UNIVERSITA' DEGLI STUDI DI PADOVA

SCUOLA DI DOTTORATO DI RICERCA IN

SCIENZE DELLE PRODUZIONI VEGETALI

INDIRIZZO PROTEZIONE DELLE COLTURE - CICLO XXIII

Dipartimento di Agronomia Ambientale e Produzioni Vegetali

Symbiotic and associated bacteria in Tephritid flies

Direttore della Scuola : Ch.mo Prof. Andrea Battisti

Supervisore : Ch.mo Prof. Vincenzo Girolami

Dottoranda : Claudia Savio

DATA CONSEGNA TESI

31 gennaio 2011

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.

January 31st, 2011 Claudia Savio

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.

31 Gennaio 2011 Claudia Savio

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

Dichiarazione

Questo lavoro di tesi è stato svolto grazie alla borsa di studio triennale finanziata da Veneto Agricoltura nell'ambito della convenzione con l'Università degli Studi di Padova.
Dall'articolo 5 della Convenzione: Veneto Agricoltura potrà utilizzare, esclusivamente ai fini della propria attività di ricerca e sperimentazione, i risultati degli studi e delle ricerche svolte dai dottorandi. Potrà altresì citare nelle proprie pubblicazioni il contenuto delle tesi finali presentate dai dottorandi al termine di ciascun ciclo di dottorato, con l'obbligo di espressamente menzionare il titolo della tesi, l'autore, il tutore e la Scuola di dottorato.

Citazione suggerita:
Savio C., 2011. Symbiotic and associated bacteria in Tephritid flies. Supervisor: Prof. Girolami V. Scuola di Dottorato di Ricerca in Scienze delle Produzioni Vegetali.

Table of contents

Riassunto

I Tefritidi, noti anche come “fruit fly”, rappresentano una vasta famiglia di Ditteri comprendenti specie dannose per l’agricoltura quali la mosca dell’olivo (*Bactrocera oleae*), la mosca del ciliegio (*Rhagoletis cerasi*), la mosca del noce (*R. completa*) e la mosca mediterranea della frutta (*Ceratitis capitata*).

L’importanza delle associazioni batteriche nella famiglia dei Tefritidi è nota sin da quando Petri, all’inizio del secolo scorso, riportò la presenza di un battere simbionte, in seguito designato “*Candidatus* Erwinia dacicola”, in un diverticolo del capo della mosca dell’olivo chiamato bulbo esofageo. I successivi studi hanno evidenziato, sia con metodi tradizionali, sia con un approccio di tipo molecolare, lo stretto legame esistente tra batteri e tefritidi, siano essi simbionti ereditari e coevoluti e non coltivabili o semplici batteri associati.

Nel presente lavoro sono stati analizzati in dettaglio alcuni aspetti delle relazioni batteriche in alcune specie di tefritidi, quali la variabilità genetica dei simbionti, la composizione della microflora batterica e la presenza di trasmissione attraverso i diversi stadi di sviluppo dell’ospite. La tesi si articola in cinque capitoli.

Il primo lavoro analizza la variabilità genetica nel battere simbionte di *B. oleae* “*Ca*. Erwinia dacicola” in diverse popolazioni italiane della mosca dell’olivo, usando il gene ribosomale 16S come marcatore. Lo studio ha evidenziato la presenza di soli due aplotipi del simbionte, evidenziando anche che la loro presenza contemporanea all’interno di uno stesso ospite sembra in base a tutti i reperti non essere non essere probabile. La distribuzione di queste due linee batteriche nelle popolazioni di *B. oleae* sembra inoltre non essere casuale, poiché le popolazioni delle due maggiori isole italiane (Sardegna e Sicilia) ospitano uno o l’altro dei due aplotipi. Al contrario, le popolazioni della penisola ospitano, in proporzioni significativamente diverse, entrambi gli aplotipi del simbionte. Non è emersa una correlazione tra gli aplotipi di “*Ca*. E. dacicola” e gli aplotipi mitocondriali del loro ospite. Tale risultato potrebbe essere spiegato ammettendo l’esistenza, oltre alla prevalente trasmissione verticale, di accidentali passaggi orizzontali del simbionte.

Nel secondo lavoro l’indagine è stata estesa a un areale più ampio circummediterraneo della mosca dell’olivo. I due aplotipi di “*Ca* Erwinia dacicola” rinvenuti in Italia sono stati riscontrati con frequenze diverse anche in Africa.

Inaspettatamente gli aplotipi del simbionte risultano essere più correlati al territorio di quanto non lo siano i numerosi aplotipi mitocondriali dell'ospite.

Nel terzo lavoro è stata indagata la composizione della microflora di *R. completa* e *R. cerasi* prelevate in natura, analizzandone i diversi stadi di sviluppo sia con approccio tradizionale coltura-dipendente che con approccio molecolare coltura-indipendente. Dal lavoro è emerso che le entità batteriche predominanti presenti nel bulbo esofageo appartengono alla famiglia delle Enterobacteriaceae. I risultati ottenuti evidenziano un meccanismo di trasmissione dei batteri diverso da quello evidenziato per i simbionti della mosca dell'olivo e della sottofamiglia Tephritinae.

Nel quarto lavoro è stata studiata l'abilità di ceppi di *Klebsiella* isolati originariamente dal bulbo esofageo di *R. completa* e *C. capitata* di incorporare il gene per l'espressione di una proteina fluorescente (GFP) e quindi la capacità del battere cosi modificato di ri-colonizzare l'ospite originario. Questa tecnica non distruttiva ha consentito il monitoraggio del destino dei batteri nel corso degli stadi di sviluppo dell'insetto. I batteri modificati sono stai ingeriti con successo dalle mosche del noce e ne hanno colonizzato in modo stabile l'intestino medio allo stadio di larva e quindi nella pupa. Si tratta del primo caso in cui un battere tipico della microflora di un tefritide è stato ingegnerizzato con successo e quindi introdotto nell'ospite nativo. La tecnica utilizzata in questo studio potrebbe costituire un valido strumento per espandere questo tipo di ricerca anche al controllo biologico di altre specie dannose in agricoltura.

L'ultimo lavoro fa parte parzialmente della tesi di dottorato della Dott.ssa Isabel Martinez-Sañudo, per il quale ho contribuito nella parte sperimentale. L'obbiettivo principale di questo studio è stato quello di indagare le relazioni filogenetiche tra le mosche della sottofamiglia delle Tefritine e i loro batteri simbionti. Alcune specie di questa sottofamiglia sono infatti note per ospitare un simbionte specifico ereditario e non coltivabile (''*Candidatus* Stammerula spp.") nell'intestino medio. Tali batteri simbionti sono presenti solo in due delle cinque tribù della sottofamiglia studiate. L'analisi della cofilogenesi ha rivelato la presenza di una congruenza, seppure imperfetta, tra ospiti e simbionti.

Summary

The Tephritidae family, commonly known as “fruit flies”, is a large Dipteran family. It includes many notorious agricultural pests, as the olive fly (*Bactrocera oleae*), the cherry fruit fly (*Rhagoletis cerasi*) and the walnut husk fly (*R. completa*).

The importance of bacteria in the life history of fruit flies is well-known. In the beginning of last century Petri (1909) was the first to report the presence of symbiotic bacteria within the olive fruit fly (*Bactrocera oleae*); recently it was designated as “*Candidatus* Erwinia dacicola”. In Tephritids flies, the bacteria are housed in the midgut and in a specialized intestinal *diverticulum*, located in the fly head, called oesophageal bulb.

In this thesis, some aspects of the relationship between the above mentioned Tephritid flies and bacteria has been investigated, such as the microflora composition, the symbiont genetic variability and the bacterial transmission.

The thesis is composed of five studies.

The first study analyzes in details the genetic variability of *Ca. E. dacicola* in various Italian olive fly populations, studying the 16S rRNA gene. The presence of only two symbiont lineages, not coexisting in the same fly individual, was clearly noticed. Interestingly, the olive fly populations of the two main Italian islands, Sicily and Sardinia, are exclusively represented by one of the two lineages, which could suggest a non-random distribution. On the other hand, the peninsular populations show both bacterial haplotypes, in different proportions. No significant correlation was found between the two symbiont haplotypes and the observed host fly haplotypes, providing evidences for a mixed model of vertical and horizontal transmission of the symbiont during the fly life cycle.

The second study extends the previous study to a wider range. The presence of only two symbiont haplotypes was still confirmed for the Mediterranean and African populations. Surprisingly the symbiont haplotypes seem to be more related to the territory than the numerous host haplotypes.

The third study deals with the identification of the microflora composition of *R. completa* and *R. cerasi*. All the life stages of the cherry fruit fly and the adult stage in *R. completa* have been taken into account, using both culture dependent and independent methods. Bacteria detected within the oesophageal bulbs of both species are affiliated to

Enterobacteriaceae family. The results on the bacterial transmission show a different mechanisms respect to the olive fly and the subfamily Tephritinae symbionts.

In the fourth study, the *Klebsiella spp*. strains isolated from the oesophageal bulb of *R. completa* and *Ceratitis capitata* were examined for their ability to incorporate the gene encoding green fluorescent protein (GFP). These bacteria were successfully labelled by conjugation with the *gfp* gene and the *gfp* gene was stably maintained in the transgenic strains. Moreover, the colonizing ability of *gfp*-tagged bacteria in the original host was tested. Here a non-invasive technique to monitor the bacterial fate during the fly life stages was used. *Gfp*-tagged bacteria were successfully ingested by walnut husk flies where they established a stable population in the fly gut over time and throughout developmental stages. This is the first report in Tephritid flies of native engineered bacteria re-introduced in its original host and the shuttle system used in this study could be a useful tool to expand and strengthen the possibility of biological control of the insect pest.

The last study is part of Isabel Martinez-Sañudo PhD thesis, for which I contributed to experimental works. The main goal of this study was to analyse the phylogenetic relationships between flies of the Tephritinae subfamily and their symbiotic bacteria. Some species of this subfamily are in effect known to host specific non-culturable symbiont bacteria (‘‘*Candidatus* Stammerula spp.”) in their midgut. The cophylogenetic analysis reveals the presence of congruence, even if imperfect, between hosts and symbionts. This non-strict congruence is probably due to events such as losses, duplications and hosts switching, which are likely to arise during the biological cycle of the fly in consideration of the extracellular status of these symbionts.

Chapter 1

Introduction

Tephritidae

Biology and distribution

The Tephritidae family, commonly known as fruit flies, is one of the largest families of Diptera, with about 4200 described species in almost 500 genera (White, 2006). Fruit flies are considered among the most attractive and biologically interesting Diptera, having patterned wings and often brightly coloured and/or patterned bodies. 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). Many species, mostly the carpophagous ones, are a notorious group of agricultural pests (*Anastrepha*, *Bactrocera*, *Ceratitis*, *Dacus* and *Rhagoletis* genera).

Tephritid larvae develop in fruit, leaf-mines or within the flower-heads of *Asteraceae*; all of these three life systems include pests, and many species associated with the *capitula* of composites are potential weed biocontrol agents (White, 1988).

Three subfamilies of Tephritidae are recognized, the Dacinae, Trypetinae and Tephritinae (Foote, 1993), but the classification of subfamilies is currently under revision. White (1988) subdivided Tephritidae into four subfamilies of Palaearctic species: subfamily Dacinae, subfamily Myopitinae, subfamily Trypetinae and subfamily Tephritinae (Fig.1.1). Dacini are mainly concentrated in the Afro-tropical region, Southeast Asia and north-eastern Australia (Drew & Hancock, 2000). Members of the subfamily Trypetini are more numerous in the Palaearctic and Oriental regions but some genera occur in the Nearctic and Neotropical (Norrbom *et al.*, 1999) regions. *Rhagoletis completa* and *R. pomonella* for example are native to North America. Freidberg (1984) reports that the subfamily Tephritinae is largely restricted to the Holarctic temperate region, higher altitude areas of the Afrotropical and Neotropical region.

