the null hypothesis of independent evolution of hosts and symbionts and to test the significance of each host-symbiont link (25 host-symbiont links) using ParaFitLink1 and ParaFitLink2 statistics. Probabilities are computed after 9999 random permutations. The global test probability was 0.00110.

Symbiont	**Host**	**Island**	**ParaFitLink1**	**ParaFitLink2**
Symbiont of *T. arboreae*	*T. arboreae*	Hawaii	0.157	0.133
Symbiont of *T. arboreae*	*T. arboreae*	Hawaii	0.147	0.122
Symbiont of *T. artemisiae*	*T. artemisiae*	Maui	0.399	0.352
Symbiont of *T. beardsleyi*	*T. beardsleyi*	Maui	0.523	0.493
Symbiont of *T. crassipes*	*T. crassipes*	Maui	0.618	0.598
Symbiont of *T. crassipes*	*T. crassipes*	Maui	0.198	0.171
Symbiont of *T. cratericola*	*T. cratericola*	Maui	0.102	0.082
Symbiont of *T. denotata*	*T. denotata*	Maui	0.114	0.097
Symbiont of *T. denotata*	*T. denotata*	Maui	0.126	0.102
Symbiont of *T. dubautiae*	*T. dubautiae*	Kauai	0.177	0.156
Symbiont of *T. dubautiae*	*T. dubautiae*	Maui	0.137	0.117
Symbiont of *T. dubautiae*	*T. dubautiae*	Oahu	0.144	0.126
Symbiont of *T. joycei*	*T. joycei*	Kauai	0.159	0.144
Symbiont of *T. joycei*	*T. joycei*	Maui	0.153	0.135
Symbiont of *T. joycei*	*T. joycei*	Molokai	0.164	0.146
Symbiont of *T. pantosticta*	*T. pantosticta*	Hawaii	0.012*	0.005*
Symbiont of *T. pantosticta*	*T. pantosticta*	Hawaii	0.005*	0.003*
Symbiont of *T. perkinsi*	*T. perkinsi*	Kauai	0.291	0.246
Symbiont of *T. "pohakuloa"*	*T. "pohakuloa"*	Hawaii	0.216	0.185
Symbiont of *T. swezeyi*	*T. swezeyi*	Kauai	0.135	0.115
Symbiont of *T. swezeyi*	*T. swezeyi*	Maui	0.086	0.070
Symbiont of *T. "waikamoi"*	*T. "waikamoi"*	Maui	0.468	0.416
Symbiont of *P. "newsp."*	*P."newsp."*	Kauai	0.002*	0.001*
Symbiont of *P. vittipennis*	*P. vittipennis*	Molokai	0.001*	0.001*
Stammerula sp. of *T. amoena*	*T. amoena*	-	0.002*	0.001*

* Significant association ($P < 0.05$).

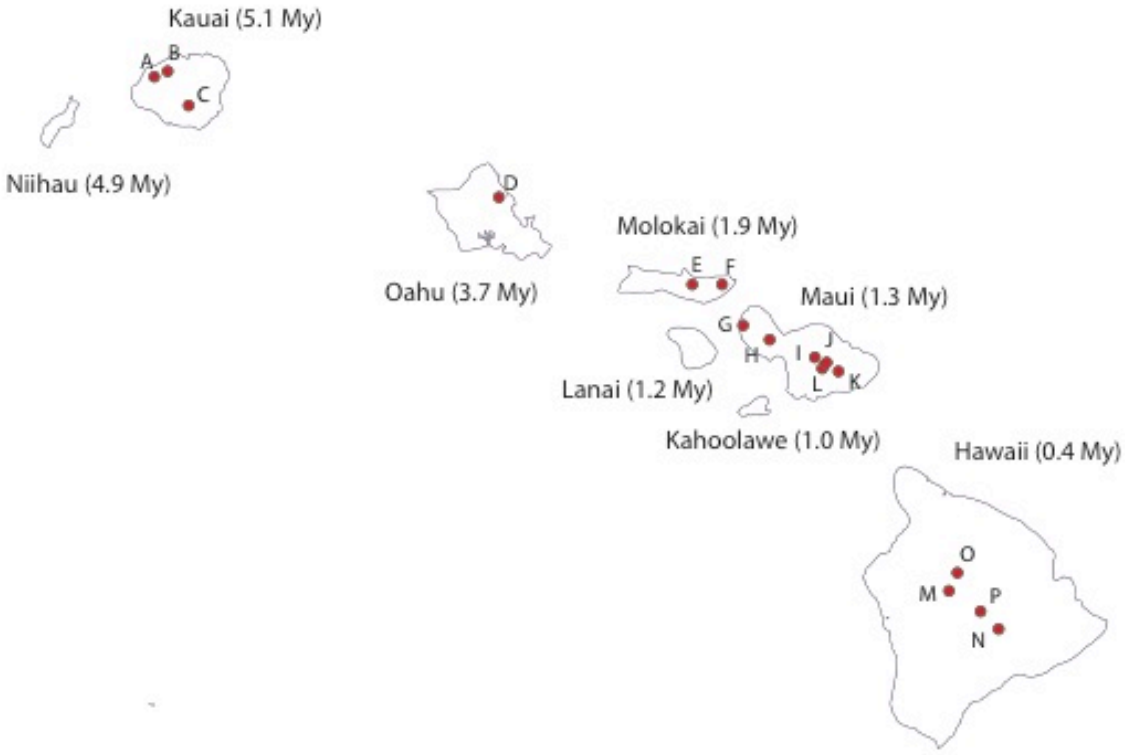

Fig. 1 Map of the Hawaiian Island Chain shows the collection site of each endemic tephritids species analyzed. A - Alakai swamp trail B -Black Pipe Trail, Koke'e, C - Kahili Ridge Trail D, - Ewa Forest Reserve, E - O'oa, F - Pu'u Kolekole, G - South Slope, Hale H - Pu'u kukui, I - Waikamoi TNC, Ko'olau Gap, J - Holua Hilton Haleakala NP, K - Kapalaoa Cabin, L – Pohaku Akala Gulch, Haleakala, M - mile 29 Saddle Road, N - Kuhuku, Havo, O - Mauna Kea State Park, P - Mauna Loa Trail.

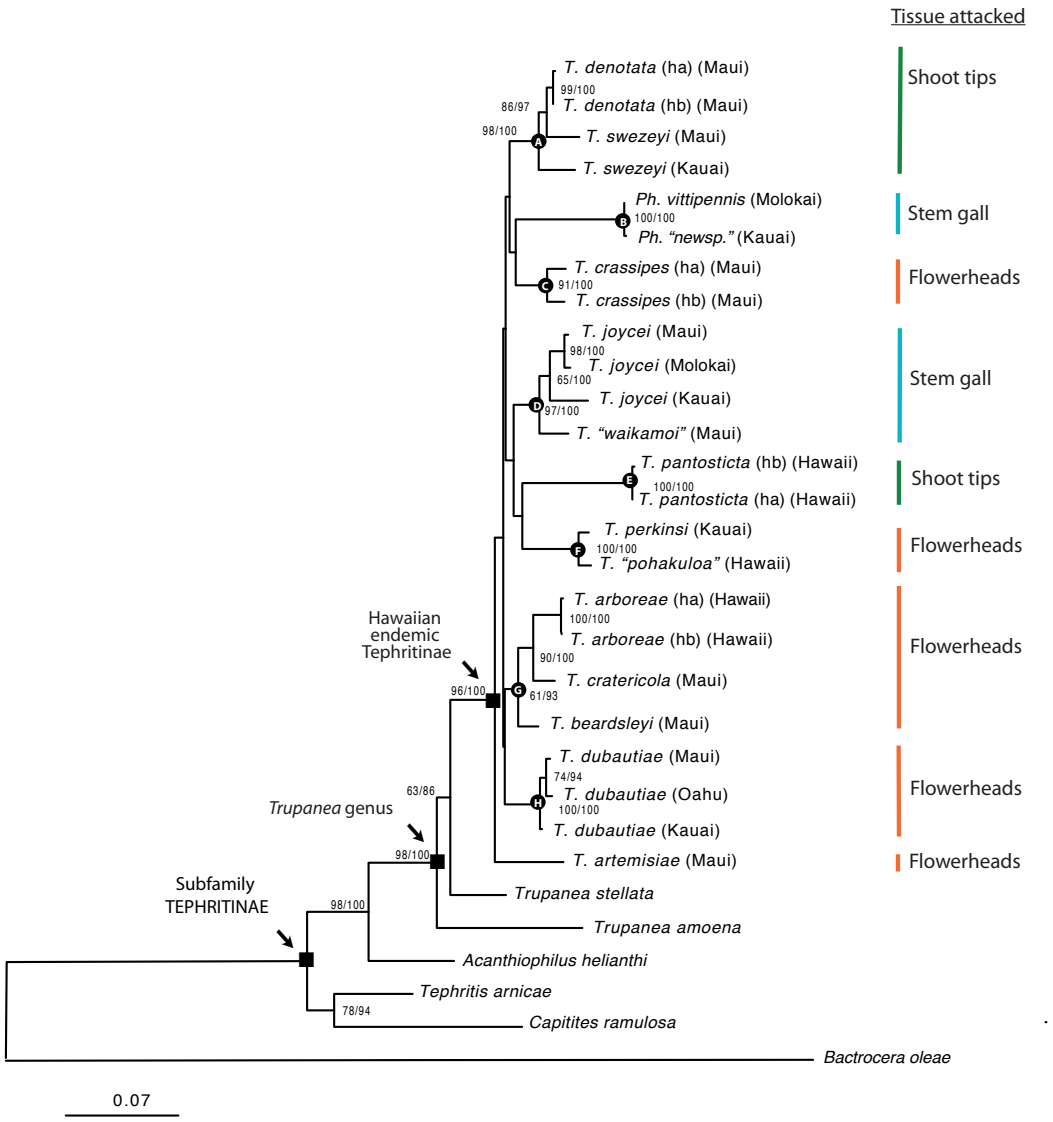

Fig. 2 Phylogenetic reconstruction of endemic tephritid samples from the Hawaiian Archipelago and several outgroups (Tab. 2) obtained by Maximum Likelihood (ML) and Bayesian inference (BI) analyses of the concatenated mitochondrial data set (16S rRNA-COI-tRNA-COII). The topology shows the best ML tree with Bootstrap support (Bs; left) and Posterior Probabilities (PP; right) from BI.

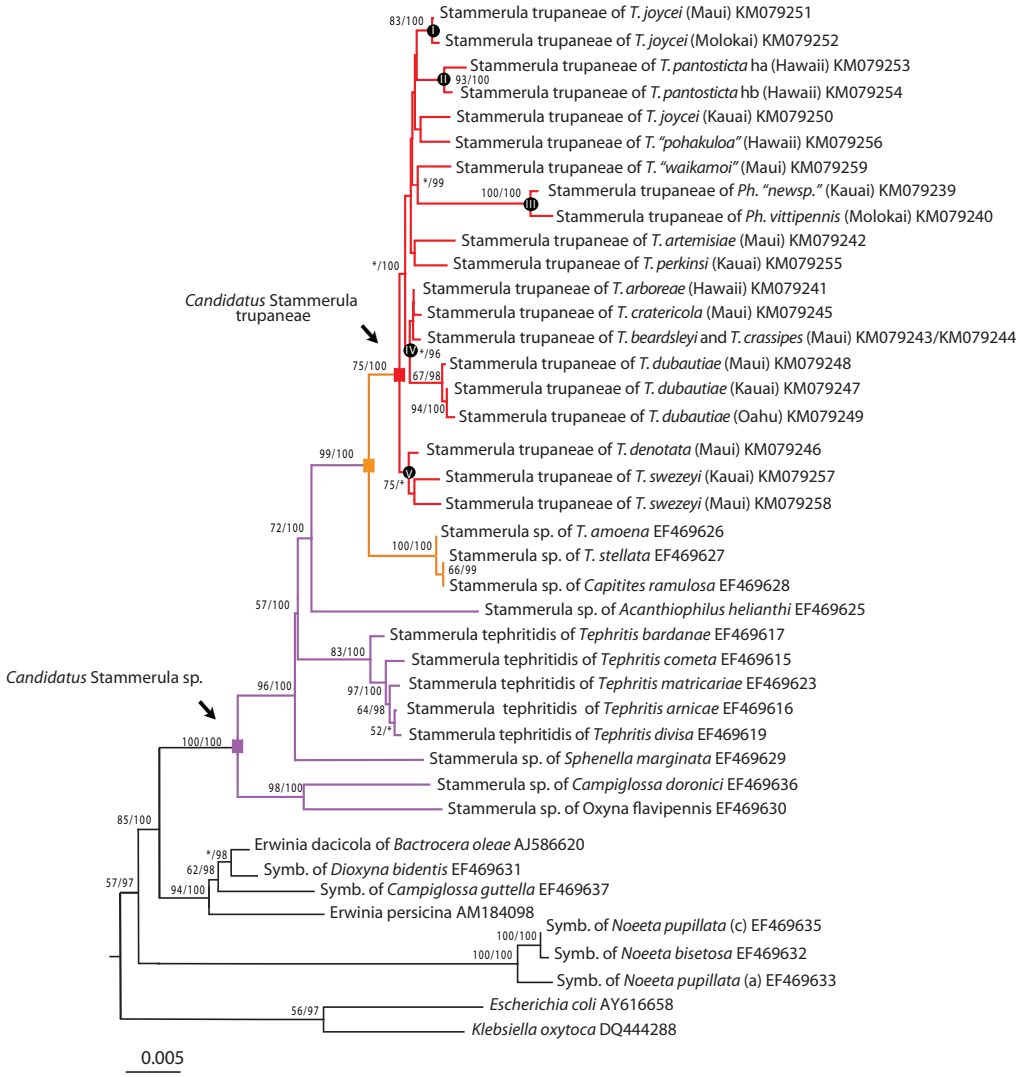

Fig 3 Phylogenetic reconstruction of the symbiont of endemic Hawaiian tephritids and several outgropus (Tab. 2) inferred with Maximum Likelihood (ML) and Bayesian inference (BI) analyses of the 16S rRNA bacterial gene. The topology shows the best ML tree with Bootstrap support (Bs; left) and Posterior Probabilities (PP; right) from BI. Asterisks denote Bs lower than 50% and PP lower than 95%.

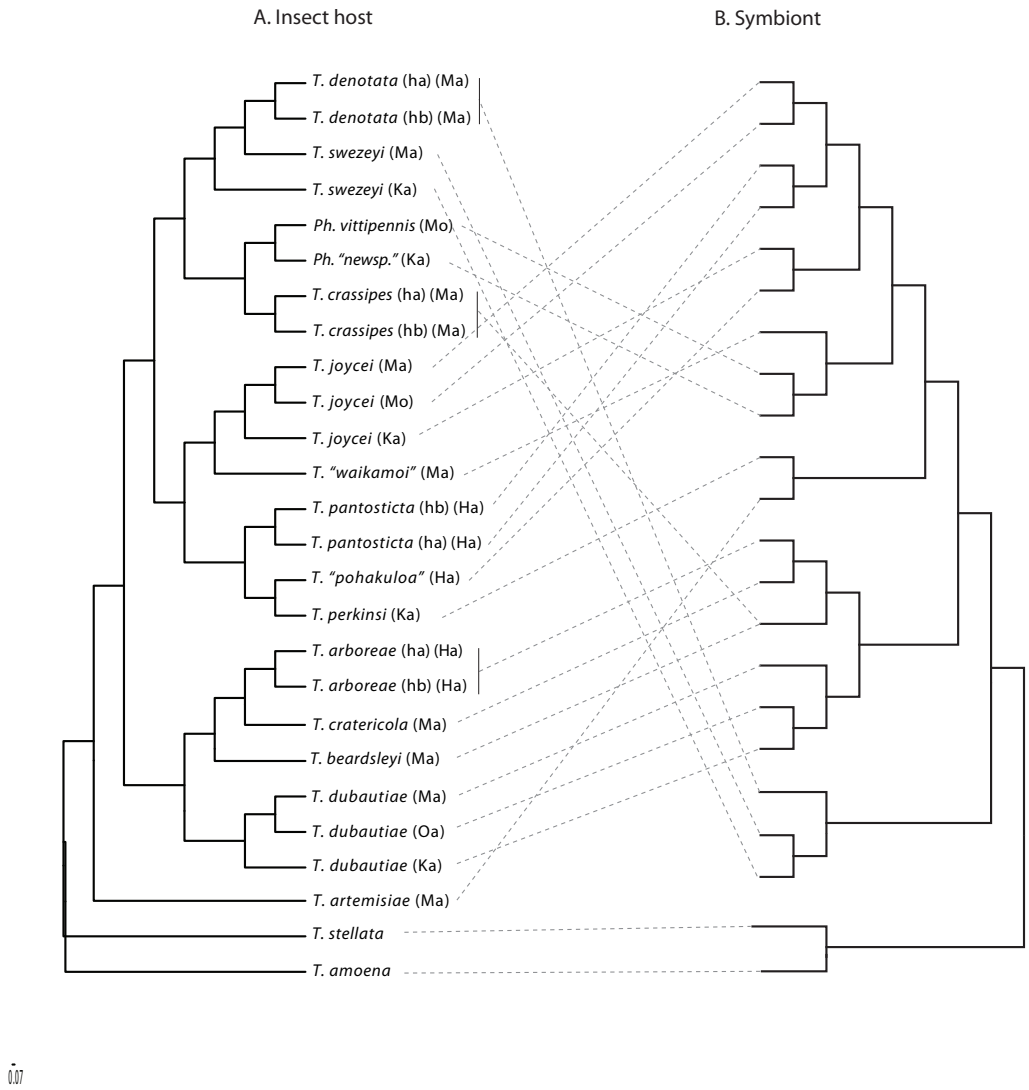

Fig. 4 Tanglegram linking the inferred phylogeny of endemic Hawaiian tephritid hosts to their symbiotic bacteria. Connecting lines illustrate host-symbiont associations. The topology shows the best ML trees based on the concatenated mitochondrial data set (16S rRNA-COI-tRNA-COII) for the Tephritinae host and the 16S rDNA bacterial gene of their symbionts.

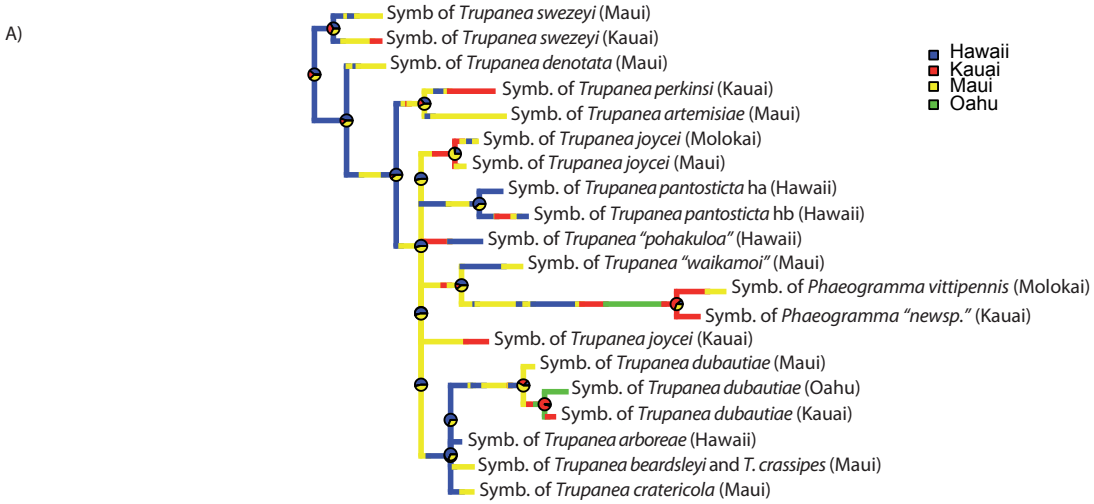

Fig. 5 Character reconstruction of host traits, island location (A), host lineage (B), and host tissue attacked (C) on the symbiont phylogeny using stochastic character mapping (Huelsenbeck et al., 2003), as implemented in the R program *phytools* v. 0.4-05 (Revell, 2012).

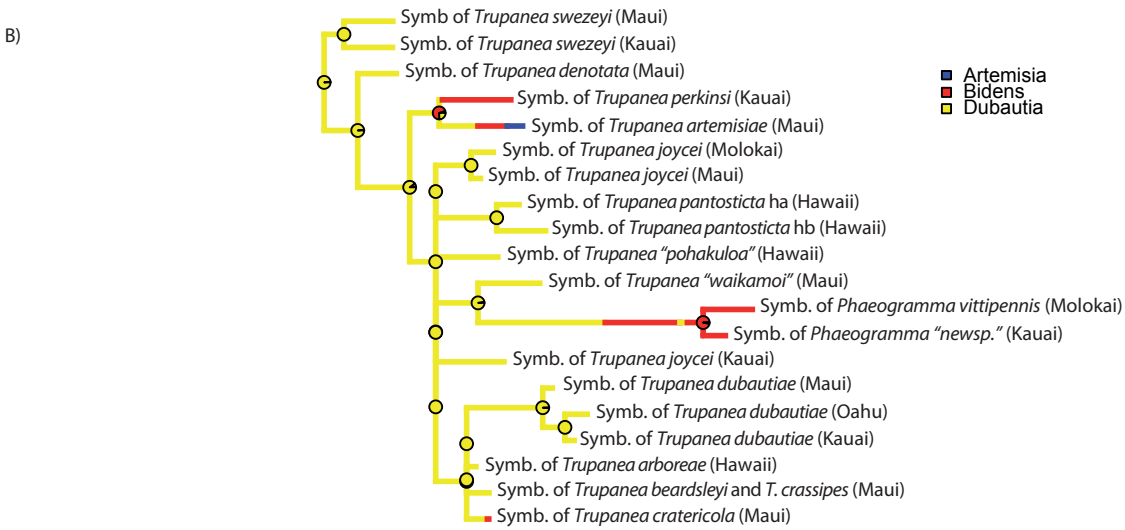

Fig. 5 Character reconstruction of host traits, island location (A), host lineage (B), and host tissue attacked (C) on the symbiont phylogeny using stochastic character mapping (Huelsenbeck et al., 2003), as implemented in the R program *phytools* v. 0.4-05 (Revell, 2012).

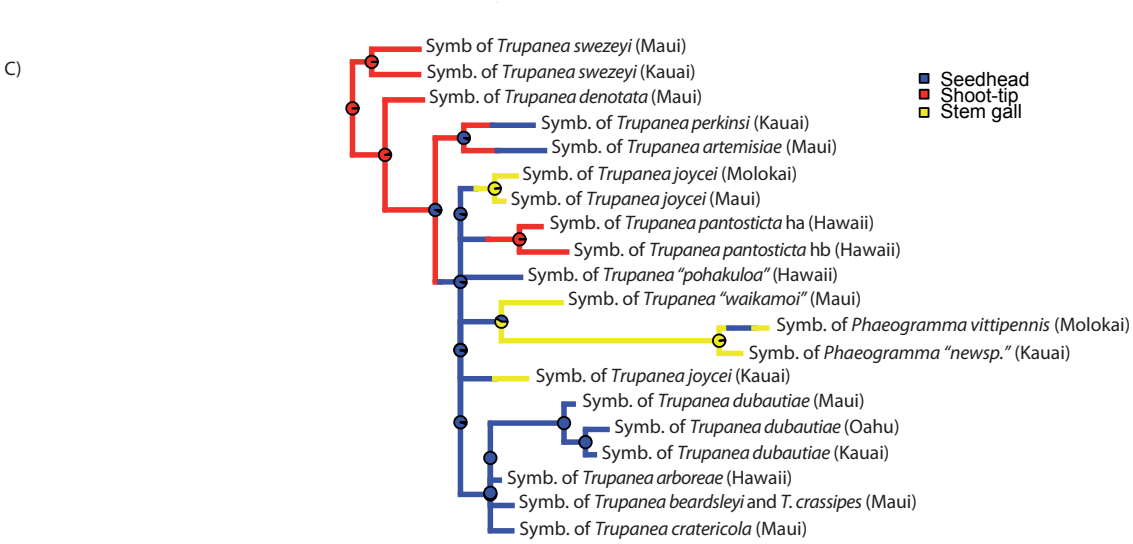

Fig. 5 Character reconstruction of host traits, island location (A), host lineage (B), and host tissue attacked (C) on the symbiont phylogeny using stochastic character mapping (Huelsenbeck et al., 2003), as implemented in the R program *phytools* v. 0.4-05 (Revell, 2012).

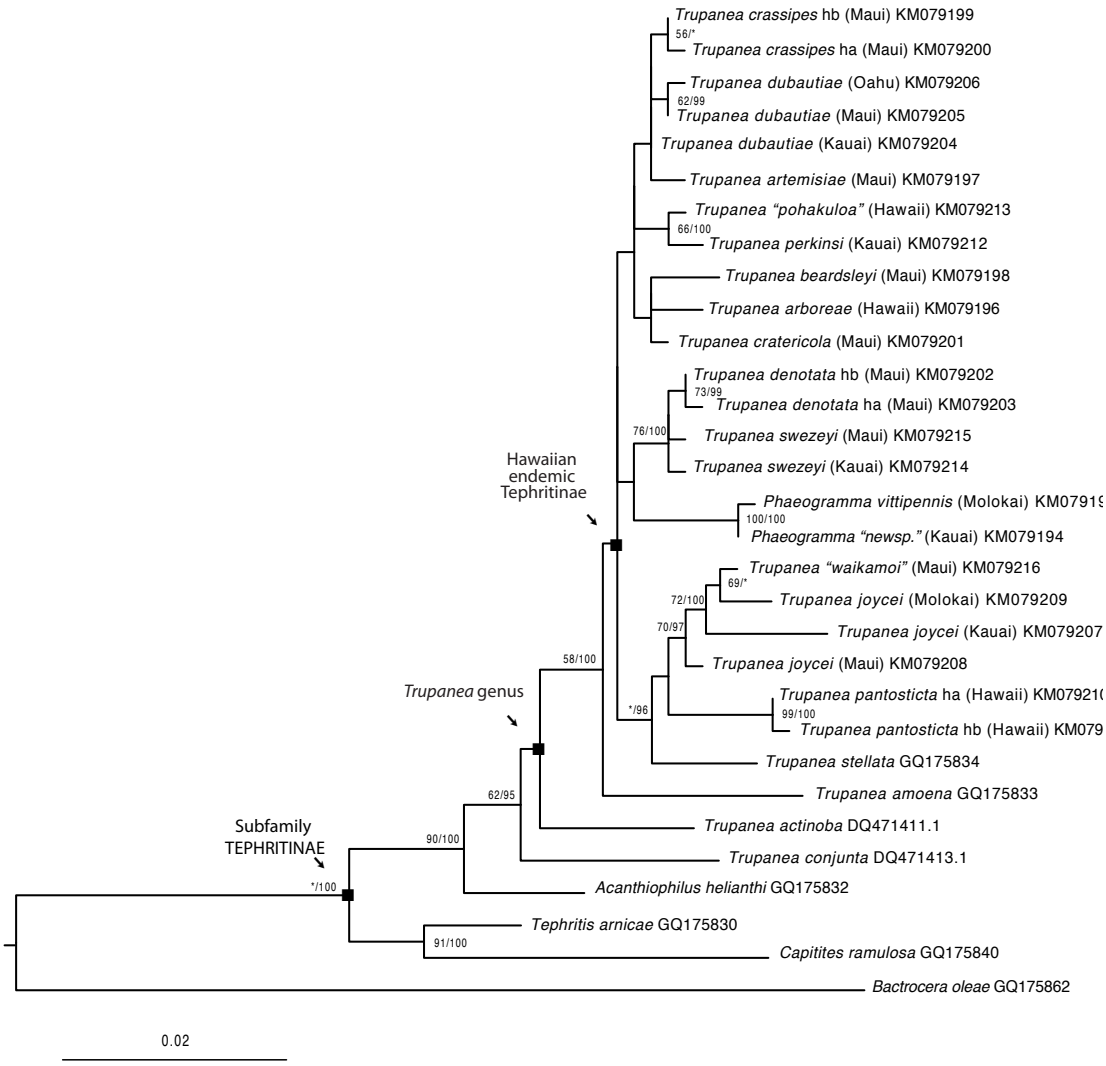

Fig S1 Inferred phylogeny of endemic tephritid samples from the Hawaiian Islands and several outgroups (Tab. 2) obtained by Maximum Likelihood (ML) and Bayesian inference (BI) analyses of the mitochondrial 16S rRNA gene. The topology shows the best ML tree with Bootstrap support (Bs; left) and Posterior Probabilities (PP; right) from BI. Asterisks denote Bs lower than 50% and PP lower than 95%.

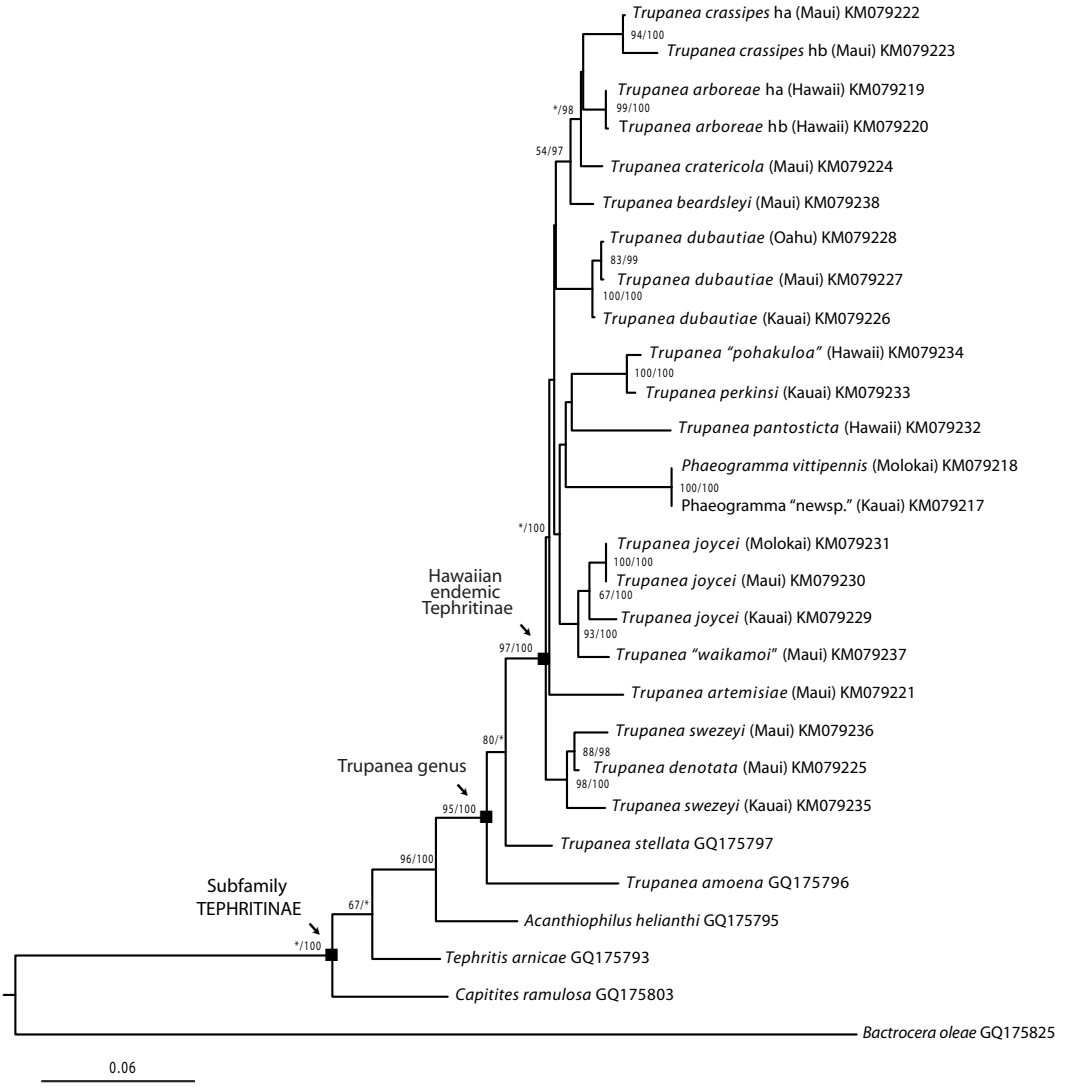

Fig S2 Inferred phylogeny of endemic tephritid samples from the Hawaiian Islands and several outgroups (Tab. 2) obtained by Maximum Likelihood (ML) and Bayesian inference (BI) analyses of the mitochondrial COI-tRNA-COII gene. The topology shows the best ML tree with Bootstrap support (Bs; left) and Posterior Probabilities (PP; right) from BI. Asterisks denote Bs lower than 50% and PP lower than 95%.