Regarding feeding behaviour the Dacinae and Trypetinae groups use the fleshy fruit of host plants from a wide variety of families as larval food sources. The Tephritinae larvae feed on the vegetative parts of host plants and flower heads, and many form galls. The tephritines, with some exceptions, feed on plants in the Asteraceae family. As can be argued, host range varies considerably, often among closely related species (Norrbom & Kim, 1988; Goeden, 1994). However some of them are strictly oligophagous or

monophagous, such as *Bactrocera oleae* (Rossi), which breeds only on olives, but other species are remarkably polyphagous, such as *Ceratitis capitata* (Widemann) which has been reported on more than 300 hosts (Liquido *et al*., 1991).

This information can be summarized following the Zwölfer subdivision (1983) based on resources exploitation strategies:

a) generalist frugivorous species: larvae feed and develop in the pulp of fleshy fruits: e.g. *Anastrepha*, *Ceratitis* and *Dacus* genera;
b) specialized frugivorous species: e.g. trypetines;
c) exploiters of vegetative structures and inflorescences: e.g. *Tephritis spp.*, leaf-mining Myopitinae.

Economic importance

About 70 species of fruit flies are considered important agricultural pests, and many others are minor or potential pests (White & Elson-Harris, 1992). The reason for heavy losses in fruit and vegetable crops (including citrus, mango, apples and many others) is due to the phytophagy of their larvae. The most injurious species belong to the *Anastrepha*, *Ceratitis*, *Bactrocera*, *Dacus* and *Rhagoletis* genera whose hosts belong to a wide variety of plant families, and include many major commercial crops (Norrbom, 1999).

- The genus *Bactrocera* is the most economically significant genus, with about 40 species considered to be important pests (White & Elson-Harris, 1992). Among them the olive fruit fly (*B. oleae*), oriental fruit fly (*B. dorsalis*), melon fly (*B. cucurbitae*), Queensland fruit fly (*B. tryoni*), and peach fruit fly (*B. zonata*).
- *Anastrepha* is the most economically important genus in the New World tropics. The Mexican fruit fly (*A. ludens*), *A. obliqua*, and South American fruit fly (*A. fraterculus* complex) are recorded as the most injurious among the 15 worst pest species by White & Elson-Harris (1992).
- The genus *Rhagoletis* includes species in both the Holarctic and Neotropical regions. The most serious are the European and eastern cherry fruit flies (*R. cerasi* and *cingulata*), walnut husk fly (*R. completa*), apple maggot fly (*R. pomonella*) and blueberry maggot fly (*R. mendax*).

- *Ceratitis* genus is restricted to African countries, with the exception of *C. capitata* which is one of the most polyphagous and widespread species of Tephritidae (Liquido *et al.*, 1991) and is by far the most notorious pest species in the genus.
- *Dacus* genus is also Afrotropical, with a few species occurring in some parts of the Palaeotropics and subtropics. The most dangerous species listed by White & Elson-Harris (1992) mainly attack Cucurbitaceae, like the pumpkin fly (*D. bivittatus*) and lesser pumpkin fly (*D. ciliatus*).

Even if Tephritidae are commonly thought of as pests, some species are beneficial: members of the genera *Proceciodochares*, *Urophora* and *Tephritis* are successfully used for biological control of weeds (Zwölfer, 1983). *U. quadrifasciata* for example has been introduced to Canada from Europe to control *Centaurea diffusa* and *C. biebersteineii* and *P. alani* has controlled *Ageratina riparia* (Regel) in some areas of Hawaii (Norrbom, 1999).

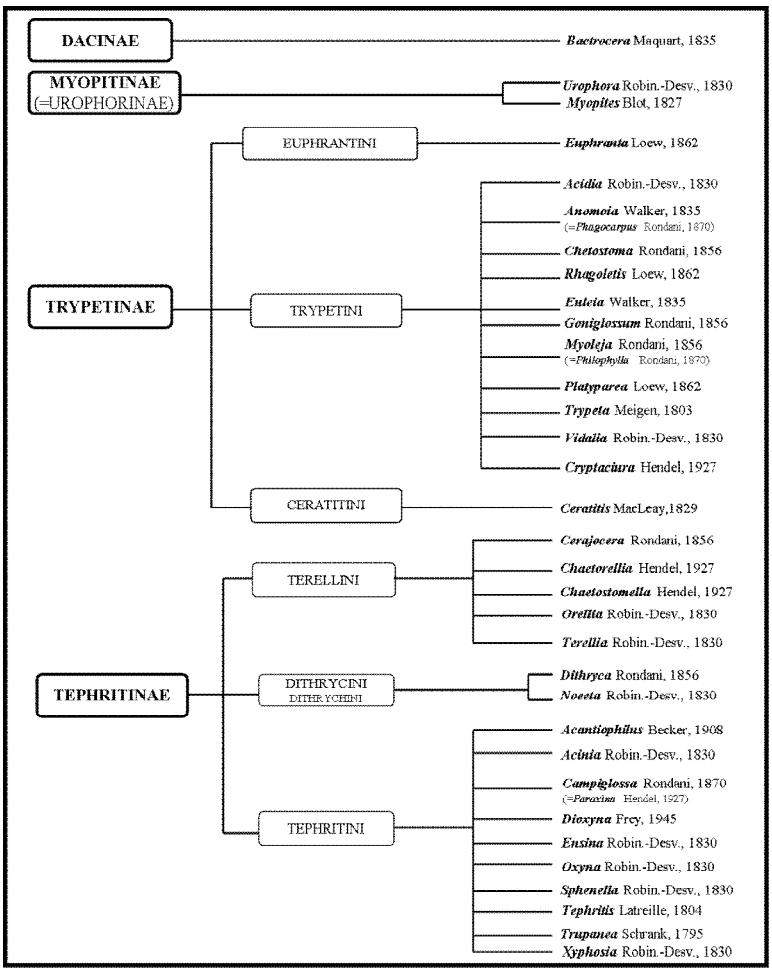

Fig. 1.1 - Systematic subdivision of Palaearctic species of the Tephritidae family (White, 1988).

***Rhagoletis cerasi* (Linnaeus)**

Biology and distribution

Rhagoletis cerasi, the European cherry fruit fly, is one of the most well-known fruit fly pests in the Mediterranean region. It occurs in almost the entire cherry producing area of Europe, where it is an important pest in some areas, while in others it causes only negligible damage. Its southern limits overlap with the Mediterranean peninsulas and Balearic, Sardinia and Sicily islands (Fimiani, 1989).

R. cerasi is an oligophagous fly, which can attack fruits of both the *Prunus* (*P. avium*, *P. cerasus*, *P. mahaleb*) and *Lonicera* (*L. xylosteum*, *L. coerulea*) genera. The cherry fruit fly life cycle includes, like all the tephritids, the following stages: egg, three larval instars, pupa (formed inside the hardened third stage larval cuticle, or *puparium*) and adult.

R. cerasi is univoltine and the pupae diapause in the soil in the immediate vicinity of the host. The species is well-equipped to survive extreme environmental conditions and some pupae may overwinter for 2 or even 3 years, thus assuring persistence of the population during years of fruit scarcity or absence. Adult emergence is synchronized with the ripening of the host fruits (from the end of May to early July) and is brought about by a diapause system closely adjusted to climate and photoperiod (Boller & Prokopy, 1976).

The first eggs are laid 10 to 15 days after flying starts, in warm weather (Fig. 1.2a). Each female lays 50 to 80 eggs that are inserted singly under the epidermis of the fruit when it is starting to turn red. After oviposition, the female deposits a non-volatile, very persistent, host marking pheromone on the fruit surface, which inhibits further egg-laying in already oviposited fruits.

Larvae are frugivorous and feed on the pulp surrounding the kernel (Fig. 1.2b). After approximately 25 days from ovipostition the larvae leave the fruit and pupate at a shallow depth in the soil.

Fruit damage may reach 100% of fruit production, so control measures are required. These usually involve using insecticides either in the form of proteinaceous bait sprays or as a cover spray applied three weeks before harvest (Katsoyannos *et al.*, 2000). Recently the neonicotinoid thiamethoxam has been registered.

***Rhagoletis completa* Cresson**

Biology and distribution

Rhagoletis completa, the walnut husk fly, native to the Midwestern United States (Fig. 1.3a), is a key pest of the common walnut in the U.S.A. (Olson & Buchner, 2002). Its presence in Europe was reported for the first time in 1991 in southern Switzerland by Merz, who identified some specimens. The first record of larval attack on *Juglans regia* in Europe was recorded in various regions of northern and central Italy (Duso, 1991; Trematerra *et al.*, 1995), and in Slovenia (Seljak, 1999). In Europe, as in the United States, the fly has one generation per year. Like *R. cerasi*, *R. completa* is a monophagous pest, attacking fruits of *Juglans* spp. (*J. regia*, *J. nigra* and their hybrids) but some attacks on peaches (*Prunus persica*) have been recorded in California (Bush, 1966).

Fly emergence lasts from early July to the second half of August. Adults feed on honeydew, yeasts and leaf exudates (Kasana & AliNiazee, 1995). Mating takes place 6-8 days after emergence followed by oviposition 1-2 weeks later, which occurs from mid-July to early September on *J. regia*, with peaks during mid-August (Duso & Dal Lago, 2006). In most cases, a single batch of eggs is laid on the fruit in a cavity created by the ovipositor. In its lifetime a female can lay up to 400 eggs (Boyce, 1934). Eggs are laid below the skin of the walnut, hatch after 3-7 days and feed for 2-5 weeks. Mature larvae leave the fruit to pupate in the soil, this being the normal overwintering stage (Fig. 1.3b). Larval attacks on the walnut pulp cause shell staining and darkened kernels.

***Bactrocera oleae* (Rossi)**

Biology and distribution

The olive fruit fly, *Bactrocera oleae* (Rossi) (Diptera: Tephritidae), was recorded attacking olives in Roman times and has long been a major pest in the Mediterranean basin. The Mediterranean basin, northern, eastern and southern Africa, Canary Islands, India, western Asia, and apparently wherever olives (the genus *Olea*) grow in the Eastern Hemisphere are the regions where *B. oleae* can be found. In 1998 the fly was recorded for the first time in the United States, in California (Rice, 1999).

In order to clarify the olive fly origins, their population genetics have been investigated since the 1980s. Initial molecular studies of the olive fruit fly compared laboratory and

natural olive fruit fly populations (Tsakas & Zouros, 1980) and later the first attempts began to characterize the genetic diversity in localized geographic regions of the Mediterranean basin (Ochando & Reyes, 2000; Augustinos *et al.*, 2002). More detailed analyses of geographically widespread olive fruit fly populations representing Africa, Europe, the Middle East, and North America used both multilocus microsatellite loci and mitochondrial DNA haplotypes. These analyses pointed out the presence of genetically distinct sub-Saharan African, Mediterranean, and Pakistani populations, the last formerly described as the variety *asiatica* by Silvestri (1916) (Nardi *et al.*, 2005; Nardi *et al.*, 2010). From these studies it seems clear that *B. oleae*'s closest relatives are of African origin, a thesis also supported by White (2006) who reported that the Mediterranean population of *B. oleae* should therefore be regarded as an introduced population, albeit a very ancient one.

B. oleae can be considered an oligophagous species because it only feeds on the fruit of some *Olea* species, including *O. europaea, O. verrucosa,* and *O. chrysophylla* (Fig.1.4a). In addition to cultivated olives, olive fly is known to attack wild olives. Infestation in these hosts could have allowed the fly to spread along the east coast of Africa as far as central South Africa where wild olives occur along with a few plantings of commercial olives (Rice, 2000).

Three to five generations of the fly per year are reported in the Middle East, two to five in different parts of Europe (Daane & Johnson, 2010). In most regions, the olive fruit fly appears to be best adapted to develop in the autumn period, when its larval food (i.e., olive fruit) is in its optimal condition for larval growth (Tzanakakis, 2006).

Eggs are laid under the surface of unripe, ripening and ripe olive fruit so that the neonate larva has direct access to food. Larvae feed and grow as fruit borers in the mesocarp of olives and become mature within a few weeks, depending on the temperature (Girolami, 1979) and then they pupate. Pupal development requires a minimum of 12 days at 24 °C (Girolami, 1979). Unlike other tephritid species, olive flies pupate within the unripe host fruit during warmer months (Fig. 1.4b), but leave the fruit to pupate in the ground or in any protected niche during autumn and winter. Adult flies can live for up to 9 months depending upon food (Girolami, 1979). Indeed, while olive fruit fly larvae are dependent on the presence of *Olea* fruit, adults feed on a variety of organic sources including insect honeydews, plant nectar, plant pollens, and fruit exudates. They may feed on nutrient

sources such as bird droppings, bacteria, and yeasts to meet their nutritional requirements (Tsiropoulos, 1977). It seems that the olive fly can achieve good fecundity without the consumption of vitamins, amino acids, proteins, etc; and that sucrose only, for example honeydew or fruit sugar is sufficient (Marzaro, 2008).

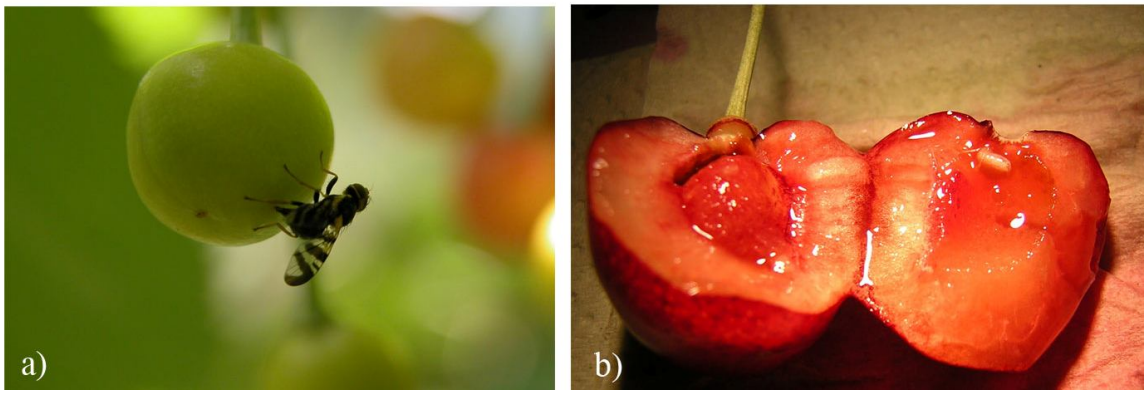

Fig. 1.2 - a) *Rhagoletis cerasi* during oviposition on a green cherry. (Photo L. Mazzon) b) Mature cherry infested by a *R. cerasi* larva. (Photo C. Savio)

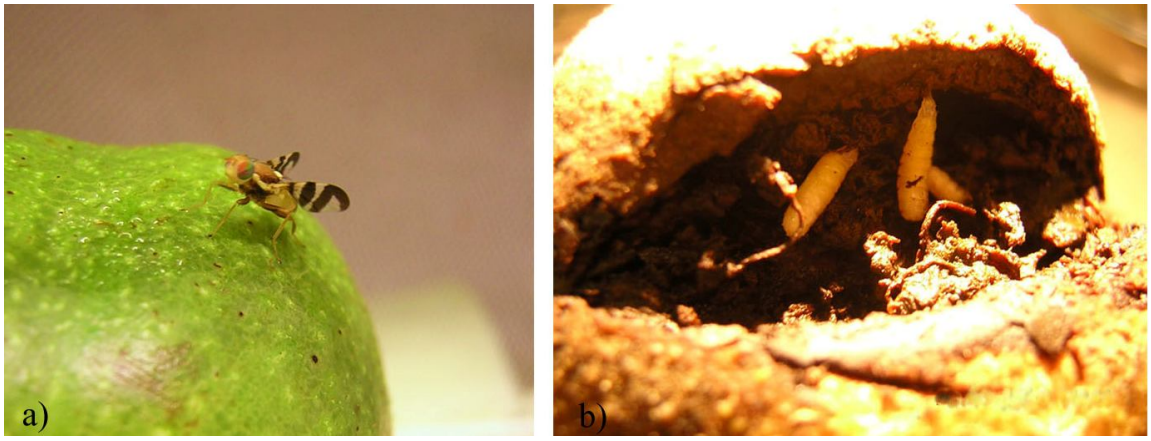

Fig. 1.3 – a) *Rhagoletis completa* on a walnut. b) Larvae of *R. completa* feeding on walnut pulp. (Photos C. Savio)

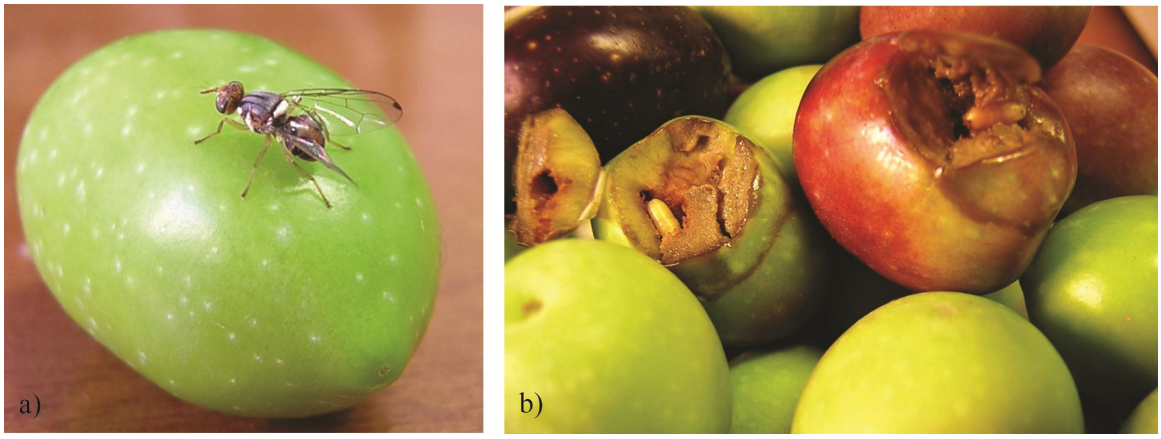

Fig. 1.4 - a) *Bactrocera oleae* on an olive. b) Olives infested by *B. oleae*; pupae and galleries are visible. (Photos C. Savio)

Bacterial symbiosis in insects

The success and the great diversity of insects has created myriad opportunities for bacteria to occupy niches created by insects, in particular in their guts (Dillon & Dillon, 2004).
The gut microbes can be defined as obligate symbionts, contributing to the host fitness, or facultative symbionts whose presence is not deemed essential for host survival (Pontes & Dale, 2006). Some facultative populations can be of a transient nature (i.e. commensals or parasites), others are mutualistic, and are known to play significant roles in their host's biology (Behar *et al.,* 2008). As reported by Zook (1998), a symbiotic relationship of a microbe with its insect host can be defined as the acquisition and maintenance of the microorganism that results in novel structures or metabolism. Such a relationship should be defined after careful consideration because many species of insects are inhabited by diverse communities of microorganisms, including bacteria, Eukaryea and Archaea.

As in vertebrates, gut microbes of most insects are extracellular, either free-living in the lumen or adhering to the gut wall (Ishikawa, 2003). It is assumed that the microbiota most often derives from the surrounding environment such as the phylloplane, or the skin of the animal host, with degree of persistence depending on the host (Dillon & Dillon, 2004). In some species the symbiont is housed within the host cells, like *Wolbachia*, known to infect not only insects but other invertebrates (mites, crustaceans, nematodes) (Werren & O'Neill, 1997) and *Buchnera*, mainly infecting aphids (Munson *et al*., 1991; Baumann *et al*., 1995). An overview on the presence of intra- and extra-cellular symbiotic bacteria is reported in Buchner's renowned treatise (1965).
Although insect bacterial associations are ubiquitous, only a few groups of host families have been well studied in relation to their associations with microbes. The first was *Buchnera* (Baumann *et al*., 1995). Other examples are the associations with cockroaches (Bandi *et al.,* 1994; Lo *et al.,* 2000), stink bugs (Prado & Almeida, 2009) and termites (Bandi *et al.,* 1995; Brune, 1998).

In the past the description of the gut microflora was based only on traditional microbiological methods, by phenotypic characterization of isolates, hence the presence of unculturable bacteria was largely ignored (Dillon & Dillon, 2004). Amann *et al.* (1995) report that about 99% of microbes in the environment cannot be cultivated. Advances in

molecular biology have allowed in-depth studies, changing the understanding of microbial diversity during the last 20 years. The best suited molecular approach involves the use of 16S rRNA genes, the most frequently used marker in bacterial identification and phylogeny, flanked by the use of other techniques such as amplified rDNA restriction analysis (ARDRA) and fluorescent *in-situ* hybridization (FISH). This kind of approach has enabled an in-depth characterization of the microbiota flora of many insects.

Bacteria in the life history of Tephritids

Originally the association of bacteria with fruit flies was investigated for an economic purpose. In fact the presence of certain bacteria may play a role in improving the performance of the SIT (Sterile Insect Technique) males, as they are known to contribute important nutritional benefit in insects (Ben Ami *et al.*, 2010). In the 1970s Girolami & Cavalloro (1972) hypothesized that the lack of symbiotic bacteria in *B. oleae* was the reason for the poor quality of adults reared on artificial substrates. A nutritionally complete diet for SIT males pre- and post-release may enhance the performance of SIT colony insects and thus the study of associated bacteria could be useful in planning pest control strategies.

Early studies on bacteria-tephitids interactions often described any bacteria isolated from life stages as symbionts (Lauzon, 2003), without considering described microorganisms as possible contaminants or not necessarily "living together". Indeed, as stated by Drew & Lloyd (1991), early studies considered fruit flies engaged in "mutually obligatory relationships which often were not demonstrated". Hence, when speaking about bacteria-tephritids interactions, we have to keep in mind that the term "symbiosis" must be intended as "obligatory relationship". This section gives a brief description of some tephritid-bacteria relationships and these are summarized in Tab. 1.1.

Petri was the first researcher working on this topic and in 1909 he described a strong relationship between an extracellular bacterium and the olive fly *Dacus* (*Bactrocera*) *oleae*. As reported in Capuzzo *et al*. (2005) "…the symbiont might be '*Bacterium'* (*Pseudomonas*) *savastanoi*, the causal agent of the olive knot disease, as it could be rescued from larvae; however, he pointed out that, if this were the case, the cultured cells would represent just a minimal fraction of the whole bacterial mass multiplying in the larval blind

sacs and in the adult's oesophageal bulb. By comparing the exiguous c.f.u. counts on nutrient plates with the microscopically visible bacterial volume filling the appropriate organs, he postulated that the bacteria hosted by the olive fly would exist in a state of prevailing non-culturability."