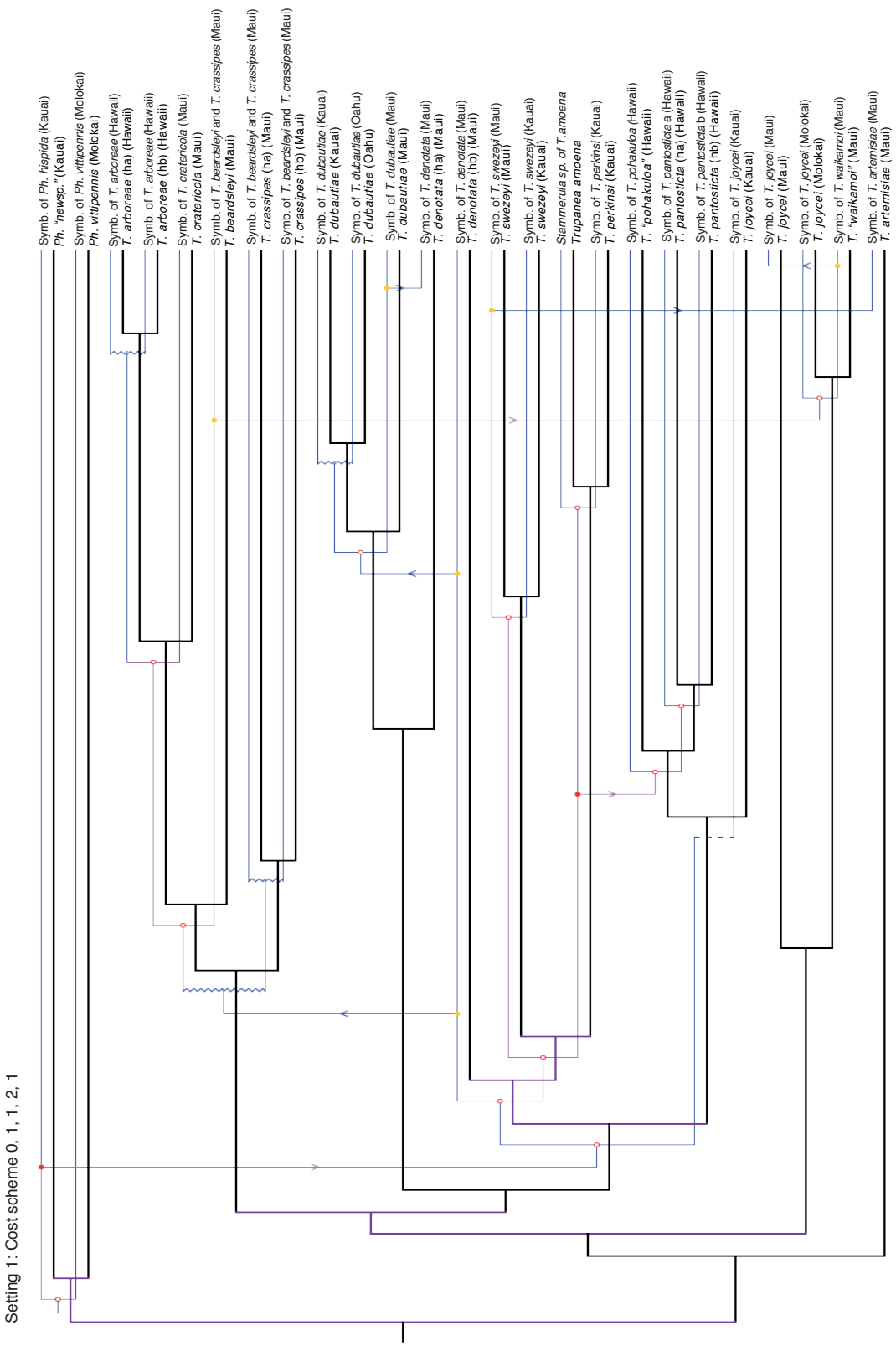

Fig. S3 Result of reconciliation analysis implemented in Jane (Conow et al. 2010) for tephritid host insects and their symbiotic bacteria phylogenies for the event cost setting 1.

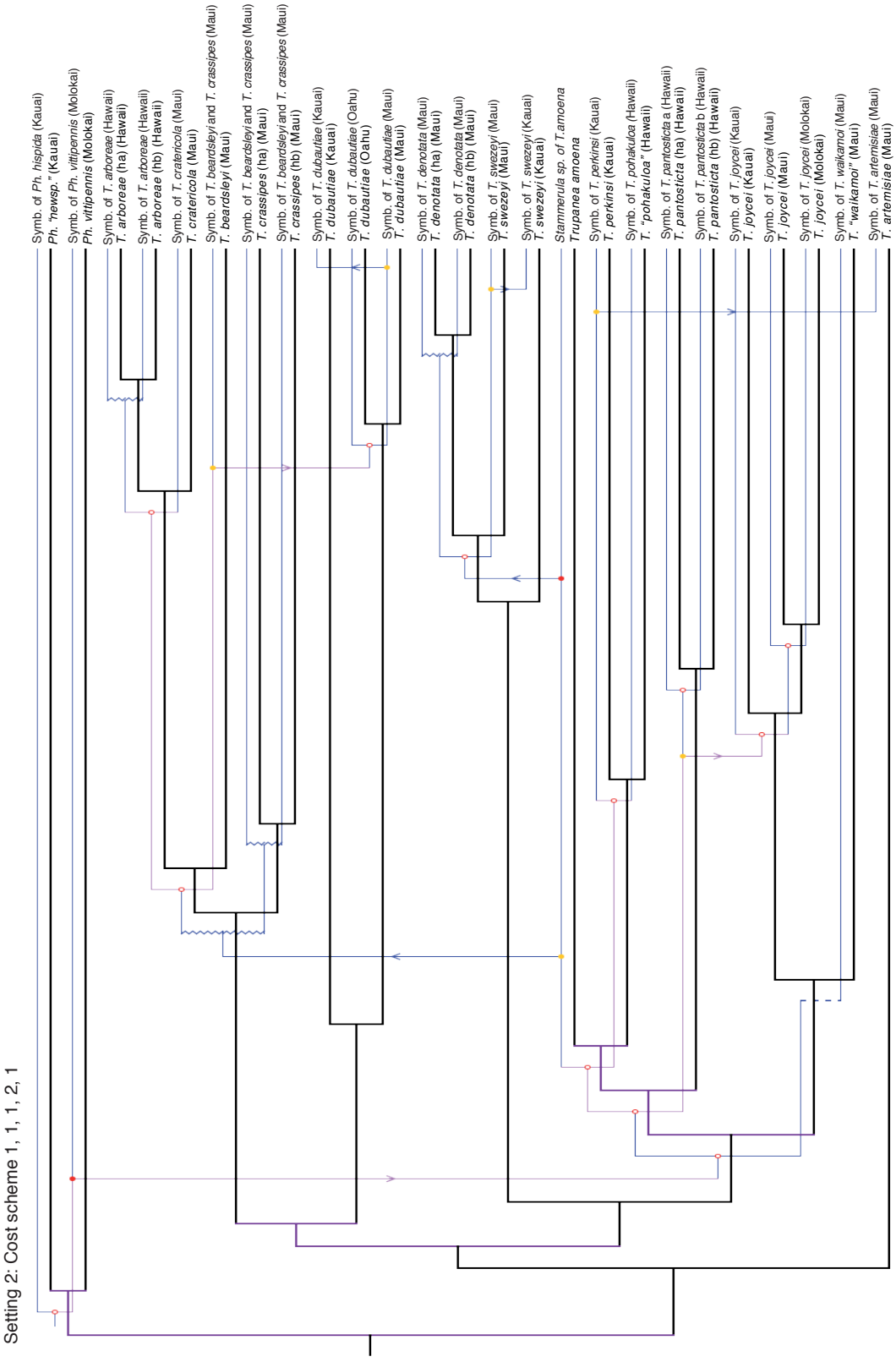

Fig S4 Result of reconciliation analysis implemented in Jane (Conow et al. 2010) for tephritid host insects and their symbiotic bacteria phylogenies for the event cost setting 2.

References

Baker, R.H. and DeSalle R., 1997. Multiple Sources of Character Information and the Phylogeny of Hawaiian Drosophilids.Syst. Biol.46, 654-673.

Baldwin, B.G. and Robichaux R.H., 1995. Historical biogeography and ecology of the Hawaiian silversword alliance (Asteraceae): New molecular phylogenetic perspectives,in: Wagner, W.L., Funk, V.A. (Eds.), Hawaiian Biogeography: Evolution on a hot spot Archipelago.Smithsonian Institution Press, Washington DC, pp. 259-287.

Baldwin, B.G., 1997. Adaptive radiation of the Hawaiian silversword Alliance: congruence and conflict of phylogenetic evidence from molecular and non-molecular investigations, in: Givnish, T.J. and Sytsma, K.J. (Eds),Molecular Evolution and Adaptive Radiation. Cambridge University Press, pp. 103-128.

Baldwin, B.G., 1998. Age and Rate of Diversification of the Hawaiian Silversword Alliance (Compositae).Proc. Natl. Acad. Sci. USA. 95, 9402-9406.

Baldwin, B.G., 2006. Contrasting patterns and process of evolutionary change in the tarweed-silversword lineage: revisiting Clausen, Keck, and Hiesey's findings. Ann. Missouri Bot. Gard. 93, 64-93.

Barrier, M., Baldwin B.G., Robichaux R.H.,Purugganan M.D., 1999. Interspecific Hybrid Ancestry of a Plant Adaptive Radiation: Allopolyploidy of the Hawaiian Silversword Alliance (Asteraceae) Inferred from Floral Homeotic Gene Duplications.Mol. Bio. Evol. 16, 1105-1113.

Brown, J.M., Todd-Thompson M., McCord A., O'Brien A., O'Fallon B., 2006. phylogeny, host association, and wing pattern variation in the endemic Hawaiian Fruit Flies (Diptera, Tephritidae).Instr. Biodiv. 7,1-16.

Caraway, V., Carr G.D., Morden, C.W., 2001. Assessment of hybridization and introgression in lava-colonizing Hawaiian *Dubautia* (Asteraceae: Madiinae) using RAPD markers. Am. J. Bot. 88, 1688-1694.

Carr, G.D., 1985. Monograph of the Hawaiian Madiinae (Asteraceae): *Argyroxiphium, Dubautia* and *Wilkesia*. Allertonia 4, 1–123.

Carr, G.D., 1998. Another New Species of Dubautia (Asteraceae-Madiinae) from Kaua'i, Hawaiian Islands. Novon. 8, 8-11.

Carson, H.L., Clague D.A., 1995. Geology and Biogeography of the Hawaiian Islands.In: Wagner, W.L., Funk, V.A. (Eds.), Evolution Evolution on A Hotspot Archipelago. Smithsonian Institution Press, Washington, DC, pp. 160– 194.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis.Mol. Bio. Evol. 17, 540-552.

Clague, D.A., Dalrymple, G.B., 1987. The Hawaiian-emperor volcanic chain. In: Decker, R.W., Wright, T.L., Stauffer, P.H. (Eds.), Volcanism in Hawaii. U.S. Geological Survey Professional Paper 1350. U.S. Government Printing Office, Washington, DC, pp. 1–54.

Clague D.A., 1996. The growth and subsidence of the Hawaiian-emperor volcanic chain.In: Keast, A., Miller, S. E. (Eds.), The Origin and Evolution of Pacific Island Biotas, New Guinea to Eastern Polynesia: Pattern and Processes.SPB Academic Press, Amsterdam, pp. 35-50.

Conow, C., Fielder D., Ovadia Y., Libeskind-Hadas R., 2010. Jane: A new tool for the cophylogeny reconstruction problem.Algorithm. Mol. Biol. 5, 16-25.

Drummond, A.J., Rambaut A., 2007. BEAST: Bayesian Evolutionary Analysis by Sampling Trees.BMC Evol. Biol. 7, 214.

Fahrenholz, H., 1913. Ektoparasiten Und Abstammungslehre.Zoologischer Anzeiger 41, 371-374.

Fleischer, R.C., McIntosh C.E., Tarr C.L., 1998. Evolution on a Volcanic Conveyor Belt: Using Phylogeographic Reconstructions and K–Ar‐based Ages of the Hawaiian Islands to Estimate Molecular Evolutionary Rates.Mol. Ecol.7, 533-545.

Fleischer, R.C., McIntosh C.E., 2001. Molecular Systematics and Biogeography of the Hawaiian Avifauna.Stud. Avian Biol. 22, 51-60.

Foote, R.H., Blanc F.L., Norrbom A.L., 1993. Handbook of the Fruit Flies (Diptera: Tephritidae) of America north of Mexico. Comstock Publishing Associates. Ithaca and London.

Gillespie, R. 2004., Community assembly through adaptive radiation in Hawaiian spiders.Science. 303, 356-359.

Girolami, V., 1973. Reperti morfo-istologici sulle batteriosimbiosi del Dacus oleae Gmelin e di altri ditteri tripetidi, in natura e negli allevamenti su substrati artificiali. Redia 54, 269–294.

Girolami, V. 1983. Fruit fly symbiosis and adult survival: general aspects. In: Cavalloro, R., (Ed.), Fruit Flies of Economic Importance, Proc. CEC/IOBC Int Symp, Athens November 1982. A.A. Balkema, Rotterdam, pp. 74–76.

Goodman, K., O'Grady P., Wicker-Thomas C., 2013. Molecular phylogeny and biogeography of the Hawaiian Craneflies *Dicranomyia* (Diptera: Limoniidae). PLoS One. 8, e73019.

Guindon, S., Gascuel O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by Maximum Likelihood.Systematic Biology. 52, 696-704.

Hardy, D.E., Delfinado, M.D., 1980.Diptera: Cyclorrhapha III, serie Schizophora. Section Acalyptrata, exclusively of family Drosophilidae. In: Zimmermann, E.C. (Eds.). Insects of Hawaii. Volume 13. University of Hawai'i Press, Honolulu. pp. 451.

Huelsenbeck,, J.P., Nielsen, R, Bollback, J.P. 2003. Stochastic mapping of morphological characters. Syst Biol 52: 131-158.

Jordan, S., Simon C., Polhemus D., 2003. Molecular systematics and adaptive radiation of Hawaii's endemic Damselfly genus *Megalagrion* (Odonata: Coenagrionidae).Syst. Biol.52, 89-109.

Koch, H., Abrol D.P., Li J., Schmid‐Hempel P., 2013. Diversity and evolutionary patterns of bacterial gut associates of Corbiculate Bees. Mol. Ecol.22, 2028-2044.

Korneyev, V.A., 1999. Phylogeny of the subfamily Tephritinae: relationships of the tribes and subtribes.In: Aluja, M., Norrbom, A.L. (Eds.), Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior. CRC Press, Boca Raton, pp. 549–580.

Law, R., Lewis D.H., 1983. Biotic environments and the maintenance of sex–some evidence from mutualistic symbioses.Biol. J. Linn. Soc.20, 249-276.

Law, R., 1985. Evolution in a mutualistic environment. In: Bouche, D.H., (Eds.), The Biology of Mutualism: Ecology and Evolution, Oxford University Press, London, pp. 145-170.

Legendre, P., M.J. Anderson.. 1998. Program Distpcoa. Département de Sciences Biologiques, Université De Montréal.

Legendre, P., 2001. Test of host-parasite coevolution: program PARAFIT user's guide.Département de Sciences Biologiques, Université de Montréal.

Legendre, P., Desdevises Y., Bazin E., 2002. A statistical test for host-parasite coevolution.Syst. Biol. 51, 217-234.

Light, J.E., Hafner M.S., 2007. Cophylogeny and disparate rates of evolution in sympatric lineages of chewing lice on pocket gophers.Molecular Phylogenetics and Evolution. 45, 997-1013.

Mazzon, L., Piscedda A., Simonato M., Martinez-Sanudo I., Squartini A., Girolami V., 2008. Presence of specific symbiotic bacteria in flies of the subfamily Tephritinae (Diptera Tephritidae) and their phylogenetic relationships: proposal of '*Candidatus* Stammerula Tephritidis'.Int. J. Syst. Evol. Microbiol.58, 1277-1287.

Mazzon, L., Martinez-Sanudo I., Simonato M., Squartini A., Savio C., Girolami V., 2010. Phylogenetic relationships between flies of the Tephritinae subfamily (Diptera, Tephritidae) and their symbiotic bacteria.Mol. Phylogenet. Evol.56, 312-326.

Mazzon, L., Martinez-Sañudo, I., Savio, C., Simonato, M., Squartini, A., 2011 - *Stammerula* and other symbiotic bacteria within the fruit flies inhabiting Asteraceae flowerheads In: *Manipulative Tenants Bacteria Associated With Arthropods*. Edited by Zchori-Fein, E. and Bourtzis, K. CRC Press, Taylor and Francis Group, pp. 1-136.

Meier-Kolthoff, J.P., Auch A.F., Huson D.H., Goker M., 2007. COPYCAT: cophylogenetic analysis tool.Bioinformatics (Oxford, England) 23, 898-900.

Mendlová, M., Desdevises Y., Civáňová K., Pariselle A., Šimková A., 2012. Monogeneans of west african cichlid fish: evolution and cophylogenetic interactions.PloS One. 7, e37268.

Nitta, J.H., O'Grady P.M., 2008. Mitochondrial phylogeny of the endemic Hawaiian craneflies (Diptera, Limoniidae,*Dicranomyia*): implications for biogeography and species formation. Mol. Phylogenet. Evol.46, 1182-1190.

Norrbom, A., 1999. Status of knowledge. In: Thompson, F.C. (Ed.), Fruit fly expert

identification system and systematic information database. Backhuys Publishers, Leiden, Myia, pp. 9–48.

Patwary, M.U., Kenchington, E.L., Bird, C.J., Zouros, E., 1994. The use of random amplified polymorphic DNA markers in genetic studies of the sea scallop Placopecten magellanicus (Gmelin, 1791). J. Shellfish Res. 13, 547–553.

Petri, L., 1909. Ricerche Sopra i Batteri Intestinali della Mosca Olearia. Memorie della Regia Stazione di Patologia Vegetale di Roma, Roma.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of AIC and bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol.Evol. 25, 1253-1256.

Price, J.P., Elliott-Fisk D., 2004. Topographic history of the Maui Nui Complex, Hawai'i, and its implications for biogeography. Pac.Sci. 58, 27-45.

Rambaut, A., Drummond A., 2008. FigTree: tree figure drawing tool, Version 1.2. 2.Institute of Evolutionary Biology, University of Edinburgh.

Randazzo, C. L., Scifò, G. O., Tomaselli, F., Caggia, C., 2009. Polyphasic characterization of bacterial community in fresh cut salads. Int. J. Food Microbiol. 128, 484-490.

Revell, L. J. (2012) phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3: 217-223.

Robichaux, R.H., Carr G.D., Liebman M., Pearcy R.W., 1990. Adaptive radiation of the Hawaiian silversword alliance (Compositae-Madiinae): ecological, morphological, and physiological diversity. Ann.Missouri Bot. Gard. 77, 64-72.

Roderick, G.K., Gillespie R.G., 1998. Speciation and phylogeography of Hawaiian terrestrial arthropods.Mol. Ecol. 7, 519-531.

Ronquist, F., Huelsenbeck J.P., 2003. MrBayes 3: bayesian phylogenetic inference under mixed models.Bioinform. (Oxford, England). 19, 1572-1574.

Roy, T., Chang, T.H., Lan, T., Lindqvist, C., 2013.Phylogeny and Biogeography of new world Stachydeae (Lamiaceae) with emphasis on the origin and diversification of Hawaiian and south american taxa.Mol. Phylogenet. Evol. 69,

218-238.

Rubinoff, D., Schmitz P., 2010. Multiple aquatic invasions by an endemic, terrestrial Hawaiian moth radiation.Proc. Natl. Acad. Sci. USA. 107, 5903-5906.

Schluter, D., 2000.The Ecology of Adaptive Radiation. Oxford University Press. Oxford, U.K.

Simon, C., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction Primers. Ann. Entomol. Soc. Am. 87, 651-704.

Simon, C.,Buckley, T.R.,Frati, F.,Stewart, J.B., Beckenbach, A.T., 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA.Annu. Rev. Ecol. Syst.. 37, 545-579.

Stammer, H.J., 1929. Die bakteriensymbiose der trypetiden (Diptera)." Zoomorphology. 15, 481-523.

Swofford, D.L., 2003. PAUP. Phylogenetic analysis using parsimony (and other methods) version 4.0. Sinauer Associates. Sunderland, Massachussetts.

Tamura, K., Dudley J., Nei M., Kumar S., 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software Version 4.0.Mol. Biol. Evol. 24, 1596-1599.

Wagner, W.L. Herbst D.R., Sohmer S.H., 1999.Manual of the flowering plants of Hawai'i. University of Hawaii and Bishop Museum Press, Honolulu.

Weisburg, W.G., Barns S.M., Pelletier D.A., Lane D.J., 1991. 16S ribosomal DNA amplification for phylogenetic study.J. Bacteriol. 173, 697-703.

Wiens, J.J., 1998. Combining Data Sets with Different Phylogenetic Histories.Syst. Biol. 47, 568-581.

Ziegler, A. C., 2002.Hawaiian Natural History, Ecology, and Evolution. University of Hawaii Press. Honolulu.

Zimmerman, Elwood Curtin. 1981.Insects of Hawaii. University of Hawaii Press. Honolulu.

Chapter III

Genetic variability of '*Candidatus* Erwinia dacicola' in *Bactrocera oleae* populations across their geographical distribution

---

Manuscript in preparation as:

V. Girolami, E. Viale, C. Savio, I. Martinez-Sañudo, M. Simonato, L. Mazzon.

Department of Agronomy, Food, Natural Resources, Animals and the Environment (DAFNAE), University of Padova - Agripolis, Viale dell'Università, 16, 35020 Legnaro (Padova), Italy

I contributed to the experimental work, data analyses and paper writing

Introduction

The olive fly *Bactrocera oleae* (Diptera, Tephritidae) is the major insect pest of olive crops worldwide. The larvae are monophagus feeding exclusively on wild and cultivated olive fruits of the genus *Olea* (Tzanakakis, 2006) (including *O. europea*, *O. verrucosa* and *O. chrystophylla*). In areas of the world where the olive fruit fly is established, it has been responsible for losses of up to 80% of oil value and 100% of some table cultivars (Neuenschwander and Michelakis, 1981; Tzanakakis, 2006). It has been estimated to damage 5% of total olive production, resulting in economic losses of approximately $800 million per year (Montiel Bueno and Jones, 2002).

The olive fly's distribution is primarily limited to regions where cultivated and wild trees are found (Daane and Johnson, 2010). Today, the olive fruit fly is reported throughout the Mediterranean basin, the Canary Islands, South and Central Africa, Middle East, California and Central America where it has been recently introduced in the 1998 (Rice, 1999; Rice and Navajas, 2003; Augustinos et al., 2002; Copeland et al., 2004). Despite its abundance and notoriety in the Mediterranean basin, the olive fruit fly is most likely associated with wild varieties of olives in the region of Central Africa from which domesticated cultivars were derived (Zohary, 1993; Nardi et al., 2005); while the North American olive fly seems to originate from the Eastern Mediterranean, highlighting the importance of long distance human mediated dispersal of this insect (Zygouridis et al., 2009; Nardi et al., 2010; Dogac et al., 2013).

Previous studies support the subdivision of the olive fly populations into three main groups: Pakistan, Africa and Mediterranean plus America, corresponding to the major geographical areas of its geographical dispersal (Nardi et al., 2005, 2010). A clear phylogeographic structure has been also reported in the Mediterranean populations, further differentiated into Eastern (Cyprus), Central (Italy and Greek) and Western (Iberian peninsula) groups (Augustinos et al., 2005; Van Asch et al., 2012). In addition, a gradual decrease of heterozygosity from the Eastern to the Western part of the Mediterranean has been demonstrated, indicating a Westward expansion of the species, which may reflect the colonization process of the olive fly in the European part of the Mediterranean basin (Augustinos et al., 2005).

The bacterial microbiota of the olive fruit fly, described for the first time by Petri (1909), has been studied for over a century (Petri, 1909; Girolami, 1973; Manousis and Ellar, 1988; Capuzzo et al., 2005, Sacchetti et al., 2008). Interest in the reciprocal influences between hosts and symbionts has increased in the last ten years because of the need to control devastating diseases, identify or develop biocontrol agents against invasive pests and improve agricultural production (Poulin and Morand, 2004). In adult olive fly, extracellular unculturable bacteria are located in the oesophageal bulb, a specialized organ in the fly head (Petri, 1909; Girolami, 1973), in which symbionts multiply rapidly, forming masses that reach the midgut. Although recent studies have identified several species of bacteria in the digestive tract of wild olive flies (Kounatidis et al., 2009), the symbiont "*Candidatus* Erwinia dacicola" is considered the most common and widespread bacterium within the olive fly population (Capuzzo et al., 2005; Estes et al., 2009). The symbiont was firstly described in Italy (Capuzzo et al., 2005) and subsequently recovered in olive flies in Spain (Silva et al., 2008), South Western USA (Estes et al., 2009) and Greece (Kounatidis et al., 2009).
Kounatidis et al. (2009) found the culturable bacterium *Acetobacter tropicalis* as predominant. The literature reports that symbiont losses are a consequence of rearing larvae on artificial media (Girolami and Cavalloro, 1972; Capuzzo et al., 2005), as well as their substitution by acidophilus bacteria (Hagen, 1966).

In a recent work (Savio et al 2012), the genetic variability of *Ca*. E. dacicola in Italian olive fly populations has been studied and two bacterial lineages were found. These symbiont haplotypes resulted randomly distributed across the Italian peninsula while they were exclusively represented in the two main island populations.

The present work extends Savio et al. (2012) study to a wider geographical sampling of the olive fly, including most regions of the Mediterranean area plus some isolated samples from Pakistan, South Africa, and California.

The aims of this study were: i) to investigate the genetic variability and haplotypes distribution of *Ca*. E. dacicola, ii) to study the correlation between the genetic variability of the olive fly symbiont and the mitochondrial haplotypes of its host fly and iii) to assess the population structure of *B. oleae* over a large areas of the insect geographical distribution.

Material and methods

Olive fly sampling and dissection

About one thousand of olive flies samples were collected from egg-infested olives in 50 major olive growing locations, selected to be representative of the entire distribution area. Populations from 17 countries around the Mediterranean basin, one population from South Africa, one from Pakistan and one from California (Tab. 1) were sampled between 2002 and 2012. In particular, the dataset includes 20 representative Italian olive fly populations from Savio et al. (2012). Infested olives were stored in transparent plastic boxes (20x15x8 cm) topped with a net until larvae emerged and developed to adulthood. Emerged insects were then transferred into net rearing cages (10x10x10 cm) and fed with a sugar diet (50 % w/v glucose solution) and water *ad libitum* according to Savio et al. (2012) protocols.

In newly emerged adults, the presence of endosymbionts is sometimes too low to be detected (Girolami, 1973; Estes et al., 2009) and so only 2–3 day-old flies were processed. This ensured that a sufficient number of bacteria was present in the esophageal bulbs.

Flies were dissected to extract the esophageal bulb hosting symbionts, following the procedure described in Capuzzo et al. (2005) but processing the whole fly head, as it give the same results in term of extracted DNA quality and sequencing.

A protocol originally developed for actinomycetes was used (Palmano et al., 2000) to extract both the insect and the bacterial DNA content of the oesophageal bulbs, when possible, with an average of 8 individuals per population.

In order to assess the quality of the DNA, aliquots from extracted samples were separated in 1% agarose gel and viewed under UV after staining with SYBR Safe DNA gel stain (Invitrogen).

Symbiont analyses: bacterial DNA amplification and sequencing

The bacterial DNA was analyzed by amplifying a fragment of 16S rRNA gene, using the following two pair of primers: fD1 (forward: 5'-AGAGTTTGATCCTGGCTCAG-3'), rP1 (revers: 5'-

ACGGTTACCTTGTTACGACTT-3') (Weisburg et al., 1991) and 63F (forward: 5'-CAGGCCTAACACATGCAAGTC-3') and 1389R (revers: 5'-ACGGGCGGTGTGTACA AG-3') (Osborn et al., 2000). Amplifications were performed in 20 µl reaction volumes according to Savio et al. (2012). The standard thermal profile for the amplification of the 16S rRNA genes included an initial denaturation at 95°C for 2 min, 27 cycles at 95°C 30 s, 56°C for 30 s, 72°C for 90 s and a final extension at 72°C for 5 min. In addition, PCR products obtained from 7 oesophageal bulbs were cloned into JM109 competent cells using the P-GEM-T Easy vector (Promega), following the manufacturer's recommendations. Transformations were verified using PCR assays with the M13-T7 universal primer pair. PCR products were checked by 1 % agarose gel stained with SYBR® Safe (Invitrogen) and purified with a mix of Exonuclease and Antarctic Phosphatase (New England Biolabs) and sequenced at the BMR Genomics service at Padova, Italy, using two additional primers fl2 (5'- GGAACTGCATTCGAAACTG - 3') (Capuzzo et al., 2005) and fL4 (5'-CGGGTGAGTAATGTCTG-3') (Mazzon et al., 2008).

Symbiont data analyses

Bacterial sequences were inspected and aligned using MEGA 5.2.2. (Tamura et al., 2007) and then analyzed with a BLAST search (www.ncbi.nlm.nih.gov) against specified databases to check the specificity of the sequences obtained.

Spatial analysis of molecular variance (SAMOVA) was performed on the *Ca.* E. dacicola sequences obtained using the program SAMOVA 12.02 (Dupanloup et al., 2002) to identify groups of populations that are phylogeographically homogeneous and maximally differentiated from each other, taking into account the geographic distances. This analysis, based on simulated annealing procedure, permits to identify the maximally differentiated groups that correspond to predefined genetic barriers by maximizing the proportion of total genetic variance due to differences between groups of populations (F_{CT}) (Crawford, 2007). Analyses were performed for a user-defined k set between 2 and 8. F_{SC} and F_{CT} patterns were examined of for each k to determine the appropriate number of population groups among the 50 olive fly populations, based on 100 simulated annealing process.

Insect host analyses: amplification and sequencing

A region of the mitochondrial DNA of the olive fly, corresponding to the NADH dehydrogenase subunit 1, the leucine tRNA and the 16S mitochondrial region, was amplified using the universal pair of primers: N1-J12261m (forward: 5'-TACTTCGTAAGAAATTGTTTGAGC-3') and LRN13398 (revers: 5'-CGCCTGTTTAACAAAAACAT-3') (Simon et al., 1994, 2006). PCR reactions were performed with the same volumes used for the symbiont DNA amplification.

The cycling program included an initial denaturation at 96 °C for 5 min, followed by 35 cycles at 96 °C for 50s, 56 °C for 50s and 72 °C for 2 min, with a final extension step at 72 °C for 5 min. PCR amplified products were purified and sequenced as performed for the symbiont using the same pairs of primers used for PCR amplification.

Insect data analyses

Insect host sequences obtained were aligned in MEGA 5.2.2. (Tamura et al., 2007) and then analyzed with BLAST (www.ncbi.nlm.nih.gov) as performed for the symbiont sequence analysis.

Haplotype distributions and frequencies within the populations (N_{HT}), gene diversity (h), nucleotide diversity (π) and mean number of pairwise differences (k) were assessed using ARLEQUIN v 3.5 software package (Excoffier et al., 2007).

Analyses of molecular variance (AMOVA) (Excoffier et al., 1992), implemented in ARLEQUIN v 3.5 (Excoffier et al., 2007), were then performed on the 50 *B. oleae* populations to test the genetic relationships between the different groups defined by SAMOVA analysis on *Ca.* E. dacicola and by geographical areas. The analyses were based on 10,000 random permutations.

Genealogical relationships between mitochondrial haplotypes were reconstructed using TCS v 1.21 (Clement et al., 2000) with the method described by Templeton et al. (1992), with a probability cut-off set at 95%. For this analysis our 400 *B. oleae* mtDNA sequences were merged with 19 previously published sequences from Nardi et al. (2010) (GenBank accession numbers: GU108459 to GU108465 and GU108476 to GU108479, see Tab. S1) to obtain a final alignment for 419 *B. oleae* flies from 62

localities. Populations were grouped according to the 6 main geographical regions investigated with AMOVA. This method organizes haplotypes into networks according to genetic distance at the intra-specific level, especially in cases of multiple haplotypes that are derived from a single ancestral sequence (Templeton et al., 1992).

Symbiont-host haplotype analysis

The relationships between the symbiont and its insect host mitochondrial haplotypes were defined by a haplotype network implemented in TCS v 1.21 (Clement et al., 2000). The statistical Fisher's exact test (Miller, 1997) was used to test the hypothesis of a significant association between *Ca*. E. dacicola and its insect host haplotypes. In addition the Pearson chi-squared test was applied to evaluate the strength of each association between symbiont-host haplotypes.

Results

Haplotype analysis and distribution of Ca. E. dacicola

To investigate the genetic variability of *Ca* Erwinia dacicola among *B. oleae* populations, a region of 780 bp of the bacterial 16S rRNA gene was obtained. Hundreds of olive fly oesophageal bulbs were analyzed and a total of 400 bacterial sequences were obtained.

In all of them the symbiont *Ca.* E. dacicola was detected as the predominant species. However, other different bacteria species were detected with frequencies less than 5%, none belonging to *Acetobacter* spp.

Three different haplotypes were identified. Two of them (htA and htB), already described (Savio et al., 2012) (GenBank acc. ns: HQ667588 and HQ667589, respectively), and the third one (the new haplotype htC) showed intermediate features between the others two haplotypes. In particular three transitions were identified. At position 11 and 282 the haplotype htC presented a T as htA while at position 657 the new haplotype showed a G as the haplotype htB. All sequences with the haplotype htC were extended in the 3' direction up to 1280 bp using the primer fl2. In addition, according to Savio et al. (2012), two further mutations were identified at nucleotide position 943 and 945 where haplotype htC presented an A in both the point mutation, as htB.

In order to check for the co-presence of different bacterial haplotypes in a single olive fly, the bacterial contents of 7 oesophageal bulbs, chosen among populations that presented both the haplotypes htA and htB, were cloned (one sample from locations 7, 40, 43, 46, 48 and 2 samples from location 36, Tab. 3). Eight to 10 amplicons for each individual were analyzed with a restriction enzyme (Apa I) or directly sequenced. The results confirmed the presence of a unique *Ca.* E. dacicola haplotype in each olive fly oesophageal bulb, as already detected by Savio et al. (2012), suggesting that htA and htB do not coexist in the same olive fly.