In 1966, Hagen tried to explain the *P. savastanoi* role, suggesting that the microorganisms provide essential amino acids for larval development. In 1929, Stammer reported the presence of a characteristic symbiosis for some species of the Tephritinae subfamily, belonging to the genera *Tephritis*, *Oxyna* spp., *Paroxyna* spp., *Campiglossa* spp., *Trupanea* spp., *Acanthiophilus* spp. and *Sphenella* spp., while no symbiotic bacteria, nor special organs devolved to contain bacteria were found in the remaining species. In adult stages of the above-mentioned genera, symbiotic bacteria are located in the first tract of the midgut in contact with epithelium but, as reported later by Girolami (1983), outside the peritrophic membrane.

In 2008 Mazzon *et al.,* in agreement with Stammer (1929), detected the presence of non-culturable symbiotic bacteria in species of the subfamily Tephritinae. Sequencing of the 16s rRNA from these bacteria indicated that they belong to the family Enterobacteriaceae and a novel candidate organism was proposed for the symbiotic bacteria of the genus *Tephritis*, under the designation of "*Candidatus* Stammerula tephritidis". A cophylogenetic study on the species revealed that (with a few exceptions) hosts and symbionts seem to have coevolved during their long history (Mazzon *et al*., 2010).

During the 1980s studies on bacterial relationships in fruit flies increased considerably (see reviews by Drew & Lloyd, 1987, 1991), taking into account the microflora inhabiting the adult fruit flies guts. These studies were mainly based on traditional culturable methods such as viable plate counts and many of their authors pointed out a frequent interaction between the Enterobacteriaceae family and tephritid flies (Drew & Lloyd, 1987; Girolami, 1986; Kuzina *et al.,* 2001; Lauzon *et al.,* 1998; Marchini *et al.,* 2002; Tsiropoulos, 1976;). The Enterobacteriaceae presence was also confirmed with the advent of biomolecular techniques, evidencing a prevalence of "associated bacteria", mainly belonging to the *Enterbacter*, *Klebsiella* and *Pantoea* genera (Behar *et al*., 2005, 2008; Ben Ami *et al*., 2010). The ease of isolation of these species from both temperate and tropical tephritids suggests that although bacterial associations may not be obligatory, these bacterial genera

likely play an important role in the life history and ecology of these pests (Drew & Lloyd, 1991).

The species most commonly found in fruit flies midguts belong to *Klebsiella* genus, first reported by Rossiter *et al*. (1983) in *R. pomonella*.

Reference	Tephritid species	Source of sample	Symbiotic bacteria	Associated bacteria
Petri, 1909	*Bactrocera oleae*	Adult oesophageal bulb	*Pseudomonas savastanoi* unculturable species	
Capuzzo *et al*., 2005		Larval blind sacs, adult oesophageal bulb, ovipositor	*Candidatus* Erwinia dacicola	
Sacchetti *et al*., 2008		Adult oesophageal bulb, gut, ovipositor	*Candidatus* Erwinia dacicola	*Asaia sp.*
Kounatidis *et al*.,2009		Adult midgut	*Candidatus* Erwinia dacicola	*Acetobacter tropicalis*
Estes *et al*., 2009		Larval and adult midgut, ovipositor	*Candidatus* Erwinia dacicola	*Enterobacter sp.*
Stammer, 1929	Tephritini tribe	Larval and adult midgut	Symbiotic bacteria	
Mazzon *et al*., 2008		Larval and adult midgut	*Candidatus* Stammerula spp.	
Marchini *et al*., 2002	*Ceratitis capitata*	Adult oesophageal bulb		*Enterobacter aglomerans* *Klebsiella oxytoca*
Behar *et al*., 2008		Adult oesophageal bulb		*Enterobacter sp.* *Klebsiella oxytoca sp*
Rossiter *et al*.,1983 Howard *et al*., 1985	*Rhagoletis completa*	Adult oesophageal bulb		*Klebsiella oxytoca* *Klebsiella ozaenae* *Klebsiella pneumoniae*
Tsiropoulos, 1976		Different life stages		*Pseudomonas sp.* *Xantomonas sp.*

Tab 1.1 – Summary on the knowledge status of bacterial relationships on some tephritid species.

Oesophageal bulb and symbiosis

Petri (1909), when studying *B. oleae* symbiosis, described for the first time the presence of a specialized organ in the fly's head, connected to the pharynx, called "pharyngeal vesicle" where the symbionts multiply and then spill out in masses into the intestine. In 1935 a similar organ, called "oesophageal bulb", was described by Dean in *Rhagoletis pomonella* but without any indication of a relationship with bacteria. The existence of a pharyngeal bulb was also considered exclusive to olive fly by Buchner (1965) and denied for *C. capitata* and *Rhagoletis*, with the puzzle of the organ described by Dean. The presence of an oesophageal bulb in all the Tephritidae was only recognized in 1973 by Girolami, who provides a description for all four morpho-histological types (Fig. 1.5 and 1.6).

- The "*Dacus* type", typical of *Bactrocera oleae*, is the largest one and it corresponds to the description of Petri (1909). It is spherical and a neck connects the pharynx, without evident muscular sheath. The symbiotic bacteria multiply next to the elongated basal epithelial cells.

- The "*Ceratitis* type" is present in all the species of the subfamily Trypetinae and Dacinae (with the exception of *B. oleae*). It is smaller than the "Dacus type" and the apex is provided with elongated cells, covered by a muscular sheath. The associated bacteria multiply in the lumen and are easily visible in Trypetinae.

- The "*Ensina* type" is egg-shaped, typical of the subfamily Tephritinae (with the exception of the tribe Terellini). It has a strong and wide muscular sheath, without articulated cell elongations. No bacteria appear within this oesophageal bulb.

- The "*Chaetorellia* type" is characteristic of the tribe Terellini and shows intermediate features between "*Ceratitis*" and "*Ensina*" types. It shows a pharyngeal outward-deflection, whose apex closely resembles the oesophageal bulb of the subfamily Trypetinae.

Bacteria were observed inside the oesophageal bulb but they are not necessarily linked to a hereditary symbiosis. Inside all of the oesophageal bulbs of Dacinae (*B. oleae* included) and Trypetinae, independently of the presence of bacteria, fibrous masses, which spill into the intestine, are produced in the pharyngeal bulb for *Dacus*, *Ceratitis* and *Chaetorellia* types.

These fibrous masses constitute the heart of the bacteria masses (Girolami, 1973) and erroneously considered as “biofilms” reasonably produced by bacteria but never demonstrated. The histological structure using an electron microscope was studied for the first time in *R. pomonella* by Ratner & Stoffolano (1984) and in *B. oleae* by Mazzini & Vita (1981).

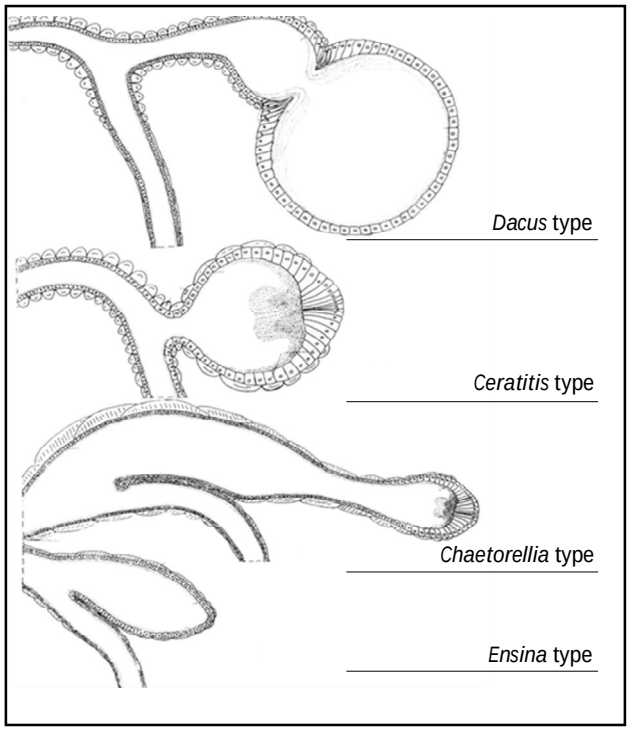

Fig. 1.5 - The four morpho-histological types of oesophageal bulbs in Tephritidae. (Modified from Girolami, 1973)

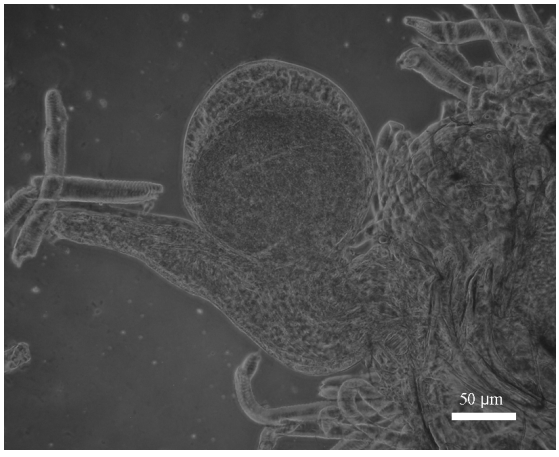

Fig. 1.6 - Oesophageal bulb of *Rhagoletis completa*, with the typical Ceratitis type shape. (Photo C. Savio)

Membranous masses

The oesophageal bulb of the species of the subfamilies Dacinae, Trypetinae and in the tribe Terellini, belonging to Tephritinae subfamily, continuously produces "membranous masses" that are discharged into the midgut regardless of whether bacteria are present or not. In the midgut they can be observed inside the peritrophic membrane (Girolami, 1973). In the Trypetinae subfamily the membranous masses adopt the shape of the oesophageal bulb (Fig. 1.7a). These masses are produced inside the oesophageal bulb, it seems from a continuous delamination of the stomodeal intima. It is known that in *B. oleae* they are produced more than ten times per day and bacteria multiply inside them (Piscedda & Girolami, 2005) (Fig.1.7b). The production of these membranous masses implies an energy cost, so it is probable that the efficient multiplication of free living bacteria represents an advantage for the adults of tephritid flies.

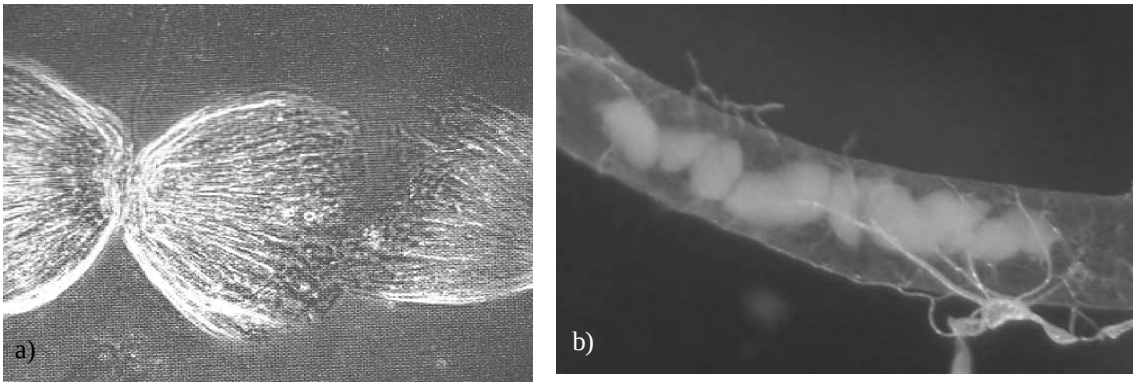

Fig. 1.7 - a) Membranous masses in *Ceratitis capitata* adopting the shape of the internal intima of the oesophageal bulb. (Photo from Girolami, 1973). b) Membranous masses in the midgut of *Bactrocera oleae* (Photo A. Piscedda)

A brief report follows of the current status of knowledge about the bacterial symbiosis/association with the tephritids species that are the subject of this thesis: *C. capitata*, *R. cerasi*, *R. completa* and *B. oleae*.