The Mediterranean populations sampled harbored haplotypes htA and htB while the htC was reported only in the Pakistani population. Fig. 1 shows the geographical distribution of the three *Ca.* E. dacicola haplotypes. South-African population and

most of the Western populations of the Mediterranean basin (Iberian Peninsula, Morocco, Algeria, and France, including Corsica and Sardinia) presented exclusively haplotype htA. On the other hand, haplotype htB was exclusive of California population as well as Tunisia and Eastern populations (Egypt, Israel, Cyprus and Turkey). Populations located in the Central Mediterranean area (Italy, Croatia, Slovenia and Greece) mostly harbor both *Ca.* E. dacicola haplotypes htA and htB. The three haplotypes were distributed among olive flies sampled with those proportions: htA 52% (209/400), htB 45% (180/400) and htC 11% (11/400) as reported in Tab. 2.

SAMOVA was performed to identify genetic groups of *Ca.* E. dacicola populations among the 50 olive fly populations analyzed. The F_{CT} value reached a plateau for k = 3 (0.819; P<0.01) while the highest F_{CT} value (0.825; P<0.01) was obtained when the 50 populations were divided into k=4 groups. For k>4 the grouping started to disappear. Fixation indices (F) are reported for each k tested in Fig. 2. For k= 3, three genetically different groups were detected corresponding to: group 1, locations 1-12, 17, 19, 21, 25, 28, 34, 36, 37, 41, 42, 48; group 2, locations 13-16, 18, 20, 22-24, 26, 27, 29-33, 35, 38-40, 43, 44, 45-47, 50; group 3, locations 49. For k= 4, the four groups were group 1, locations 27 and 31, group 2, locations 13-16, 18, 20, 22-24, 26, 29, 30, 32, 33, 35, 38-40, 43 and groups 3 and 4 corresponding to group 1 and 3 from the previous (k=3) grouping. In k=3 Group 1 correspond with South-Africa, Iberian Peninsula, Morocco, Algeria, France including Corsica, Sardinia, and Crete; group 2 correspond with populations from Tunisia, Egypt, Italy, including Sicily, Slovenia, Croatia, Israel, Cyprus and Turkey and group 3 includes only Pakistan. These results show an East-West distribution of the *Ca.* E. dacicola haplotypes, with the exception of Crete (Fig. 1).

Mitochondrial haplotype analysis and population structure of B. oleae

A portion of 757 bp of the mitochondrial DNA of the olive fly, corresponding to the NADH dehydrogenase subunit 1, the leucine tRNA and the 16S mitochondrial region, was obtained on the same 400 individuals in which the symbiotic bacteria were amplified and sequenced. No deletions or insertions were detected and 25 polymorphic sites (3.3% of total length) were identified.

A total of 39 haplotypes were observed and designated from H1 to H39 (Tab. 3). Fig. 3 shows the geographic distribution of the mitochondrial haplotypes among the populations of *B. oleae*. Twenty haplotypes were unique of a population.

Three types of indices including haplotype distribution by locality and frequency among the populations analyzed, gene diversity (h), nucleotide diversity (π), average number of parwise differences (k) within populations and number of haplotypes (N_{HT}) were calculated to measure the genetic variability within olive fly populations (Tab. 3). Population from Marrakech revealed the highest number of haplotypes (N_{HT}=7) followed by Anopolis and Avignon (N_{HT}=6). On the other hand, populations from Serra de Leomil, Cuenca, Cairo, Fanna and Antakya exhibit only one haplotype. An average of 3.26 haplotypes per population was recovered. Considering populations with at least 5 samples, most of populations showed high level of genetic variability. Haplotype diversity (h) ranged from 0 (Antakya and Fanna) to 0.9333 (Athens, Greece) with the average overall value of 0.5753 while sequence divergence (π) among the haplotypes changed from 0 (Antakya and Fanna) to 0.0037 (Avignon, France) with the overall mean of 0.0016.

Frequencies and distribution of the 39 haplotypes found in this analysis plus one haplotype identified from GenBank sequences (Nardi et al., 2010) are listed and illustrated in Table S2 and Fig. 4. The 95% parsimony network (Fig. 4) was constructed using TCS v1.21 (Clement et al., 2000) grouping populations by geographical areas. Reticulations of this network were resolved following common theoretical predictions about network structure (Crandall and Templeton, 1993; Posada and Crandall, 2001). Three haplotypes, H6, H2 and H1, were found at higher frequencies comprising the 65.4% of all the individuals (274/419 individuals). Haplotype H6, comprising 29.1% was the most common haplotype shared by 122 individuals from 34 populations, where only Central Mediterranean populations represent the 43.6%. The others two most widespread haplotypes, H2 and H1, were found at a frequency of 24.8% (104/419 individuals) and 11.5% (48/419 individuals) respectively. In haplotype H2 64.7% (55/85 individuals) of the Western Mediterranean and 16.9% (42/248 individuals) of the Central Mediterranean samples were grouped together. All Eastern Mediterranean populations share H1, the third most common

haplotype, at a frequency of 72.5% (29/40 individuals) while it was not present in the Western Mediterranean area. African individuals were represented mostly by haplotype H35 and haplotype H39. In Pakistan four unique haplotypes (H33, H34, H36 and H40) were found (Tab. 3).

The haplotype network does not reveal a clear geographical structure. Nevertheless some populations as Africa and Pakistan are well structured and differentiated from the Mediterranean and American haplotypes. Moreover, as reported in Nardi et al. (2005) and Dogac et al. (2013), the network suggests an Eastern origin of Californian olive fly populations.

Tests of homogeneity among populations were performed using AMOVA. For this purpose, population groupings were chosen according to the SAMOVA results obtained for *Ca.* E. dacicola analyses (k=3 and k=4) and to the 6 major geographical regions of the sampling area (West, Central and Eastern Mediterranean areas, South Africa, California and Pakistan) (see Tab. 1).

AMOVA results, summarized in Tab. 4, revealed that significant differences among groups of populations were observed for the three SAMOVA hierarchical groups ($p=0.02$) and when olive fly populations were divided based on geographical areas ($p<0.001$). However, a significant population structure among groups was not detected for the SAMOVA groups k=4 ($p=0.07$).

Relationships between Ca. Erwinia and B. oleae haplotypes

The association between *Ca.* Erwinia and *B. oleae* haplotypes was investigated in order to detect the presence of strict vertical transmission. A haplotype network (Fig. 5) was constructed based on the three symbiont lineages and the insect haplotypes observed in our dataset to define symbiont-host relationships.

Considering all sampled localities, Fischer's exact test revealed a significant association between symbiotic bacteria and host haplotypes, with a phylogeographyc correlation corresponding to the main geographical regions ($p<0.01$). The relationships between the symbiont and the insect haplotypes are reported in Table 5. Pearson chi-squared test measured the strength of the single association between symbiont-host haplotypes. The most striking association concerned insect haplotypes H1, H2, H7,

H20 and H35 with bacterial haplotypes htA and htB. Haplotypes H33, H34 and H36, exclusively to the Pakistan population, were instead strongly related to htC. Furthermore, excluding from the analysis the Pakistani, South African and Californian populations, which could influence the significance of association due to their geographical isolation, and single and rare haplotypes, forcedly correlated to one of the symbiont lineages, the Fisher exact test was still significative ($p<0.01$).

Most of the populations comprised in the most frequent *B. oleae* haplotypes (H1, H6 and H2), as described in the insect mtDNA data analysis, reflect the subdivision of the Mediterranean area into three main regions: Eastern Mediterranean-America, Central Mediterranean, and Western Mediterranean, respectively (Fig. 4). A non random subdivision of the Mediterranean basin is also observed in the symbiont-host haplotype distribution. Olive fly populations sharing haplotype H2, harbor at a frequency of 69.3% (70/101), the bacterial haplotype htB; in contrast populations within haplotype H1 present 93.2% (41/44) of the bacterial genetic variant htA. In populations with haplotype H6, mostly located in the Central Mediterranean region, 55.4% and 44% of symbiotic bacteria show the haplotypes htB and htA respectively (Tab. 5). The South African population, sharing the haplotype H35 at a frequency of 93.8%, pronominally harbors the bacterial haplotype htA (15/16) while the Californian olive fly population harbors only the variant htB (8/8). Populations with mitochondrial haplotypes H33, H34, H35 and H36 plus one sample with H6 harbored the third haplotype htC.

Discussion

In the present study, we investigate the genetic variability of the olive fly symbiont, *Ca.* E. dacicola, together with the patterns of mitochondrial genetic differentiation of *B. oleae* over a large area of the insect geographical distribution, including most regions of the Mediterranean area, plus South Africa, California and Pakistan.

Genetic variability and pattern of distribution of Ca. E. dacicola

Three bacterial haplotypes of *Ca.* E. dacicola were identified with a significative geographic distribution. Haplotypes htA and htB, previously recovered in Italian populations (Savio et al., 2012), were distributed among Mediterranean, South African and Californian olive fly populations while the new one (htC) was exclusive of the Pakistani population (Fig. 1, Tab. 2). The co-existence of different *Ca.* E. dacicola haplotypes in a single fly was not detected, as reported in Savio et al. (2012), although it is common to find more than one larva inside the same olive and therefore different haplotypes can frequently come in contact.

Our data (SAMOVA) revealed a significant genetic structure of the symbiont haplotypes distribution in the sampling populations showing a clear East-West genetic differentiation in the Mediterranean basin (Fig. 1). Western and South African populations mainly harbor the bacterial symbiont haplotype htA while most of the olive flies from Eastern Mediterranean, California and Tunisia harbor the genetic variant htB. The two populations sampled in Crete are the most striking exception of *Ca.* E. dacicola haplotype distribution. The prevalence of htA in these two island populations could be explained by the geographical isolation, which contributed to the absence of external gene flow. Although the commercial trades with its neighbors have been extensive since the introduction of the olive cultivation, it is likely that local fly populations were largely contained in the island. This hypothesis is also supported by the presence of exclusive mitochondrial haplotypes of *B. oleae,* H15 and H16, in the island. A restricted gene flow is also observed in the Pakistani populations in which a significant association between the exclusive symbiont lineage (htC) and insect haplotypes was detected. The Pakistani group is isolated from the rest of the

distribution range by natural barriers, such as mountains and desert areas, which probably strongly reduce the gene flow.

On the other hand, in the Central Mediterranean area a combination of both haplotypes (htA and htB) has been found among the populations. This region seems to be a confluence region between Eastern and Western area, where the presence of both symbionts haplotypes in the same populations could be the result of losses and new acquisitions events bypassing the model of strict vertical transmission. It could be hypnotized that the horizontal transmission could happen between larvae living on the same olive.

Population structure of B. oleae

A total of 39 haplotypes based on mitochondrial markers were observed in the *B. oleae* populations analyzed (Fig. 3, Tab. S1). Mitochondrial haplotype analysis (AMOVA) reveals a clear structural pattern supporting the subdivision of this area into different geographical groups. According to some authors, the geographical topography, together with human mediated activities, have played an important role in shaping the genetic structure of this fly, a monophagous species tightly related to cultivated olive trees and wild relatives (Nardi et al., 2005). Consistent with all available data, including mitochondrial and nuclear markers (Nardi et al., 2005, 2010), Pakistan and South African haplotypes are well structured and differentiated from the Mediterranean and American haplotypes, showing the phylogeographic structure of those populations. Furthermore, strong evidence for this differentiation is provided by the identification of six unique mitochondrial haplotypes (H33-H36, H39, H40) forming an isolated clade separated from all the other populations (Fig. 4).

It was suggested that historical gene flow might have been negligible or absent in these areas, where olive trees are spottily distributed, in contrast with the Mediterranean region, where olive groves have a continuous distribution. Thus, long term isolation and low population numbers, suggested by the low levels of intra group differentiation, support the hypothesis that drift might have played a significant role in the evolution of those populations. Moreover, the haplotype distribution indicates that Californian populations are related to the Eastern Mediterranean olive fly where Turkey is considered the main possible source of American populations (Zygouridis et

al., 2009; Dogac et al., 2013), probably caused by a single introduction event and later dispersion (Segura et al., 2008). The presence of symbiont haplotype htB, found in both America and all Eastern Mediterranean populations, strongly support this hypothesis.

The analyzed populations are characterized by a substantial degree of genetic differentiation that seems to be characteristic of this species on a larger geographic scale (Nardi et al., 2005, 2010; Segura et al., 2008; Zygouridis et al., 2009), although an extensive gene flow among olive fly populations has been found in the Mediterranean basin. The high variability of *B. oleae* is probably due to the time that elapsed since it became established in the Mediterranean and the high population densities characteristic of this species.

It is interesting to note that the clear genetic differentiation from the Eastern and Western regions of the Mediterranean basin found in the olive fly symbiont and in its host has been also observed in wild olives trees (Lumaret et al., 2004; Besnard et al., 2007), which was explained by a re-colonization of the insect from different glacial refugia (Besnard et al., 2002). Even though it has been found that the bacterial genetic variant htA was present in both the Western Mediterranean and South African populations, suggesting a more extensive explanation for this phenomenon.

Conclusion

The present study shades light on the genetic differentiation of the olive fruit fly populations together with their symbiotic bacteria variability and offers an important opportunity for studying the long host-symbiont coevolutionary life history. The vertical transmission seems to be the predominant mechanism of symbiotic bacterial transmission but it reasonable to hypothesize the presence of some horizontal events. In this regard, further investigations could be useful to elucidate the details of the bacterial transmission. Moreover, the symbiont haplotypes distribution is well structured and phylogeographically differentiated in contrast to the high genetic variability found in the insect host lineages. This represents an useful tool to reconstruct the debated olive fly's historical origin and its colonization route and could open new scenarios of population movements related to glaciation, Messinian salinity crisis, or even more extensive migrations concerning the African continent.

Acknowledgements

We thank all colleagues who helped us with sampling collection: Y. Ben-Dov, A. Bozzolo, V. Caleca, E. Cargnus, S. Deliperi, N. Mori, C. Duso, M.M. El Idrissi, I. Elmaghraby, M. Faccoli, P. Fontana, D. Fornasiero, G. Frigimelica, H. Genç, J. Giliomee, C. Manoli, M. Marzaro, B. Massa, H. Mladi, F. Mraichi, M.E. Noce, G. Pellizzari, E. Petrucco Toffolo, C. Rapisarda, M.A. Rashid, C. Rempoulakis, P. Salvato, C. Sañudo, C. Sartori, V. Vacante, C. Villari, L. Vivan and M. Borin.

Tab. 1 - Collection details for the 50 *B. oleae* populations sampled in the Mediterranean basin, California, South Africa and Pakistan. The Mediterranean area is divided in 3 geographical areas (West, Central and East region). Geographic coordinates, expressed in decimal degrees, and the collection date are reported.