Ceratitis capitata

The Mediterranean fruit fly, is one of the tephritids whose microflora has recently been under scrutiny. The first accurate study goes back to 2002 when Marchini *et al.,* analysed the oesophageal bulb content using culture-dependent methods. As in other tephritids a large number of Enterobacteriaceae was found (e.g. *Enterbacter* sp., *Citrobacter freundii, Klebsiella oxytoca*), but also members of other bacterial families, e.g. *Pseudomonas* sp. (Behar *et al.*, 2005, 2008, 2009). The bacterial transmission in the medfly has also been investigated and the presence of vertical transmission from larvae to adults was hypothesized in 1986 by Girolami and later by Marchini *et al.* (1991) and Behar *et al*. (2008).

The microorganisms functions is not easy to understand but recent studies suggest that that they occur in nitrogen fixation (Behar *et al*., 2005) and could increase the fitness during various stages of the host's life cycle (Ben-Yosef *et al.*, 2008) and in sterile males (Ben Ami *et al*., 2010).

Rhagoletis cerasi

There are no clear data up to now about *R. cerasi* microflora, except for a citation in Fitt & O'Brien (1985) about the detection of *Pseudomonas* sp. as a symbiont of *R. cerasi*.

In 1989, Blümel & Masalmeh detected the presence of gram-negative bacteria in pupae and adults of the cherry fruit fly but not in the eggs and larvae.

On the other hand the existence of *Wolbachia,* an endosymbiont of arthropods that can cause alterations to host reproduction, is certain and well-studied. In *R. cerasi* it is responsible for cytoplasmic incompatibility (Riegler & Stauffer, 2002; Riegler *et al*., 2004).

Rhagoletis completa

The first detailed study on the microflora associated with *R. completa* dates back to 1976. Tsiropoulos isolated 15 morphologically different bacteria from both rotted walnut pulp and different life stages of the fly. Only *Pseudomonas* sp. and *Xantomonas* sp. were found to be associated with all life stages. Later, Howard *et al.*, in a compared study among seven *Rhagoletis* species in 1985, noticed that no bacterium has entered into an obligate symbiotic relationship with *R. completa,* but *Klebsiella oxytoca* was the most common inhabitant of its oesophageal bulb.

Bactrocera oleae

The associations between the olive fly, *Bactrocera oleae*, and its intestinal digestive-system bacteria has been studied for more than a century. In 1909, Petri was the first, based on microscopic observations, to describe the presence of bacterial masses, hypothesized to be *Pseudomonas savastanoi*, a pathogen causing the olive knot disease. In the adult they multiply inside the oesophageal bulb, a specific organ present in the foregut of the adult olive fly. From this organ the bacteria are released into the oesophagus and thence to the intestine where they are eventually digested (Petri, 1909). In his studies Petri showed the presence of specific contractile perianal glands in the ovipositor of female flies (Fig.1.8), allowing the transmission of symbionts to the offspring; indeed, a bacterial cap-like mass is typically found around the egg's micropyle. Petri supposed that in this way the bacteria could eventually enter the micropyle and stay until they establish within the larval midgut. Here they are located in 4 gastric intestinal caeca (Fig.1.9), or larval blind sacs, within the peritrophic membrane and thus in the intestinal lumen direction. The peritrophic membrane develops from the blind sacs and has a thin lining layer, which can erroneously be identified as cellular membranes (Estes *et al.*, 2009). According to Petri, bacteria remain during the pupal stage (but their location, except the prepupa, is still unknown) and they re-appear in the adult fly multiplying in the oesophageal bulb.

The basis of the symbiotic advantage for the flies has been postulated to be a nutritional effect, both in terms of enhanced dietary protein hydrolysis and the synthesis of required amino acids lacking in the olive pulp (Tsiropoulos, 1980). Studies by Hagen (1966) and Hagen & Tassan (1972), reported that larvae developing from eggs laid by adults deprived of their bacteria are unable to develop in the olives. On the basis of these observations Girolami (1973) pointed out a specific correlation between symbiont losses, reduced fecundity, fertility and survival of the species and pharyngeal bulb largeness.

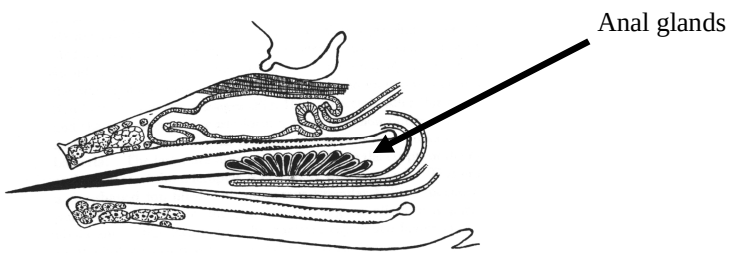

Fig. 1.8 – Longitudinal section of the female olive fly ovipositor (Modified from Petri, 1909)

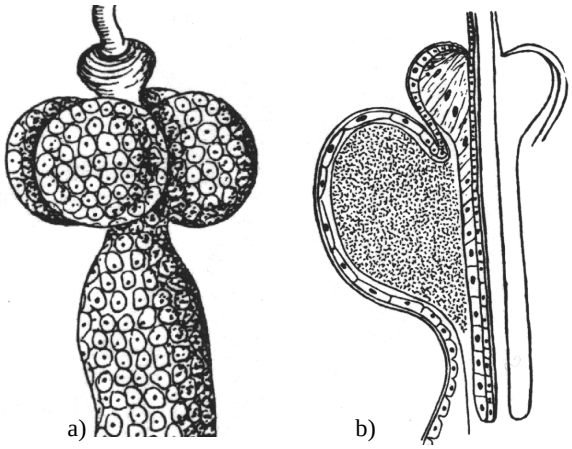

Fig. 1.9 – a) Larval intestinal caeca of *Bactrocera oleae*; b) longitudinal section. (Modified from Petri, 1909)

Thanks to biomolecular techniques, Capuzzo *et al.,* (2005) suggested that the bacterium housed within the oesophageal bulb and the midgut of *B. oleae* is not *P. savastanoi* but displays marked similarity to *Erwinia* species. The name *Candidatus* Erwinia dacicola was thus proposed (Fig. 1.10). Different cultivability tests have been conducted but the bacterium still remains uncultivable (Capuzzo *et al.*,, 2005; Estes *et al.*,, 2009). Other studies have identified temporary gammaproteobacterial and alphaproteobacterial associations acquired during feeding (Belcari *et al.,* 2003; Sacchetti *et al.,* 2008; Kounatidis *et al.,* 2009) in the olive fly. This could happen because flies progressively lose bacteria in artificial media rearing, causing not only lower insect vitality and fertility, but also shrinking of the oesophageal bulb to about one-third of its normal volume. As a consequence this cavity, lacking its typical microflora, becomes much more prone to invasion by other microbial species (Girolami & Cavalloro, 1972; Capuzzo *et al.,* 2005).

Ca. Erwinia dacicola has also been found, other than in Italy where it was first described, within adult *B. oleae* in Greece (Kounatidis *et al.,* 2009), the south-western United States (Estes *et al.,* 2009), Spain (Moret *et al*., unpublished) and in other life stages (Estes *et al.,* 2009).

In conclusion, the high frequency of association of *Ca.* Erwinia dacicola with the olive fly, its presence in different populations and in all life stages, its inability to be cultured on

standard microbiological media, plus its vertical transmission to offspring, suggest that this bacterium is a strictly associated symbiont of olive flies (Estes *et al*., 2009).

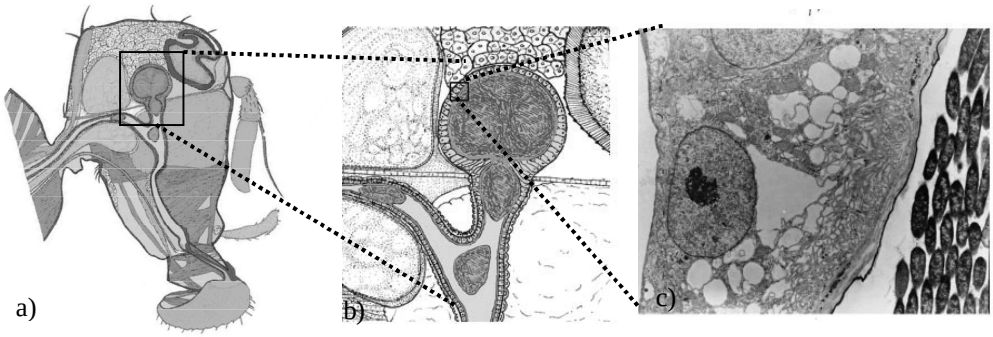

Fig. 1.10 – *Candidatus* Erwinia dacicola location within adult *Bactrocera oleae.* a) Olive fly head in longitudinal section with indication of the oesophageal bulb. b) Oesophageal bulb, lumpy masses of bacteria in transit towards the oesophagus are visible. c) Transmission electron micrograph showing an insect cell (left) from the pharyngeal bulb epithelium and the bacteria stored within (right). (Modified from Capuzzo *et al*., 2005). Image c) is used courtesy of M. Mazzini.

Objectives and content of the thesis

In this thesis I further developed studies on the olive fly symbiont, *Candidatus* Erwinia dacicola, started by Capuzzo *et al.*, 2005, particularly focusing on its genetic variability within Italian olive fly populations (Chapter 2) and within a wider worldwide context (Chapter 3) covering most of the olive fly range. The results obtained in both studies give an interesting insight into the long *Ca*. E. dacicola - *B. oleae* co-evolutionary history and could be a very useful tool to clarify the debated olive fly origin.

In the third study (Chapter 4) the microflora composition inhabiting the oesophageal bulbs and midguts of *Rhagoletis cerasi* and *R. completa* has been analysed in detail for different years and in all life stages using both culturable and unculturable methods, revealing the presence of associated bacteria mainly belonging to Enterobacteriacee family and the lack of a specific symbiont in both fly species.

The fourth study (Chapter 5) investigates, using a molecular approach, the fate of bacteria associated with *Rhagoletis completa* during the different life stages of the fly. This is the first time for the tephritids that a bacterium typical of the fly microflora has been successfully engineered with the *gfp* technique and detected after feeding acquisition in larvae and pupae.

The last chapter (Chapter 6) deals with the phylogenetic relationships of flies of the Tephritinae subfamily and their symbiotic bacteria. The cophylogenetic analysis reveals a substantial congruence between hosts flies and symbionts.

References

Amann R.I., Ludwig W., Schleifer K.H. 1995. Phylogenetic identification and *insitu* detection of individual microbial cells without cultivation. *Microbiol. Rev*.**59**:143–69.

Augustinos A.A., Stratikopoulos E.E., Zacharopoulou A., Mathiopoulos K.D. 2002. Polymorphic micro satellite markers in the olive fruit fly, *Bactrocera oleae*. *Mol. Ecol. Notes* **2**:278–80.

Bandi C., Damiani G., Magrassi L., Grigolo A., Fani R., Sacchi L. 1994. Flavobacteria as intracellular symbionts in cockroaches. *Proc. R. Soc. Lond. B* **257**:43–48.

Bandi C., Sironi M., Damiani G., Magrassi L., Nalepa C.A., Laudani U., Sacchi L. 1995. The establishment of intracellular symbiosis in an ancestor of cockroaches and termites. *Proc R Soc Lond B.* **259**:293–299.

Baumann P., Baumann L., Lai C., Rouhbakhsh D., Moran N., Clark M. 1995. Genetics, physiology, and evolutionary relationships of the genus *Buchnera*: intracellular symbionts of aphids. *Ann. Rev. Microbiol*. **49**:55–94.

Behar A., **Yuval, B., Jurkevitch, E. 2005**. Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly *Ceratitis capitata*. *Mol. Ecol*. **14:**2637–2643.

Behar A., **Jurkevitch, E., Yuval, B. 2008**. Bringing back the fruit into fruit fly–bacteria interactions. *Mol. Ecol*. **17:**1375–1386.

Behar A., **Ben-Yosef M., Lauzon C.R., Yuval B., Jurkevich E. 2009**. Structure and function of the bacterial community associated with the Mediterranean fruit fly. In: Bourtzis K, Miller T (eds). Insect Symbiosis, Vol. 3. CRC Press: Boca Raton, pp 251–271.

Ben Ami E., Yuval B. & Jurkevitch E. 2010 Manipulation of the microbiota of mass-reared Mediterranean fruit flies *Ceratitis capitata* (Diptera: Tephritidae) improves sterile male sexual performance. *The ISME Journal*. **4:**28–37.

Ben-Yosef M., **Jurkevitch E., Yuval B. 2008a**. Effect of bacteria on nutritional status and reproductive success of the Mediterranean fruit fly *Ceratitis capitata*. *Physiol Entomol* **33:**145–154.

Ben-Yosef M., Behar A., Jurkevitch E., Yuval B. 2008b. Bacteria-diet interactions affect longevity in the medfly *Ceratitis capitata*. *J. Appl. Entomol.* **132:** 690–694.

Blümel S. & Masalmeh M.A. 1989. Primary investigation on the microbial endoflora of the European cherry fruit fly, *Rhagoletis cerasi* L. (Trypetidae) pp. 339-344. Edited by R. Cavalloro. Proceedings of the CEC/IOBC International Symposium. Rome April 1987.

Boller E.F. & Prokopy R.J. 1976. Bionomics and management of *Rhagoletis. Annu. Rev. Entomol.,* **21:**223-246.

Boyce A.M. 1934 Bionomics of the walnut husk fly, *Rhagoletis completa*. *Hilgardia* **8:**363-579.

Buchner P. 1965. Endosymbiosis of Animals with Plant Microorganisms. New York: Interscience Publishers, Inc.

Bourtzis K. **& O'Neill S.L. 1998**. *Wolbachia* infections and arthropod reproduction. *Bioscience*, **48:**287-293.

Bush G.L. **1966**. The taxonomy, cytology and evolution of the genus *Rhagoletis* in North America (Diptera: Tephritidae). *Bulletin of the Museum of Comparative Zoology* **134:**431-526.

Daane K.M. **& Johnson M.W. 2010**. Olive fruit fly : Managing an ancient pest in modern times. *Annu. Rev. Entomol.* **55:**151-169.

Dean R. W. 1935. Morphology of the digestve tract of the apple maggot fly, *Rhagoletis pomonella* Walsh, N. Y. State Agr. Exp.Station (Geneva, N.Y.) Bull. 229.

Dillon R.J. **& Dillon V.M. 2004**. The gut bacteria of insects: nonpathogenic interactions. *Annu. Rev. Entomol.* **49:**71–92.

Drew R.A.I. & Lloyd A.C. 1987. Relationship of fruit flies (Diptera: Tephritidae) and their bacteria to host plants. *Ann. Entomol. Soc. Am.* **80:**629–636.

Drew R.A.I. & Hancock D.L. **2000**. Phylogeny of the tribe Dacini (Dacinae) based onmorphological, distributional and biological data, p 491-504. In Aluja, M., and Norrbom, A.L. (eds), Fruit flies (Tephritidae): Phylogeny and Evoution behaior CRC Press, Boca Raton, FL.

Duso C. 1991. Sulla comparsa in Italia di un Tefritide neartico del noce: *Rhagoletis completa* Cresson (Diptera: Tephritidae). *Bollettino di Zoologia Agraria e di Bachicoltura* **23:**203- 209.

Duso C. & Dal Lago G. 2006. Life cycle, phenology and economic importance of the walnut husk fly *Rhagoletis completa* Cresson (Diptera: Tephritidae) in northern. *Italy Ann. soc. entomol. Fr.* **42:**245-254.

Fimiani P. 1989. Mediterranean Region. In: Robinson, A. S. and Hooper, G. (eds), Fruit Flies: Their Biology, Natural Enemies and Control, Vol. 3A, Elsevier, Amsterdam, pp. 37—50.

Fitt G.P. & O'Brien R.W. 1985. Bacteria associated with four species of *Dacus* (Diptera: Tephritidae) and their role in the nutrition of the larvae. *Oecologia*, **67:**447-454.

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. 571pp.

Freidberg A. 1984. Gall Tephritidae (Diptera), p. 129-167. In T. N. Ananthakrishnan, ed., Biology of gall insects. Oxford&IBH Publishing Co., New Delhi.

Girolami V. & 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.

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

Girolami V. 1979. Studi biologici e demoecologici sul *Dacus oleae* (Gmelin). 1. Influenza dei fattori ambientali abiotici sull'adulto e sugli stadi preimmaginali. *Redia*, **62:**147-191.

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.

Girolami V. 1986. Mediterranean fruit fly associated bacteria: transmission and larval survival. In *Pest Control: Operations and Systems Analysis in Fruit Fly Management*, NATO ASI Series, Vol. G 11, (M. Mangel, Ed.), pp. 135.146. Springer-Verlag, Berlin.

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

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.

Hagen K.S. & Tassan, R.L. 1972. Exploring nutritional roles of extracellular symbiotes on the reproduction of honeydew feeding adult chrysopids and tephritids. In Insect and Mite Nutrition: Significance and Implications in Ecology and Pest Management, pp. 323–351. Edited by J. G. Rodriguez. Amsterdam: North-Holland.

Howard D.J., Bush, G.L., and Breznak, J.A. 1985. The evolutionary significance of bacteria associated with *Rhagoletis. Evolution*. **39:**405–417.

Ishikawa H. **2003**. Insect symbiosis: an introduction, p. 1–16. *In* K. Bourtzis and T. Miller (ed.), Insect symbiosis. CRC Press, Boca Raton, FL.

Kasana A. & AliNiazee M.T. 1995. Adult flight dynamics of walnut husk fly (Diptera: Tephritidae) in the Willamette Valley of Oregon. *Pan-Pacific Entomology* **71:**142-148.

Katsoyannos B.I., Papadopoulos N.T., Stavridis D. 2000. Evaluation of trap types and food attractants for *Rhagoletis cerasi* (Diptera, Tephritidae). *J. Econ. Entomol*. **93:**1005–1010.

Kuzina L.V., Peloquin J.J., Vacek D.C., Miller T.A. 2001. Isolation and identification of bacteria associated with adult laboratory Mexican fruit flies, *Anastrepha ludens* (Diptera: Tephritidae). *Curr. Microbiol*. **42:**290-294.

Lauzon C.R., Sjogren R.E., Wright S.E., Prokopy R.J. **1998**. Attraction of *Rhagoletis pomonella* (Diptera: Tephritidae) flies to odor of bacteria: apparent confinement to specialized members of Enterobacteriaceae. *Environ. Entomol*. **27:**853–857.

Lauzon C. R. 2003. Symbiotic Relationships of Tephritids. In: Bourtzis K, Miller T (eds). Insect Symbiosis, Vol. 1. CRC Press: Boca Raton, pp 115–129.

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.

Lo N., **Gaku T., Watanabe H., Rose H., Slaytor M., Maekawa K., Bandi C., Noda H. 2000**. Evidence from multiple gene sequences indicates that termites evolved form wood-feeding cockroaches. *Curr. Biol*. **10:**801–804.

Manousis T. **& Moore N.F. 1987**. Control of *Dacus oleae*, a major pest of olives. *Insect Sci. Appl.* **8:**1–9.

Marchini D., Bernini L.F., Marri L., Giordano P.C., Dallai R. 1991. The female reproductive accessory glands of the medfly *Ceratitis capitata*: antibacterial activity of the secretion fluid. *Insect Biochem* **21:**597–605.

Marchini D., Rosetto M., Dallai R., Marri L. 2002. Bacteria associated with the oesophageal bulb of the medfly *Ceratitis capitata* (Diptera: Tephritidae). *Curr. Microbiol.* **44:**120–124.

Marzaro M. **2008**. Possibili apporti nutrizionali della simbiosi batterica in *Bactrocera oleae* (Rossi). M.S. thesis, University of Padova, Italy.

Mazzini M. & Vita G. 1981. Identificazione submicroscopica del meccanismo di trasmissione del batterio simbionte in *Dacus oleae* (Diptera, Trypetidae). *Redia* **64:**277–301.

Mazzon L., **Piscedda A., Simonato M., Martinez-Sañudo 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. Micr*. **58:**1277–1287.

Mazzon L., Martinez-Sañudo 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:**312-26.

Merz B. **1991**. *Rhagoletis completa* Cresson und *Rhagoletis indifferens* Curran zwei wirtschaftlich bedeutende nordamerikanische Fruchtfliegen, neu fur Europa (Diptera: Tephritidae). *Mitteilungen der Schweizerischen Entomologischen Gesellschaft* **64:** 55-57.

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

Nardi F., **Carapelli A., Dallai R., Roderick G.K., Frati F. 2005**. Population structure and colonization history of the olive fruit fly, *Bactrocera oleae* (Diptera, Tephritidae). *Mol. Ecol.* **14:**2729–38.

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.

Norrbom A.L., **Carroll L.E., Freidberg 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.

Norrbom A.L. & 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 81-52:114 p.

Ochando M.D. & Reyes A. 2000. Genetic population structure in olive fruit fly *Bactrocera oleae* (Gmelin): gene flow and patterns of geographic differentiation. *J. Appl. Entomol.* **124:**177– 83.

Olson W.H. & Buchner R.P. 2002. Leading edge of plant protection for walnuts. *Horttechnology* **12:**615-618.

Peloquin J.J., Kuzina L., Lauzon C., Miller T. 2000. Transformation of internal extracellular bacteria isolated from *Rhagoletis completa* Cresson gut with enhanced green fluorescent protein. *Curr. Microbiol*. **40:**367–371.

Peloquin J.J., Lauzon C, Potter S., Miller T. 2002. Transformed bacterial symbionts re-introduced to and detected in host gut. *Curr. Microbiol.* **45:**41-45.

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

Piscedda A. & Girolami V. 2005. Ghiandola faringea e produzione di agglomerati membranosi nei tefritidi: prima quantificazione in *Bactrocera oleae* (Gmelin). Proceedings del XX Congresso Nazionale Italiano di Entomologia p.50.

Pontes M.H. & Dale C. 2006. Culture and manipulation of insect facultative symbionts. *Trends Microbiol.* **14:**406–412.

Prado S.S. & Almeida R.P.P. 2009 Phylogenic placement of pentatomid stink bug gut symbionts. *Curr. Microbiol.* **58:**64–69.