Region	Country	Location	Coordinates (lat./long.)	Date
West Mediterranean	Portugal	Bragança	41°48'N 6°45'W	10/2007
West Mediterranean	Portugal	Lisboa	38°31'N 8°55'W	10/2007
West Mediterranean	Portugal	Serra de Leomil	40°57'N 7°40'W	10/2007
West Mediterranean	Portugal	Tavira	37°07'N 7°38'W	10/2009
West Mediterranean	Spain	Cuenca	40°04'N 2°07'W	02/2007
West Mediterranean	Spain	Almeria	36°50'N 2°28'W	10/2007
West Mediterranean	Spain	Valencia	39°28'N 0°23'E	09/2008
West Mediterranean	Spain	Tarragona	41°06'N 1°13'E	11/2009
West Mediterranean	Spain	Barcelona	41°24'N 2°09'E	10/2010
West Mediterranean	Morocco	Oujda	34°37'N 1°56'E	12/2009
West Mediterranean	Morocco	Marrakech	31°37' N 8°00'W	03/2012
West Mediterranean	Algeria	Constantine	36°21'N 6°36'E	12/2009
Central Mediterranean	Tunisia	Chaffar	34°34'N 10°33'E	01/2009
Central Mediterranean	Italy	Bitetto	41°02'N 16°45'E	11/2002
Central Mediterranean	Italy	Taormina	38°51'N 15°17'E	08/2007
Central Mediterranean	Italy	Montegrotto Terme	45°19'N 11°47'E	12/2007
Central Mediterranean	Italy	Imperia	43°53'N 8°03'E	01/2008
Central Mediterranean	Italy	Marsala	37°47'N 12°26'E	07/2008
Central Mediterranean	Italy	Alghero	40°33'N 8°19'E	10/2008
Central Mediterranean	Italy	Thiene	45°42′N 11°28′E	10/2008
Central Mediterranean	Italy	Bardolino	45°32'N 10°43'E	12/2008
Central Mediterranean	Italy	Ancona	43°37'N 13°31'E	09/2009
Central Mediterranean	Italy	Portici	40°48'N 14°20'E	09/2009
Central Mediterranean	Italy	Rossano Calabro	39°34'N 16°38'E	09/2009
Central Mediterranean	Italy	Cagliari	39°13′N 9°07′E	10/2009
Central Mediterranean	Italy	Portoferraio	42°82'N 10°32'E	10/2009
Central Mediterranean	Italy	Campione	45°58'N 08°58'E	10/2009
Central Mediterranean	Italy	Otranto	40°08'N 18°29'E	10/2009
Central Mediterranean	Italy	Roma	41°53'N 12°29'E	12/2009

Central Mediterranean	Italy	Gioia Tauro	38°27′N 15°54′E	12/2009
Central Mediterranean	Italy	Camogli	44°20'N 9°09'E	11/2010
Central Mediterranean	Italy	Fanna	46°11′N 12°45′E	12/2011
Central Mediterranean	Italy	Sciacca	37°31'N 13°03'E	12/2011
Central Mediterranean	Greece	Cephalonia	38°10'N 20°34'E	10/2007
Central Mediterranean	Greece	Athens	37°56'N 23°01'E	12/2007
Central Mediterranean	Greece	Heraklion	35°19'N 25°08'E	03/2009
Central Mediterranean	Greece	Anopolis	35°12'N 24°06'E	04/2010
Central Mediterranean	Slovenia	Strunjan	45°31'N 13°34'E	09/2008
Central Mediterranean	Croatia	Pag	44°29'N 14°57'E	08/2008
Central Mediterranean	Croatia	Pula	44°52'N 13°50'E	10/2008
Central Mediterranean	France	Avignon	43°55'N 4°48'E	11/2010
Central Mediterranean	France	Ile-Rousse	42°38′N 8°56′E	12/2011
Central Mediterranean	Turkey	Çanakkale	40°09'N 26°24'E	10/2008
East Mediterranean	Turkey	Antakya	36°12′N 36°09′E	02/2013
East Mediterranean	Egypt	Cairo	30°03′N 31°13′E	06/2011
East Mediterranean	Cyprus	Limassol	32°57'N 34°45'E	11/2008
East Mediterranean	Israel	Bet Dagan	31°59'N 34°49'E	12/2007
South-Africa	South-Africa	Stellenbosch	33°56'S 18°51'E	05/2008
Pakistan	Pakistan	Swat	35°21'N 72°11'E	10/2010
USA	California	San Francisco	34°25'N 119°42'W	12/2011

Tab. 2 - '*Ca.* Erwinia dacicola' haplotype frequencies among the olive fly sampled populations. *N*, number of individuals analyzed.

	Country	Location	N	htA	htB	htC		Country	Locality	N	htA	htB	htC
1	Portugal	Bragança	5	5	0	0	26	Italy	Portoferraio	6	2	4	0
2	Portugal	Lisboa	9	9	0	0	27	Italy	Campione	7	3	4	0
3	Portugal	Serra de Leomil	4	4	0	0	28	Italy	Otranto	7	6	1	0
4	Portugal	Tavira	7	7	0	0	29	Italy	Roma	6	2	4	0
5	Spain	Cuenca	2	2	0	0	30	Italy	Gioia Tauro	3	1	2	0
6	Spain	Almeria	6	6	0	0	31	Italy	Camogli	8	4	4	0
7	Spain	Valencia	11	10	1	0	32	Italy	Fanna	5	0	5	0
8	Spain	Tarragona	3	3	0	0	33	Italy	Sciacca	7	2	5	0
9	Spain	Barcelona	7	7	0	0	34	Greece	Cephalonia	6	5	1	0
10	Morocco	Oujda	4	4	0	0	35	Greece	Athens	6	1	5	0
11	Morocco	Marrakech	25	22	3	0	36	Greece	Heraklion	15	15	0	0
12	Algeria	Constantine	2	2	0	0	37	Greece	Anopolis	13	13	0	0
13	Tunisia	Chaffar	14	0	14	0	38	Slovenia	Strunjan	9	2	7	0
14	Italy	Bitetto	9	1	8	0	39	Croatia	Pag	10	0	10	0
15	Italy	Taormina	8	0	8	0	40	Croatia	Pula	10	2	8	0
16	Italy	M. Terme	8	2	6	0	41	France	Avignon	13	13	0	0
17	Italy	Imperia	6	6	0	0	42	France	Ile-Rousse	14	12	2	0
18	Italy	Marsala	6	0	6	0	43	Turkey	Çanakkale	9	0	9	0
19	Italy	Alghero	8	8	0	0	44	Turkey	Antakya	5	0	5	0
20	Italy	Thiene	8	0	8	0	45	Egypt	Cairo	1	0	1	0
21	Italy	Bardolino	7	5	2	0	46	Cyprus	Limassol	14	0	14	0
22	Italy	Ancona	7	0	7	0	47	Israel	Bet Dagan	7	0	7	0
23	Italy	Portici	6	0	6	0	48	South-Africa	Stellenbosch	16	16	0	0
24	Italy	R. Calabro	6	1	5	0	49	Pakistan	Swat	11	0	0	11
25	Italy	Cagliari	6	6	0	0	50	California	San Francisco	8	0	8	0
								Tot.		400	209	180	11

Tab. 3 – Summary of genetic diversity indices of *B. oleae* populations analyzed. *N*, number of individuals studied; N_{HT}, total number of haplotypes for each sampling location; h, haplotype diversity; π, nucleotide diversity; k, mean number of pairwise differences per sequence.

Country	Location	Haplotypes distribution	*N*	N_{HT}	h	π	k
Portugal	Bragança	2H2, 1H4, 2H7	5	3	0.8000	0.0013	1.000
Portugal	Lisboa	4H2, 2H4, 2H6, 1H7	9	4	0.7778	0.0015	1.166
Portugal	Serra de Leomil	4H2	4	1	0.0000	0.0000	0.000
Portugal	Tavira	6H2, 1H22	7	2	0.2857	0.0003	0.285
Spain	Cuenca	2H2	2	1	0.0000	0.0000	0.000
Spain	Almeria	4H2, 1H4, 1H5	6	3	0.6000	0.0008	0.666
Spain	Valencia	8H2, 3H7	11	2	0.4364	0.0005	0.436
Spain	Tarragona	2H2, 1H4	3	2	0.6667	0.0008	0.666
Spain	Barcelona	5H2, 2H6	7	2	0.4762	0.0012	0.952
Morocco	Oujda	1H2, 2H4, 1H6	4	3	0.8333	0.0013	1.000
Morocco	Marrakech	2H1, 16H2, 3H4, 1H6, 1H7, 1H31, 1H32	25	7	0.5867	0.0009	0.753
Algeria	Constantine	1H2, 1H4	2	2	1.000	0.0013	1.000
Tunisia	Chaffar	1H2, 2H4, 7H6, 1H20, 3H26	14	5	0.7253	0.0015	1.175
Italy	Bitetto	1H1, 2H2, 4H6, 1H14, 1H20	9	5	0.8056	0.0022	1.666
Italy	Taormina	2H1, 2H4, 3H25, 1H26	8	4	0.8214	0.0023	1.750
Italy	M. Terme	1H2, 5H6, 1H9, 1H26	8	4	0.6429	0.0015	1.178
Italy	Imperia	5H6, 1H7	6	2	0.3333	0.0013	1.000
Italy	Marsala	1H1, 2H2, 1H4, 2H6	6	4	0.8667	0.0019	1.466
Italy	Alghero	1H2, 5H6, 1H7, 1H27	8	4	0.6429	0.0017	1.357
Italy	Thiene	2H1, 2H2, 3H6, 1H28	8	4	0.8214	0.0024	1.821
Italy	Bardolino	2H1, 2H6, 3H10	7	3	0.7619	0.0026	2.000
Italy	Ancona	5H2, 2H6	7	2	0.4762	0.0012	0.952
Italy	Portici	2H6, 1H9, 1H10, 2H26	6	4	0.8667	0.0032	2.466
Italy	R. Calabro	5H2, 1H6	6	2	0.3333	0.0008	0.666
Italy	Cagliari	2H2, 3H6, 1H21	6	3	0.7333	0.0018	1.400
Italy	Portoferraio	1H1, 3H6, 2H10	6	3	0.7333	0.0030	2.333
Italy	Campione	1H4, 3H6, 2H10, 1H22	7	4	0.8095	0.0031	2.380
Italy	Otranto	1H2, 1H4, 4H6, 1H24	7	4	0.7143	0.0013	1.047
Italy	Roma	1H2, 5H6	6	2	0.3333	0.0008	0.666
Italy	Gioia Tauro	1H2, 1H8, 1H20	3	3	1.000	0.0035	2.666

Italy	Camogli	2H2, 5H6, 1H23	8	3	0.6071	0.0020	1.571
Italy	Fanna	5H20	5	1	0.0000	0.0000	0.000
Italy	Sciacca	2H2, 5H6	7	2	0.4762	0.0012	0.952
Greece	Cephalonia	1H4, 4H7, 1H17	6	3	0.6000	0.0020	1.533
Greece	Athens	1H1, 1H2, 2H4, 1H8, 1H14	6	5	0.9333	0.0027	2.066
Greece	Heraklion	1H1, 4H2, 8H6, 2H16	15	4	0.6667	0.0017	1.333
Greece	Anopolis	1H1, 2H2, 5H6, 1H8, 3H10, 1H15	13	6	0.8205	0.0029	2.230
Slovenia	Strunjan	1H1, 3H2, 4H6, 1H7	9	4	0.7500	0.0020	1.555
Croatia	Pag	3H2, 4H6, 1H7, 1H18, 1H19	10	5	0.8000	0.0025	1.911
Croatia	Pula	1H2, 8H6, 1H8,	10	3	0.3778	0.0007	0.600
France	Avignon	2H6, 5H8, 2H10, 1H11, 2H12, 1H13	13	6	0.8333	0.0037	2.846
France	Ile-Rousse	2H2, 1H4, 9H6, 1H8, 1H9	14	5	0.5934	0.0012	0.912
Turkey	Çanakkale	2H1, 1H2, 5H6, 1H37	9	4	0.6944	0.0020	1.555
Turkey	Antakya	5H1	5	1	0.0000	0.0000	0.000
Egypt	Cairo	1H1	1	1	1.000	0.0000	0.000
Cyprus	Limassol	12H1, 1H2, 1H3	14	3	0.2747	0.0003	0.285
Israel	Bet Dagan	5H1, 1H29, 1H30	7	3	0.5238	0.0010	0.761
South-Africa	Stellenbosch	15H35, 1H39	16	2	0.1250	0.0001	0.125
Pakistan	Swat	1H6, 4H33, 1H34, 1H35, 4H36	11	5	0.7818	0.0020	1.563
California	San Francisco	4H1, 1H6, 3H38	8	3	0.6786	0.0021	1.607

Tab. 4 Analysis of molecule variance (AMOVA) for three different groups of *B. oleae* populations based on SAMOVA results of *Ca.* E. dacicola analyses and geographic subdivisions.

Structure	Source of variation	Variance (%)	Fixation indices	p value
Grouping by SAMOVA results (k=3)	Among groups	7.89	F_{CT}=0.07885	0.02
	Among populations within groups	32.30	F_{SC}=0.35068	<0.001
	Within populations	59.81	F_{ST}=0.40188	<0.001
Grouping by SAMOVA results (k=4)	Among groups	6.37	F_{CT}= 0.06369	0.07
	Among populations within groups	33.25	F_{SC}=0.35512	<0.001
	Within populations	60.38	F_{ST}=0.39620	<0.001
Grouping by geographical regions	Among groups	41.80	F_{CT}=0.41797	<0.001
	Among populations within groups	7.51	F_{SC}=0.12900	<0.001
	Within populations	50.69	F_{ST}=0.49305	<0.001

Tab. 5 – Relationship between *Ca.* Erwinia dacicola (htA, htB, htC) and *B. oleae* haplotypes (H1-H39). N, number of individuals for each insect haplotype.

***B.oleae* haplotypes**	***N***	***Ca.* Erwinia dacicola haplotypes**		
		htA	**htB**	**htC**
H1	44	3	41	0
H2	101	70	31	0
H3	1	0	1	0
H4	22	14	8	0
H5	1	1	0	0
H6	119	52	66	1
H7	15	14	1	0
H8	10	8	2	0
H9	3	2	1	0
H10	13	10	3	0
H11	1	1	0	0
H12	2	2	0	0
H13	1	1	0	0
H14	2	0	2	0
H15	1	1	0	0
H16	2	2	0	0
H17	1	0	1	0
H18	1	0	1	0
H19	1	0	1	0
H20	8	0	8	0
H21	1	1	0	0
H22	2	1	1	0
H23	1	1	0	0
H24	1	1	0	0
H25	3	3	0	0
H26	7	0	7	0
H27	1	1	0	0
H28	1	0	1	0
H29	1	0	1	0
H30	1	0	1	0
H31	1	1	0	0
H32	1	1	0	0
H33	4	0	0	4
H34	1	0	0	1
H35	16	15	0	1
H36	4	0	0	4
H37	1	0	1	0
H38	3	0	3	0
H39	1	1	0	0
Tot.	400	207	182	11

Tab. S1 - *B. oleae* haplotypes distribution.

	Country	Location	H 1	H 2	H 3	H 4	H 5	H 6	H 7	H 8	H 9	H 10	H 11	H 12	H 13	H 14	H 15	H 16	H 17	H 18	H 19	H 20	H 21	H 22	H 23	H 24	H 25	H 26	H 27	H 28	H 29	H 30	H 31	H 32	H 33	H 34	H 35	H 36	H 37	H 38	H 39
1	Portugal	Bragança	0	2	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Portugal	Lisboa	0	4	0	2	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Portugal	Serra Leomil	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Portugal	Tavira	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	Portugal	Paradela	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	Spain	Cuenca	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	Spain	Almeria	0	4	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	Spain	Valencia	0	8	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	Spain	Tarragona	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	Spain	Barcelona	0	5	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	Morocco	Oujda	0	1	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	Morocco	Marrakech	2	17	0	3	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
13	Algeria	Constantine	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	Algeria	Sig City	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	Tunisia	Chaffar	0	1	0	2	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0
16	Italy	Bitetto	1	2	0	0	0	4	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	Italy	Bari	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	Italy	Catania	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	Italy	Vaggia	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	Italy	Taormina	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0
21	Italy	Montegrotto T.	0	1	0	0	0	5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
22	Italy	Imperia	0	0	0	0	0	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	Italy	Marsala	1	2	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	Italy	Alghero	0	1	0	0	0	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
25	Italy	Thiene	2	2	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

Tab. S1 - *B. oleae* haplotypes distribution.

26	Italy	Bardolino	2	0	0	0	0	2	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	Italy	Ancona	0	5	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	Italy	Portici	0	0	0	0	0	2	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
29	Italy	Rossano C.	0	5	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	Italy	Cagliari	0	2	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	Italy	Portoferraio	1	0	0	0	0	3	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	Italy	Campione	0	0	0	1	0	3	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	Italy	Otranto	0	1	0	1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	Italy	Roma	0	1	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	Italy	Gioia Tauro	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	Italy	Camogli	0	2	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	Italy	Fanna	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	Italy	Sciacca	0	2	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	Greece	Cephalonia	0	0	0	1	0	0	4	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	Greece	Athens	1	1	0	2	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	Greece	Heraklion	1	4	0	0	0	8	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42	Greece	Anopolis	1	2	0	0	0	5	0	1	0	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	Slovenia	Strunjan	1	3	0	0	0	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	Croatia	Pag	0	3	0	0	0	4	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	Croatia	Pula	0	1	0	0	0	8	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
46	France	Avignon	0	0	0	0	0	2	0	5	0	2	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	France	Ile-Rousse	0	2	0	1	0	9	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	Turkey	Çanakkale	2	1	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
49	Turkey	Antakya	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	Turkey	Osmaniye	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	Egypt	Cairo	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Tab. S1 - *B. oleae* haplotypes distribution.

52	Cyprus	Limassol	12	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
53	Israel	Bet Dagan	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
54	Israel	Haifa	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
55	S. Africa	Stellenbosch	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	1
56	S. Africa	Paarl Mtn.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0
57	Pakistan	Swat	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1	1	4	0	0	0
58	Pakistan	Cherat	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
59	Pakistan	Malakand	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
60	California	San Francisco	4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
61	California	Santa Barbara	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
62	California	Oroville	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
		Tot.	48	10 4	1	22	1	12 2	15	10	3	13	1	2	1	2	1	2	1	1	1	8	1	3	1	1	3	7	1	1	1	1	1	1	5	1	18	6	1	5	1

Tab. S2 - List of sequences from Nardi et al. (2010) included in our analyses.

* new haplotype.