Ratner S.S. & Stoffolano J.G. 1984. Ultrastructural changes of the esophageal bulb of the adult female apple maggot, *Rhagoletis pomonella* (Walsh) (Diptera: Tephritidae). *Int. J. Insect. Morphol. Embryol.* **13:**191–208.

Rice R.E. **1999**. Olive fruit fly, *Bactrocera* (*Dacus*) *oleae*. *UC Plant Protection Quarterly*, **9:**2.

Rice R.E. 2000.Bionomics of the olive fruit fly *Bactrocera* (*Dacus*) *oleae*. Plant Prot. Q. **10:**1–5.

Riegler M. & Stauffer C. 2002. *Wolbachia* infections and superinfections in cytoplasmically incompatible populations of the European cherry fruit fly *Rhagoletis cerasi* (Diptera, Tephritidae). *Mol. Ecol.* **11:**2425–2434.

Riegler M., Charlat S., Stauffer C., Merçot H. 2004. *Wolbachia* Transfer from *Rhagoletis cerasi* to *Drosophila simulans*: Investigating the Outcomes of Host-Symbiont Coevolution. *Appl. Environ. Microbiol.***70:**273–279.

Rossiter M.C., Howard D.J. & Bush G.L. 1983. Symbiotic bacteria of *Rhagoletis pomonella*, p. 77-84. 4223 In R. Cavalloro, ed., Fruit flies of economic importance. Proceedings of the CEC/IOBC International Symposium, Athens, Greece, November 16-19, 1982. A. A. Balkema, Rotterdam.

Seljak G. 1999. Orehova muha (*Rhagoletis completa* Cresson) nov nevaren skodljivec orehov v Sloveniji. *Revija za Sadjarstvo Vinogradnistvo in Vinarstvo* **10:**12-15.

Silvestri F. **1916**. Prima notizia sulla presenza della mosca delle olive e di un parasita di essa in India. *Rendiconti Della Reale Accademia dei Lincei*, **25:**424.

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

Tsakas S.C. **& Zouros E. 1980**. Genetic differences among natural and laboratory-reared populations of the olive fruit fly *Dacus oleae* (Diptera, Tephritidae). *Entomol. Exp. Appl.* **28:**268–76.

Tsiropoulos G.J. 1976. Bacteria associated with the Walnut Husk fly, *Rhagoletis completa*. *Environ. Entomol.* **5:**83-86.

Tsiropoulos G.J. 1977. Reproduction and survival of adult *Dacus oleae* feeding on pollens and honeydews (Diptera: Tephritdae). *Environ. Entomol.* **6:**390–92.

Tsiropoulos G.J. **1980**. Major nutritional requirements of adult *Dacus oleae*. *Ann. Entomol. Soc. Am.* **73:**251–253.

Trematerra P., Paparatti B., Girgenti P. 1995. Attenzione alla presenza della mosca delle noci. *Informatore agrario* **47:**74-76.

Tzanakakis M.E. 2006. Insects and Mites Feeding on Olive: Distribution, Importance, Habits, Seasonal Development and Dormancy. Leiden: Brill Acad. Publ. 182 pp.

Werren J.H. & O'Neill S.L. 1997. The evolution of heritable symbionts. In *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction* (S.L. O.Neill, A.A. Hoffmann, and J.H. Werren, Eds.), pp. 1.41. Oxford University Press, Oxford, U.K.

White I.M. **1988**. Tephritid flies (Diptera: Tephritidae). Handbooks for the Identification of British Insects 10 (5a): Dorset Press, Dorchester, Dorset, 134 pp.

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

White I.M. & Elson-Harris M.M. 1992. Fruit flies of economic significance: their identification and bionomics. CAB International, Wallingford, United Kingdom.

Zook D. 1998. A new symbiosis language. *Symbiosis News* **1:**1–3.

Zwölfer H. 1983. Life systems and strategies of resource exploitation in tephritids. pp. 16-30. *In:* E. Cavalloro, ed. Fruit Flies of Economic Importance. Proc. CEC/ IOBC Intl. Symp. Rotterdam, Balkema.

Chapter 2

Evidence of two lineages of the symbiont "*Candidatus* Erwinia dacicola" in Italian populations of *Bactrocera oleae* (Rossi) based on 16S rRNA gene sequence

---

Submitted as:

Savio C., Mazzon L., Martinez-Sañudo I., Simonato M., Squartini A. and Girolami V. Evidence of two lineages of the symbiont "*Candidatus* Erwinia dacicola" in Italian populations of *Bactrocera oleae* (Rossi) based on 16S rRNA gene sequence. International Journal of Systematic and Evolutionary Microbiology.

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

Abstract

A close association between the olive fly *Bactrocera oleae* (Rossi) (Diptera: Tephritidae) and bacteria has been known for more than a century. Recently, the presence of an host specific, hereditary, unculturable symbiotic bacterium, designated as '*Candidatus* Erwinia dacicola', has been described inside a cephalic organ of the fly, called oesophageal bulb.
In the present work we analyzed the 16s rRNA variability of "*Ca.* E. dacicola" within and among 26 Italian olive flies populations sampled across areas where olive trees occur in the wild, and areas where cultivated olive trees were introduced in historical times.
The bacterial content of the oesophageal bulbs of 314 olive flies was analyzed and a minimum of 781bp of 16S rRNA gene was sequenced. The corresponding host fly genotype was assessed by sequencing a 776 bp portion of its mitochondrial genome. Two "*Ca*. E. dacicola" haplotypes were found (htA and htB), one being slightly more prevalent than the other (57%). The two haplotypes do not co-exist in the same fly individual, as confirmed by cloning. Interestingly the olive fly populations of the two main Italian islands, Sicily and Sardinia, appeared exclusively represented respectively by htB and htA while peninsular populations showed both bacterial haplotypes in different proportions. No significant correlation emerged between the two symbiont haplotypes and the 16 host fly haplotypes observed, suggesting evidences for a mixed model of vertical and horizontal transmission of the symbiont during the fly life cycle.

Introduction

The microbiota associated with Tephritid flies (Diptera, Tephritidae) has been the object of many studies (Behar *et al.*, 2009). In particular, the associations between the olive fly *Bactrocera oleae*, the most important olive crop pest, and its symbiotic bacteria have been explored for more than a century. Petri (1909), based on microscopic observations, described for the first time the presence of bacterial masses inside an intestinal diverticulum located in the fly head called oesophageal bulb and the intestinal lumen of the midgut of *B. oleae*. In a previous study (Capuzzo *et al.*, 2005), using molecular techniques and microbiologically controlled conditions, we determined the nature of such unculturable and

host-specific bacterium, which, under the name of "*Candidatus* Erwinia dacicola", hereafter *E. dacicola,* was designated as novel taxon within the family Enterobacteriaceae.

The olive fly symbiont has also been found in Italy, where it was firstly described (Capuzzo *et al.,* 2005; Sacchetti *et al.,* 2008), in Spain (Silva *et al.,* 2008), in southwestern United States (Estes *et al.,* 2009) and in Greece (Kounatidis *et al.,* 2009).

The olive fly larvae feed upon the pulp of fruits of the genus *Olea,* both wild (oleaster) and domesticated forms (Tzanakakis, 2006). Botanically, the oleaster and the olive tree cultivated form correspond, respectively, to *Olea europaea* subsp. *europaea* var. *sylvestris* and var. *europaea* (Breton *et al*., 2006).

The olive fly is found nowadays in East and South Africa, the Mediterranean basin and Pakistan (Tzanakakis, 2006). Recently, the fly has been accidentally introduced in California (Rice, 1999). In Italy, as in others Mediterranean countries where the pest is recorded, the olive fly range traditionally matches with the olive tree range, including both oleaster and the cultivated form.

On the basis of these statements the Italian olive fly distribution could be divided in: a) areas where olive trees are grown extensively and the presence of oleaster is diffused (these areas overlap with historic olive cultivated regions and with the Italian phytoclimatic sub-zone of warm *Lauretum* according to Pavari (1916)); b) areas where olive trees are marginally cultivated and oleaster is absent. It is well-know that in the former of these two zones *B. oleae* populations can perform several generations all-year round, also exploiting residual fruits on wild olive trees present during spring time (Blando & Mineo, 2006).

To date the genetic diversity within the olive fly symbiont has never been investigated. For this purpose in this study we examined the olive fly symbiont *E. dacicola* 16S rDNA diversity, focusing on Italian olive fly populations with the following approaches: i) verifying the 16S rDNA *E. dacicola* sequence diversity all over Italian territory, ii) assessing possible correlations between *E. dacicola* genetic variability and mitochondrial haplotypes of the host fly.

Materials and methods

Olive flies sampling

Infested olives were collected from 26 different locations throughout Italy from 2002 to 2009 in the current cultivation areas of olive trees. Sampling sites were chosen in order to represent both wild olive trees areas according to Pavari (1916) and Carrión *et al*. (2010) (Liguria coast, south Italy, Sardinia and Sicily, loc. 6 and 13-26) and areas where cultivated olive was introduced and diffused through human-activities in historical time and where the wild olive (oleaster) is absent (North and Central Italy, loc. 1-5 and 7-12) (Tab. 2.1 and Fig. 2.1). In addition, two populations from both the areas were collected for three consecutive years and in the same season in order to detect possible variations on the distribution of different *E. dacicola* haplotypes in time (loc. 5, from 2007 to 2009 and 6, from 2006 to 2008) (Tab. 2.1).

The olives were kept into transparent plastic boxes (20x15x8 cm) topped with a net until the adults emerged. Insects were then transferred into rearing net cages (10x10x10 cm) and fed with a sugar diet (50% glucose solution) and water *ad libitum*.

Insects dissection and symbionts analysis

As in newly-emerged adults the endosymbiont presence is sometime too low to be detected (Girolami, 1973; Estes *et al.*, 2009) only 2-3 days-old flies were processed, ensuring the presence of sufficient number of bacteria in their oesophageal bulbs.

Flies were aseptically dissected (Capuzzo *et al.*, 2005) under a laminar flow hood in physiological saline solution (0.9% NaCl) under a stereomicroscope, extracting the oesophageal bulbs. Dissecting tools (forceps and tweezers) were sterilized before each individual insect processing. The oesophageal bulbs were gently transferred into Eppendorf tubes and kept at -20°C until further treatments.

A protocol originally developed for actinomycetes was used (Palmano *et al.*, 2000) to extract the microbial DNA content of the oesophageal bulbs. The bacterial 16S rRNA gene was partially amplified by PCR using two couples of universal bacterial primers: fd1 and rp1 (Weisburg *et al.*, 1991) or 63f and 1389r (Osborn *et al.*, 2000). The reaction mixture contained, in a total volume of 20 µl, 1x PCR GoTaq Flexi Buffer (Promega), 2.5

mM MgCl2, 0.1 mM dNTPs, 0.5 µM concentrations of each primer, 1 U of GoTaq Flexi DNA polymerase (Promega), and 1 µl of a 1:30 dilution of the DNA extract. All reaction mixtures were prepared at 4°C in 0.2 ml reaction tubes to avoid nonspecific priming. The standard thermal profile for the amplification of 16S rRNA genes was as follows: initial denaturation 95°C, 2 min; 27 cycles of: denaturation 95°C 30 s; annealing 56°C 30 s and extension 72°C 90 s; final extension 72°C, 5 min.

The 23S rRNA region was also tested for some individuals using the universal primers pair 1623f and 2490r (Hunt *et al*., 2006) under the following conditions: initial denaturation 94°C, 3 min, 30 cycles of: denaturation 94°C 60 s, annealing 51°C 60s and extension 72°C 90 s; final extension 72°C 5 min.