GeneBank a/n	**Country**	**Location**	**Identical haplotypes in our study**
GU108467	Pakistan	Cherat	H36
GU108463	Pakistan	Cherat	H36
GU108477	Pakistan	Malakand	H33
GU108469	Pakistan	Malakand	H40*
GU108478	South-Africa	Paarl Mtn.	H35
GU108462	South-Africa	Paarl Mtn.	H35
GU108465	Turkey	Osmaniye	H1
GU108461	Turkey	Osmaniye	H1
GU108460	Israel	Haifa	H1
GU108472	Israel	Haifa	H1
GU108474	Morocco	Marrakesh	H2
GU108468	Algeria	Sig City	H2
GU108473	Portugal	Paradela	H22
GU108459	Italy	Bari	H6
GU108470	Italy	Bari	H6
GU108471	Italy	Catania	H2
GU108464	Italy	Vaggia	H6
GU108479	California	Santa Barbara	H38
GU108475	California	Oroville	H38

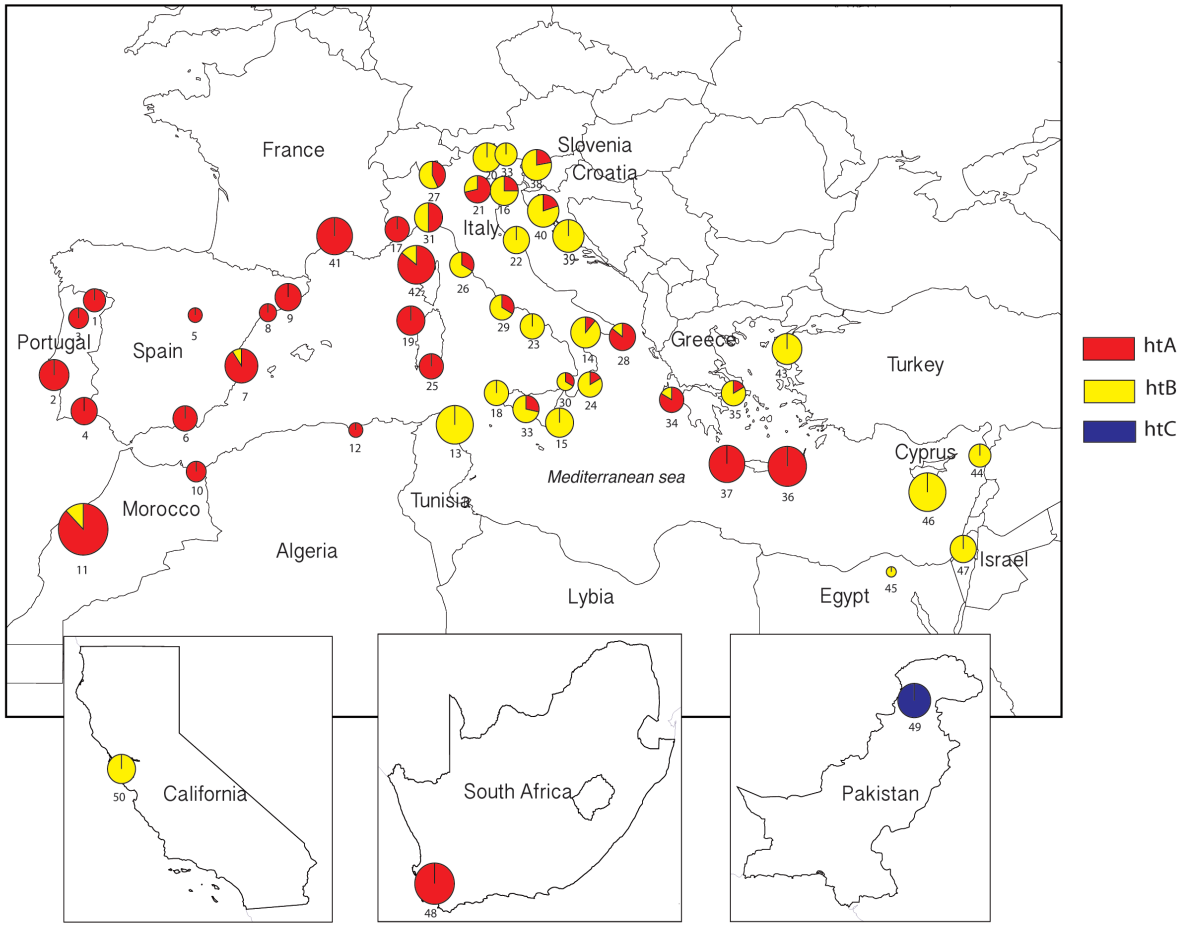

Fig. 1 - Geographic distribution of the three '*Ca.* Erwinia dacicola' haplotypes among the 50 *B. oleae* populations sampled in the Mediterranean basin, California, South Africa and Pakistan coded according to Tab. 2. The areas of the circles are proportional to the sampling.

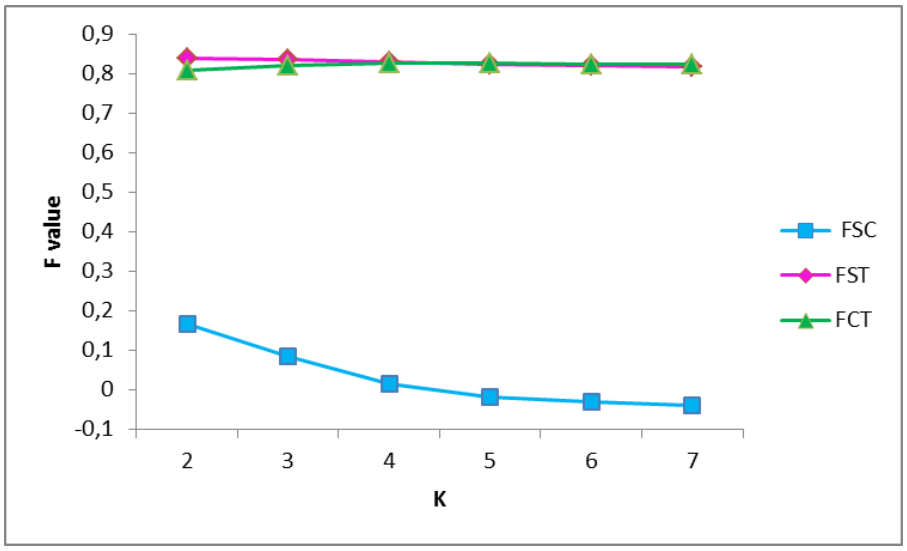

Fig. 2 Values of fixation indices (F), obtained from SAMOVA for *Ca.* Erwinia dacicola as a function of increasing number of groups (k). F_{CT}: differentiation between groups; F_{ST}: differentiation between populations among groups; F_{SC}: differentiation between populations within groups.

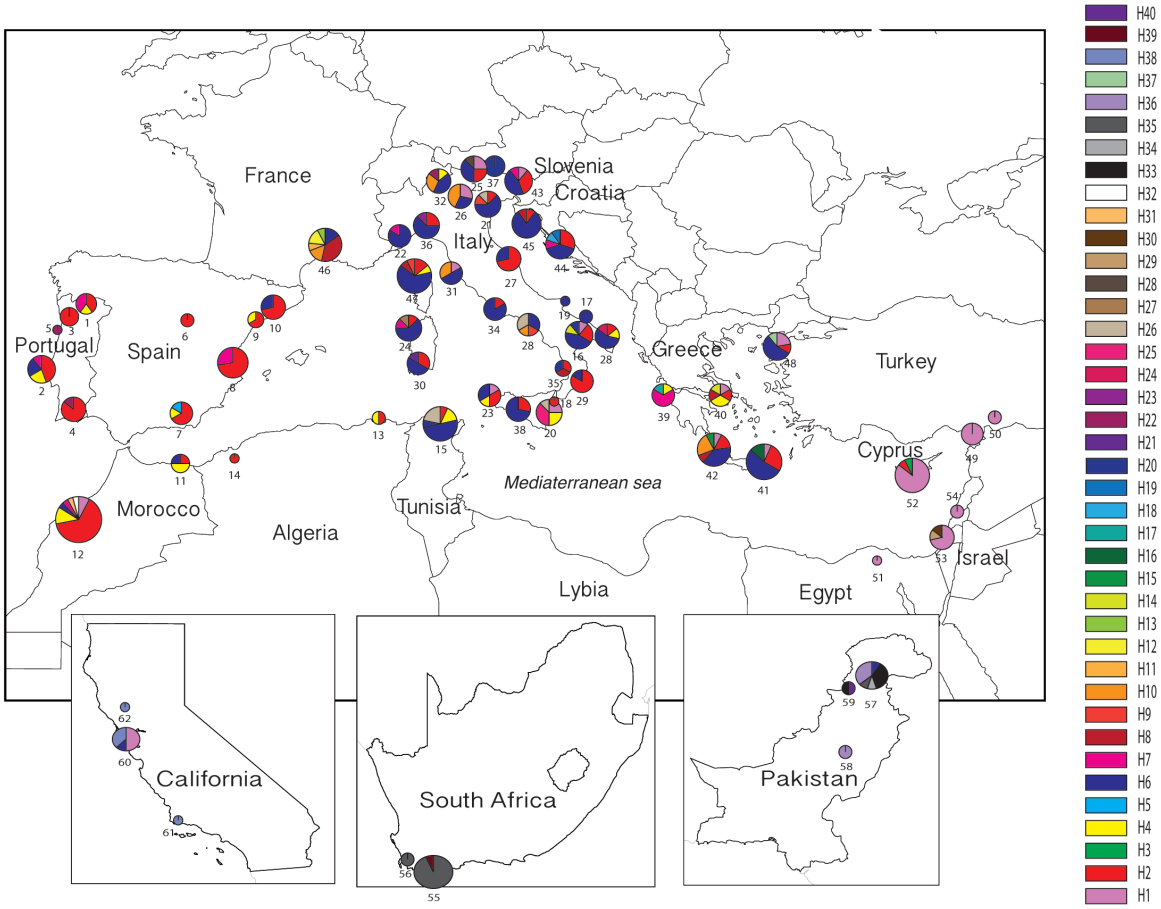

Fig. 3 - Geographic distribution of mitochondrial haplotypes observed among Mediterranean, Californian, South African and Pakistani populations of *B. oleae* coded according to Tab. S2. The areas of the circles are proportional to the sampling.

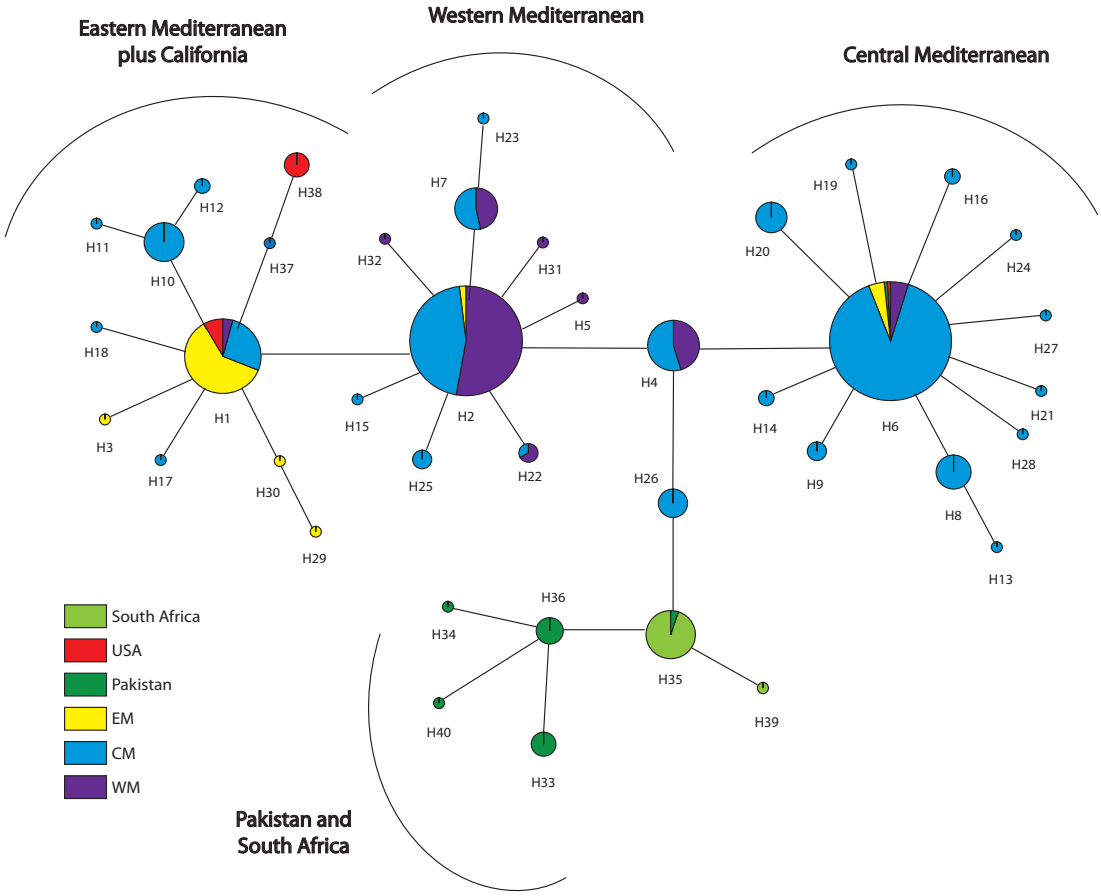

Fig. 4 - Haplotype network of *B. oleae* populations inferred by TCS v1.21 (Clement et al., 2000). The size of the circles is proportional to the haplotype frequency. Sampling region of each haplotype is color coded as in the legend.

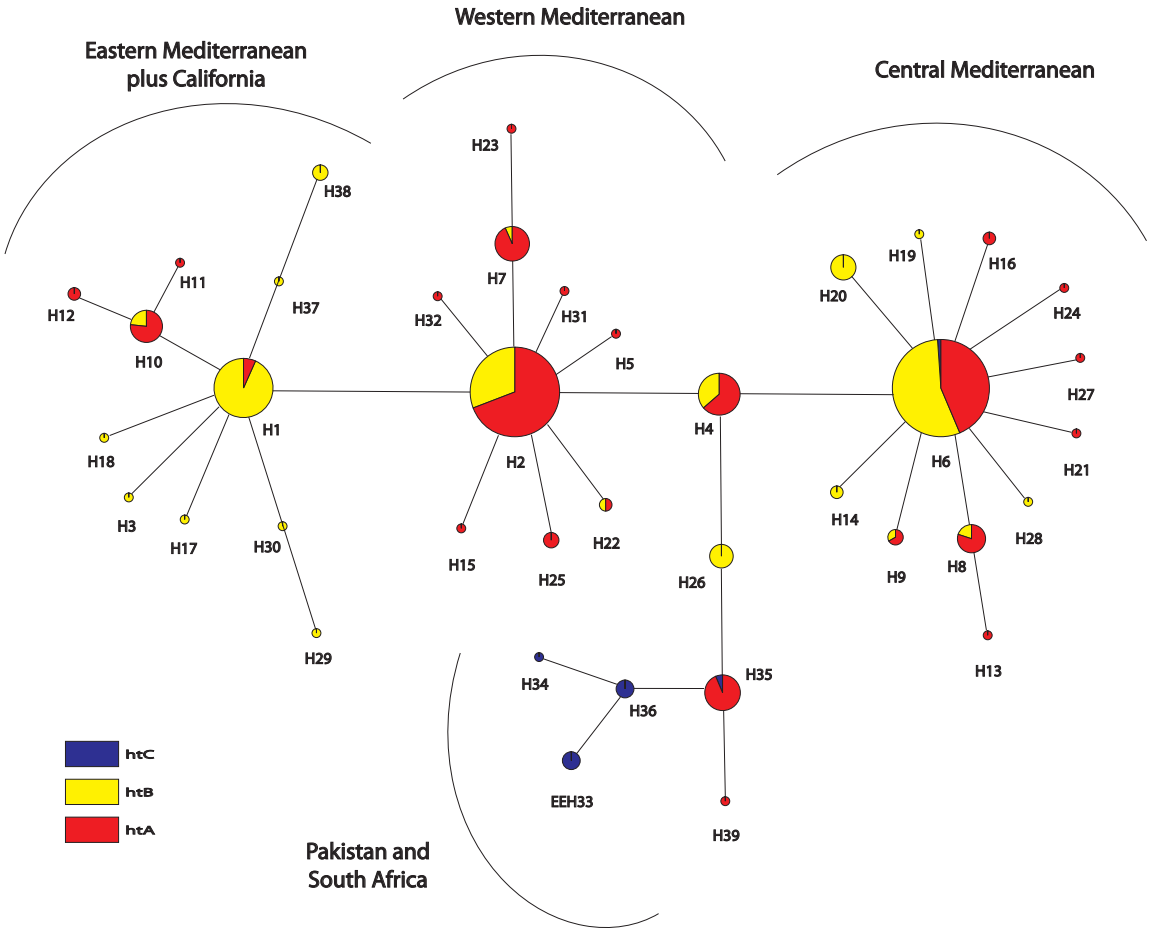

Fig. 5 - Haplotype network of *Ca.* Erwinia dacicola among *B. oleae* populations inferred by TCS v1.21 (Clement et al., 2000). The size of the circles is proportional to the haplotype frequency and color coded as in the legend.

References

Augustinos, A.A., Stratikopoulos E.E., Zacharopoulou A., Mathiopoulos K.D., 2002. Polymorphic microsatellite markers in the olive fly, *Bactrocera oleae*. Mol. Ecol. Notes. 2, 278–280.

Augustinos, A.A., Mamuris, Z., Stratikopoulos, E.E., D'Amelio, S., Zacharopoulou, A., Mathiopoulos, K.D., 2005. Microsatellite analysis of olive fly populations in the Mediterranean indicates a westward expansion of the species. Genetica. 125, 231-241.

Besnard, G., Khadari, B., Baradat, P., Bervillè, A., 2002. *Olea europea* (Oleaceae) phylogeography based on chloroplast DNA polymorphism. Theor. Appl. Genet. 104, 1353–1361.

Besnard, G., Rubio de Casas, R., Vargas, P., 2007. Plastid and nuclear DNA polymorphism reveals historical processes of isolation and reticulation in the olive tree complex (Olea *europaea)*. J. Biogeogr. 34, 736–752.

Crandall, K.A. and Templeton, A.R., 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phytogeny reconstruction. Genetics. 134, 959-969.

Capuzzo, C., Firrao, G., Mazzon, L., Squartini, A., Girolami, V., 2005. '*Candidatus* Erwinia dacicola', a coevolved symbiotic bacterium of the olive fly *Bactrocera oleae* (Gmelin). Int. J. Syst. Evol. Microbiol. 55, 1641–1647.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1659.

Copeland, R.S., White, I.M., Okumu, M., Machera, P., Wharton, R.A., 2004. Insects associated with fruits of the oleaceae (Asteridae, Lamiales) in Kenya, with special reference to the Tephritidae (Diptera). Bishop Mus. Bull. Entomol. 12, 135-164.

Crawford, M.H., 2007. Genetic structure of circumpolar populations: a synthesis. American J. Hum. Biol. 19, 203–217

Daane, K.M. and Johnson, M.W., 2010. Olive Fruit Fly: Managing an Ancient Pest in Modern Times. Annu. Rev. Entomol. 55, 151–69.

Dogac, E., Kandemir, I., Taskin, V., 2013. The genetic polymorphisms and colonization process of olive fly populations in Turkey. PloS One 8, e56067.

Estes, A.M., Hearn, D.J., Bronstein, J.L., Pierson, E.A., 2009. The olive fly endosymbiont, "*Candidatus* Erwinia dacicola", switches from an intracellular existence to an extracellular existence during host insect development. Appl. Environ. Microbiol. 75, 7097–7106.

Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 131, 479–491.

Excoffier, L., Laval, G., Schneider, S., 2007. ARLEQIN (ver. 3.0): An integrated software package for population genetics data analysis. Evol. Bioinform. Online. 1, 47–50.

Lumaret, R., Ouazzani, N., Michaud, H., Vivier, G., Deguilloux, M.F., Di Giusto, F., 2004. Allozyme variation of oleaster populations (wild olive tree) (Olea *europea* L.) in the Mediterranean Basin. Heredity. 92, 343–351.

Girolami, V., 1973. Reperti morfo-istologici sulle batteriosimbiosi del Dacus oleae Gmelin e di altri ditteri tripetidi, in natura e negli allevamenti su substrati artificiali. Redia 54, 269–294.

Girolami, V. and Cavalloro, R., 1972. Aspetti della simbiosi batterica di Dacus oleae (Gmelin) in natura e negli allevamenti di laboratorio. Ann. Soc. Entomol. Fr. 8, 561–571 (in Italian).

Hagen, K. S., 1966. Dependence of the olive fly, Dacus oleae, larvae on symbiosis with *Pseudomonas savastanoi* for the utilization of olive. Nature. 209, 423–424.

Kounatidis, I., Crotti, E., Sapountzis, P., Sacchi, L., Rizzi, A., Chouaia, B., Bandi, C., Alma, A. Daffonchio, D., Mavragani-Tsipidou, P., Bourtzis, K., 2009. *Acetobacter tropicalis* is a major symbiont in the olive fruit fly *Bactrocera oleae.* Appl. Environ. Microbiol. 75, 3281–3288.

Manousis, T. and Ellar, D.J., 1988. Dacus *oleae* microbial symbionts. Microbiological Sciences. 5, 149-152.

Mazzon, L., Piscedda A., Simonato M., Martinez-Sanudo I., Squartini A., Girolami V., 2008. Presence of specific symbiotic bacteria in flies of the subfamily Tephritinae

(Diptera Tephritidae) and their phylogenetic relationships: proposal of '*Candidatus* Stammerula Tephritidis'. Int. J. Syst. Evol. Microbiol. 58, 1277-1287.

Miller, M.P., 1997. RxC. A program for the analysis of contingency tables. Northern Arizona University, Flagstaff, AZ.

Montiel Bueno, A. and Jones, O., 2002. Alternative methods for controlling the olive fly, *Bactrocera oleae,* involving semiochemicals. IOBC Wprs. Bull. 25, 1-11.

Nardi, F., Carapelli A., Dallai R., Roderick G.K., Frati F., 2005. Population structure and colonization history of the olive fly, *Bactrocera oleae* (Diptera, Tephritidae). Mol. Ecol. 14, 2729–2738.

Nardi, F., Carapelli A., Boore J.L., Roderick G.K., Dallai R., Frati F., 2010. Domestication of olive fly through a multi-regional host shift to cultivated olives: Comparative dating using complete mitochondrial genomes. Mol. Phylogenet. Evol. 57, 678-686.

Neuenschwander, P., Michelakis, S., 1981. Olive fruit drop caused by *Dacus oleae* (Gmel) (Dipt., Tephritidae). J. Appl. Entomol. 91, 193-205

Osborn, A.M., Moore, E.R., Timmis, K.N., 2000. An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. Environ. Microb. 2, 39-50.

Palmano, S., Firrao, G., Locci, R., 2000. Sequence analysis of domains III and IV of the 23S rRNA gene of verticillate streptomycetes. Int. J. Syst. Evol. Microbiol. 50, 1187-1191.

Petri, L., 1909. Ricerche Sopra i Batteri Intestinali della Mosca Olearia. Memorie della Regia Stazione di Patologia Vegetale di Roma, Roma.

Posada, D., Crandall, K.A., 2001. Selecting the best-fit model of nucleotide substitution. System. Biol. 50, 580-601.

Poulin, R. and Morand, S., 2004. Parasite Biodiversity. Smithsonian Books.

Rice, R.E., 1999. Olive fruit fly, *Bactrocera* (Dacus) *oleae.* UC Plant Prot. Q. 9, 2.

Rice, R., Phillips, P., Stewart-Leslie, J. , Sibbett, G., 2003. Olive fruit fly populations measured in Central and Southern California. Calif. Agr. 57, 122-127.

Sacchetti, P., Granchietti, A., Landini, S., Viti, C., Giovannetti, L., Belcari, A., 2008. Relationships between the olive fly and bacteria. J. Appl. Entomol. 132, 682–689.

Savio, C., Mazzon, L., Martinez-Sañudo, I., Simonato, M., Squartini, A., Girolami, V., 2012. Evidence of two lineages of the symbiont 'Candidatus Erwinia dacicola'in Italian populations of *Bactrocera oleae* (Rossi) based on 16S rRNA gene sequences. Int. J. System. Evol. Microbiol. 62, 179-187.

Segura, M.D., Callejas, C., Ochando, M.D., 2008. *Bactrocera oleae*: a single large population in northern mediterranean basin. J. Appl. Entomol. 132, 706-713.

Silva, F.J., Moret, A., Neef, A., Belda, E., 2008. Bacterial microbiota associated with a *Bactrocera oleae* population from Eastern Spain. In First meeting of TEAM, Palma de Mallorca.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.

Simon, C., Buckley, T.R., Frati, F., Stewart, J.B., Beckenbach, A.T., 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 37, 545–79.

Tamura, K.J., Nei, D.M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596-1599.

Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics. 132, 619–633.

Tzanakakis, M.E., 2006. Insects and mites feeding on olive: distribution, importance, habits, seasonal development and dormancy. Leiden: Brill Acad. Publ. 182 pp.

Van Asch, B., Castro, I.P., Rei, F. , Costa, L., 2012. Mitochondrial haplotypes reveal olive fly (*Bactrocera oleae*) population substructure in the Mediterranean. Genetica. 140, 181-187.

Weisburg, W.G., Barns, S.M., Pelletier, D.A., Lane, D.J., 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. 173, 697–703.

Zygouridis, N.E., Augustinos, A.A., Zalom, F. G., Mathiopoulos, K.D., 2009. Analysis of olive fly invasion in California based on microsatellite markers. Heredity. 102, 402-412.

Zohary, D., 1993. The wild genetic resources of the cultivated olive.

Conclusion

Adaptive radiations are among the most spectacular processes in organismal evolution and this is amplified for island taxa, which display increased speciation rates and elevated levels of morphological diversity (Schluter, 2000; Jordan et al., 2003). This evolutionary process is exemplified by endemic species of the Hawaiian Islands, which provide an ideal setting for studies of evolution and speciation (Roy et al., 2013; Roderick and Gillespie, 1998). In such a scenario, this thesis investigates the bacterial symbioses that occur in tephridis endemic to this Archipelago. Our results support the monophyly of all the endemic Hawaiian tephritids, as reported in Brown et al. (2006), suggesting that adaptive radiation of this group likely occurred within the Islands from a single colonizing ancestor.

The presence of obligate symbioses has been identified in all endemic Hawaiian Tephritinae tested, as previously recovered in some European species within this subfamily (Mazzon et al., 2008). This has allowed us to test the proposed model of speciation in a relatively simple and recent radiation. The detected symbionts are included in the monophyletic clade, described by Mazzon et al. (2008), represented by *Ca.* Stammerula spp. and further grouped within *Stammerula* subclade of non-Hawaiian (Paleartic) *Trupanea* spp. On the basis of these results, the designation of a candidate species '*Candidatus* Stammerula trupaneae' has been proposed to include all Hawaiian symbionts of the *Trupanea* and *Phaeogramma* genera.

Symbiont evolution is greatly influenced by host evolution, and close congruence of host and symbiont phylogenies can indicate cospeciation (Fahrenholz, 1913). Studying the possibility of a coevolution between Hawaiian tephritids and their symbiotic bacteria, cophylogenetic analyses suggest a significant fit between the insect and the symbiont phylogenies, even if not perfect. The lack of strict congruence of this host-symbiont association could be the result of a mix of coevolutionary independent events such as host switching, duplications and sorting events during the biological cycle of the fly. Indeed as described by Girolami (1973) and Mazzon et al. (2008), the extracellular condition of these symbionts in larvae, adults and eggs are important triggers for contacts with the

outer environment. Nevertheless, host-symbiont co-speciation, even if not strict, suggests that vertical transmission appears to be the primary force shaping the topological congruence between insect and bacterial phylogenies. Moreover, the conservatism of plant host lineage and plant host tissue evolution, reconstructed on the phylogeny of the symbiont, also suggests tight fly-symbiont coevolution, as these characters are conservative over the fly lineage.

This approach enables us to better understand the essential mechanism of symbiosis associated with this group of tephritid flies and could represent an useful tool for studying the mechanism responsible for Hawaiian tephritid flies radiation. Future studies, including a larger sample host size, in particular multi-island and oligophagous species, could increase our knowledge of the host-symbiont coevolutionary events.

The study of the interactions between the genetic variability of the bacterial symbiont and its insect host offers an important opportunity for studying the evolution of the fly, as showed in this thesis for the olive fly symbiont, *Candidatus* Erwinia dacicola and its host, *Bactrocera oleae.* Our results reveal the presence of three symbiont haplotypes with a significant phylogeographic distribution related to the territory. Considering the Mediterranean populations, only two lineages have been found and a significative East-West genetic differentiation has been shown. In the South African population only one of these two bacterial haplotypes was recovered while the other was identified in Californian olive flies. In addition, the Pakistani olive fly population harbored a third haplotype, with mixed features between the other two. In contrast, according to some authors (Augustinos et al., 2005; Nardi et al., 2005, 2010; Zygouridis et al., 2009), the olive fly is characterized by an extensive gene flow among the Mediterranean populations.

We think that the combined analyses of host and symbiont genetic variability enable us to provide new interesting insight into the debated colonization routes of the olive fly and to better understand the long host-symbiont coevolutionary life history. For these purpose a further study, in particular focused on Central/Eastern African populations, where the species is supposed to be originated, is necessary to verify this hypothesis. Moreover, this approach could be a model for many studies

of the evolutionary insect-symbiont interdependency and could be extended, for example, to other arthropod-symbiont interactions having similar pattern of symbiotic relationship.

References

Augustinos, A.A., Mamuris, Z., Stratikopoulos, E.E., D'Amelio, S., Zacharopoulou, A., Mathiopoulos, K.D., 2005. Microsatellite analysis of olive fly populations in the Mediterranean indicates a westward expansion of the species. Genetica. 125, 231-241.

Brown, J.M., Todd-Thompson M., McCord A., O'Brien A., O'Fallon B., 2006. phylogeny, host association, and wing pattern variation in the endemic Hawaiian Fruit Flies (Diptera, Tephritidae).Instr. Biodiv. 7,1-16.

Jordan, S., Simon C., Polhemus D., 2003. Molecular systematics and adaptive radiation of Hawaii's endemic Damselfly genus *Megalagrion* (Odonata: Coenagrionidae).Syst. Biol.52, 89-109.

Girolami, V., 1973. Reperti morfo-istologici sulle batteriosimbiosi del Dacus oleae Gmelin e di altri ditteri tripetidi, in natura e negli allevamenti su substrati artificiali. Redia 54, 269–294.

Mazzon, L., Piscedda A., Simonato M., Martinez-Sanudo I., Squartini A., Girolami V., 2008. Presence of specific symbiotic bacteria in flies of the subfamily Tephritinae (Diptera Tephritidae) and their phylogenetic relationships: proposal of '*Candidatus* Stammerula Tephritidis'.Int. J. Syst. Evol. Microbiol.58, 1277-1287.

Nardi, F., Carapelli A., Dallai R., Roderick G.K., Frati F., 2005. Population structure and colonization history of the olive fly, *Bactrocera oleae* (Diptera, Tephritidae). Mol. Ecol. 14, 2729–2738.

Nardi, F., Carapelli A., Boore J.L., Roderick G.K., Dallai R., Frati F., 2010. Domestication of olive fly through a multi-regional host shift to cultivated olives: Comparative dating using complete mitochondrial genomes. Mol. Phylogenet. Evol. 57, 678-686.

Roderick, G.K., Gillespie R.G., 1998. Speciation and phylogeography of Hawaiian terrestrial arthropods.Mol. Ecol. 7, 519-531.

Roy, T., Chang, T.H., Lan, T., Lindqvist, C., 2013.Phylogeny and Biogeography of new world Stachydeae (Lamiaceae) with emphasis on the origin and diversification of Hawaiian and south american taxa.Mol. Phylogenet. Evol. 69,

218-238.

Schluter, D., 2000.The Ecology of Adaptive Radiation. Oxford University Press. Oxford, U.K.

Zygouridis, N.E., Augustinos, A.A., Zalom, F. G., Mathiopoulos, K.D., 2009. Analysis of olive fly invasion in California based on microsatellite markers. Heredity. 102, 402-412.

Acknowledgements

I am grateful to a very large community of people that have made this research possible in many ways. In particular, I would like to thank my supervisors Prof. Vincenzo Girolami and Dott. Luca Mazzon who assisted and supported me during these three years of experience.

A special thanks go to my research colleague, mate of office and friend, Isabel Martinez-Sañudo, that has supported me both academically and personally and all the community of friends, colleagues and professors I had the fortune to met at the University of Padova. I'm grateful to their assistance on my work and, most importantly, I appreciate all of their friendship over the years. In particular, I wish to thanks Mauro Simonato for his patient guidance and invaluable help with various aspects of genetic data analysis.

I'm extremely grateful to Dott. Alberto Bressan that provided me the opportunity to visit his lab for five mounts at the Department of Plant and Environmental Protection Sciences, College of Tropical Agriculture and Human Resources**,** University of Hawaii at Manoa, Honolulu, HI, USA, and discovered that special place. I wish also to thank colleagues and friends I worked with during this experience, especially, Shizu Watanabe, Classon Higasci and Massimo Faccoli. I am grateful to Prof. Russell Messing for introducing me to the field sites in the island of Kauai and providing his valuable assistance. Prof. Jonathan Brown was enormously important for his collaboration and assistance with sample collection, data analysis and paper review.

My family has always provided loving support throughout this process; especially I would like to thank my parents for their constant belief in me. Finally, I thank my boyfriend for his patience and important support and all of my friends for their great encouragements.