PCR products were checked by 1.0% agarose gel stained with SYBR® Safe (Invitrogen) and purified with the ExoSAP-IT kit (Amersham Biosciences) before sequencing.

The content of 8 oesophageal bulbs after PCR was cloned into JM109 competent cells using the P-GEM-T Easy vectors (Promega), following the manufacturer's recommendations. Transformation was verified using PCR assays with the M13-T7 universal primers pair.

Insect host analysis

A portion of the mitochondrial DNA of the olive fly corresponding to part of the NADH dehydrogenase subunit 1, the leucine tRNA and the 16S mitochondrial region, was amplified and sequenced using the universal primer pair N1-J12261m and LRN13398 (Simon *et al*., 1994; 2006). This analysis was carried on the same flies whose oesophageal bulb content was sequenced. A total of 15 olive fly populations out of 26, representative of the main Italian regions, was processed for this task. Amplifications were performed in 20 µl reactions (1x PCR Go Taq Flexi buffer (Promega), 2.5 mM MgCl2, 0.1 mM dNTPS, 0.2 µM each primer, 0.5 U of Taq polymerase (Promega), 2µl template DNA). Thermal cycling condition were 5 min at 96°C followed by 35 cycles of 96°C for 50 s, 56°C for 50 s, and 72°C for 2 min, with a final extension of 72°C for 5 min.

Sequencing and data analysis

Sequencing was performed at the BMR Genomics service (Padova, Italy) on automated DNA sequencers employing the primers used for PCR amplification and, in the case of *E. dacicola*, two additional primers for the 16S rRNA gene: fL2 (Capuzzo *et al*., 2005) and fL4 (Mazzon *et al*., 2008).

Sequences were edited and aligned using MEGA 4.0.2 (Tamura *et al*., 2007) and chromatograms were visually inspected. A BLASTN GenBank analysis of the sequences obtained was run through the NCBI website (www.ncbi.nlm.nih.gov) to assess the identity of the sequences obtained. The software package Arlequin Version 3.1 (Excoffier *et al*., 2005) was used to perform the following genetic analyses on the olive fly symbiont: i) Analysis of population differentiation was done by conducting exact tests of population differentiation with 100,000 steps in Markov chain, with 10,000 dememorization steps. ii) An analysis of molecular variance (AMOVA) was performed to define the partition of genetic variability among populations and among groups of populations. iii) The correlation between geographical and genetic distances (F_{ST} values) among populations was assessed by the nonparametric Mantel test using 1,000 permutations. For this analysis the pairwise F_{ST} values were transformed as log [F_{ST} /(1-F_{ST})] (Rousset, 1997).

A nested clade phylogeographic analysis (NCPA) was implemented by the program package ANeCa ver. 1.2 (Panchal, 2007), which includes TCS 1.21 (Clement *et al*., 2000) and GeoDis 2.5 (Posada *et al*., 2000). The haplotype network for the mitochondrial olive fly region was constructed using the statistical parsimony algorithm (Templeton *et al*., 1992) which outputs the 95% plausible set of most parsimonious linkages among haplotypes. The NCPA first tests the null hypotheses of no association between geography and the haplotype network. Only when this null hypothesis is rejected at the 5% level of significance it is possible to infer likely historical and geographical events (Templeton *et al*., 1995). Fisher's exact test was applied to observe associations between haplotype of *E. dacicola* and haplotype of insect host.

Results and Discussion

Erwinia dacicola *diversity*

A total of 314 olive fly oesophageal bulbs from 26 Italian populations were analyzed with an average of 11.5 individuals per population. GenBank similarity search (Blastn) showed that the 781 bp 16S rRNA fragment matched with 16S ribosomal RNA gene of *E. dacicola*. Only in three olive flies other bacteria were predominant (data not shown). When aligned by MEGA, *E. dacicola* sequences showed the presence of three associated point mutations. Two *E. dacicola* haplotypes were thus identified, hereafter referred as haplotype A (htA) and haplotype B (htB). Sequences were respectively deposited under the accession nos. HQ667588 and HQ667589. In particular three transitions were identified: at positions 11 and 282 htA showed a C and htB showed a T, and at position 657 htA showed a G and htB showed an A. Totally htA was present in 43% of the cases (134/311), htB was present in the remaining 57% (177/311).

Moreover, the 16S rRNA sequences of 43 *E. dacicola* samples coming from 9 localities, 21 of which belonging to htA and 22 to htB, were extended to 3' direction using the fL2 primer up to 1280 bp. Two more mutations were identified: at positions 943 htA showed a C and htB showed an A, and at position 945 bp htA showed a T and htB showed an A.

The presence of these point mutations allowed us to use a restriction enzyme (*Apa* I) to distinguish successfully and faster the 2 different bacterial haplotypes. The digestion technique was used in addition to standard sequencing.

To date 10 sequences attributed to *E. dacicola* 16S rRNA have been deposited in GenBank (5 from USA; 4 from Spain and one from Italy). Eight of them present a full length coverage with the sequences of our two haplotypes hence only these sequences were compared with our results. All the American sequences (GQ478373, GQ478377, GQ478378), 2 from Valencia, Spain (FM958429, FM958431) and one from Italy (AJ586620) matched with htB, while the remaining 2 sequences (FM958428, FM958430) both from Valencia, are identical to htA.

The 23S rRNA region of ten oesophageal bulbs contents was also tested. These flies were chosen from different populations (locations: 2, 3, 6, 10, 12, 13, 17, 20, 22, 24). From the DNA amplification and sequencing with the universal bacterial primers, a 761 bp fragment

was obtained (accession no. HQ667590). Upon aligning these sequences no nucleotide differences appeared among them. The 16S rRNA sequencing of the same samples showed that six of them corresponded to htA, four of them corresponded to htB.

In order to check the co-presence of both bacterial haplotypes in a single oesophageal bulb, the content of 8 *B. oleae* bulbs was cloned. Flies whose oesophageal bulb was cloned were mostly chosen from populations that exhibited both bacterial haplotypes (locations: 3, 5, 6, 9, 10, 16, 17, 24, Tab. 2.1). From 7 to 12 amplicons for each individual were sequenced or analyzed with restriction enzyme (*Apa* I). The results confirmed the presence of a unique *E. dacicola* haplotype in every olive fly oesophageal bulb suggesting that htA and htB do not appear to coexist inside the same olive fly. Intermediate haplotypes between htA and ht B have never been found, thus implying an ancient separation of the two bacterial variants probably originated after a prolonged period of isolation in different geographic areas. From the literature, the olive fly is supposed to have reached the Mediterranean area in historical times from the Middle East with the olive cultivated form (Augustinos *et al.*, 2005), as well as from Africa with the wild olive trees (oleaster) at the beginning of the Quaternary and remained associated with its host plant during the retreat to Pleistocenic glacial refugia (Nardi *et al.*, 2005; 2010). Moreover, a clear genetic diversification was described in wild olives between the Eastern and Western Mediterranean area as a consequence of Pleistocenic fragmentation in different glacial refugia (Besnard *et al.*, 2002; 2007). To obtain a rough estimate of the time of split between the two haplotypes we considered *Stammerula* sp., the phylogenetically closest symbiont of *E. dacicola* (Mazzon *et al.*, 2008; 2010), that shows a similar kind of symbiotic relationship within fruit flies. Using the substitution/site per million year obtained from *Stammerula* sp. (0.00118-0.00147, calculated with the absolute substitution rate of *Stammerula* sp. and the time of divergence of the insect host, (Martinez-Sañudo *et al., in preparation*) we tentatively estimated the divergence time between the two *E. dacicola* haplotypes in 1.3-1.6 million years ago. This period is compatible with the pleistocenic recolonization of the Mediterranean basin by the wild olives and *B. oleae* from different glacial refugia, while it is inconsistent with the last glaciation age.

Erwinia dacicola *haplotypes distribution*

With the exception of Marostica (locality n. 4) in all the peninsular populations both htA and htB are present, at different rate frequencies (Tab. 2.1). Surprisingly the olive fly populations of the two main Italian islands (Sicily and Sardinia) solely harbour a unique symbiont haplotype: in Sardinia only htA was found, while htB was the only one recorded in Sicily. The geographical distribution of the two *E. dacicola* haplotypes htA and htB is shown in Fig. 2.1. Geographic isolation could be the main explanation to this evident homogeneous distribution within the islands, as opposed to the haplotype mix observed in the peninsula populations. However, it is unlikely to assume that throughout the ages accidentally introduction in the islands of olive flies carrying the other symbiont haplotype never happened, especially for a species like *B. oleae* tightly connected with cultivation and commercial human-mediated exchanges.

The presence of population differentiation was confirmed by the exact tests of population differentiation ($P < 0.001$). When analyzing the *E. dacicola* haplotypes distribution among the 26 Italian olive fly populations by AMOVA, a clear geographic pattern does not appear. AMOVA was performed to analyze the origin of molecular variability in the different hierarchical levels and groups (Tab. 2.2). In the first, all the 26 populations analyzed were considered as forming a single group. In this case AMOVA showed that about 39% of variation was explained by differences among populations although the highest variation occur within the populations themselves (60.83%). When only the peninsular populations were considered, the percentage of variance among populations decreased to 26.25. In the successive AMOVA tests the populations were grouped according to the Italian phytoclimatic zones proposed by Pavari (1916) or according to the presence of geographic isolation. A variance value of 38% was observed when considering the sea as geographical barrier and thus combining the populations into three groups (peninsular populations vs. Sicily populations vs. Sardinia populations) (Tab. 2.2a). Moreover, in order to test the additional influence of the orographic barrier of Apennines range the populations were clustered into eight groups: north-east vs. north-west vs. central-west vs. central-east vs. south-west vs. south-east vs. Sicily island vs. Sardinia island (Tab. 2.2b). This grouping explained a lower but significant proportion of variation (34.53%), whereas it decreased (19.42%) upon removing the two islands groups (Tab. 2.2c). Finally we grouped only the

populations belonging to the phytoclimatic sub-zones of the warm *Lauretum* (supposed to be the refugial areas of olive tree during the cold periods) into 4 clusters that presently appear geographically isolated from each other (Liguria region vs. south Italy vs. Sicily island vs. Sardinia island). In this case AMOVA test showed a very high and significant percentage of variation (51.15%) (Tab. 2.2d).

According to the Mantel test based on 26 populations, genetic and geographical distances were significantly correlated (r = 0.13, P = 0.026). If peninsular populations were split into two groups, separated by the orographic barrier of the Apennine mountains, the Mantel test revealed an evidence for isolation by distance when considering the olive flies populations located along the Tyrrhenian coast (locations 6, 7, 8, 9, 10, 13, 14, 16, 17) (r = 0.35, P= 0.03), showing an increasing proportion of htB going southern but becomes non-significant for locations near the Adriatic coast (locations 11, 12, 18, 19, 20) (r = -0.18, P= 0.61).

In order to assess variations on the distribution of the two haplotypes over time two olive flies populations, each composed by both *E. dacicola* haplotypes, were monitored for three consecutive years (locality 5 from 2007 to 2009 and locality 6 from 2006 to 2008) (Tab. 2.1). The proportions of the two bacterial haplotypes in each population appeared not to be random as Fisher's exact test showed no significant differences in the htA-htB distributions over the years for both populations (locality 5, Fisher's exact test, df=2, P=0.56; locality 6, Fisher's exact test, df=2, P=0.98). This results suggests that the distribution of the two haplotypes seems constant over time and specific of the population.

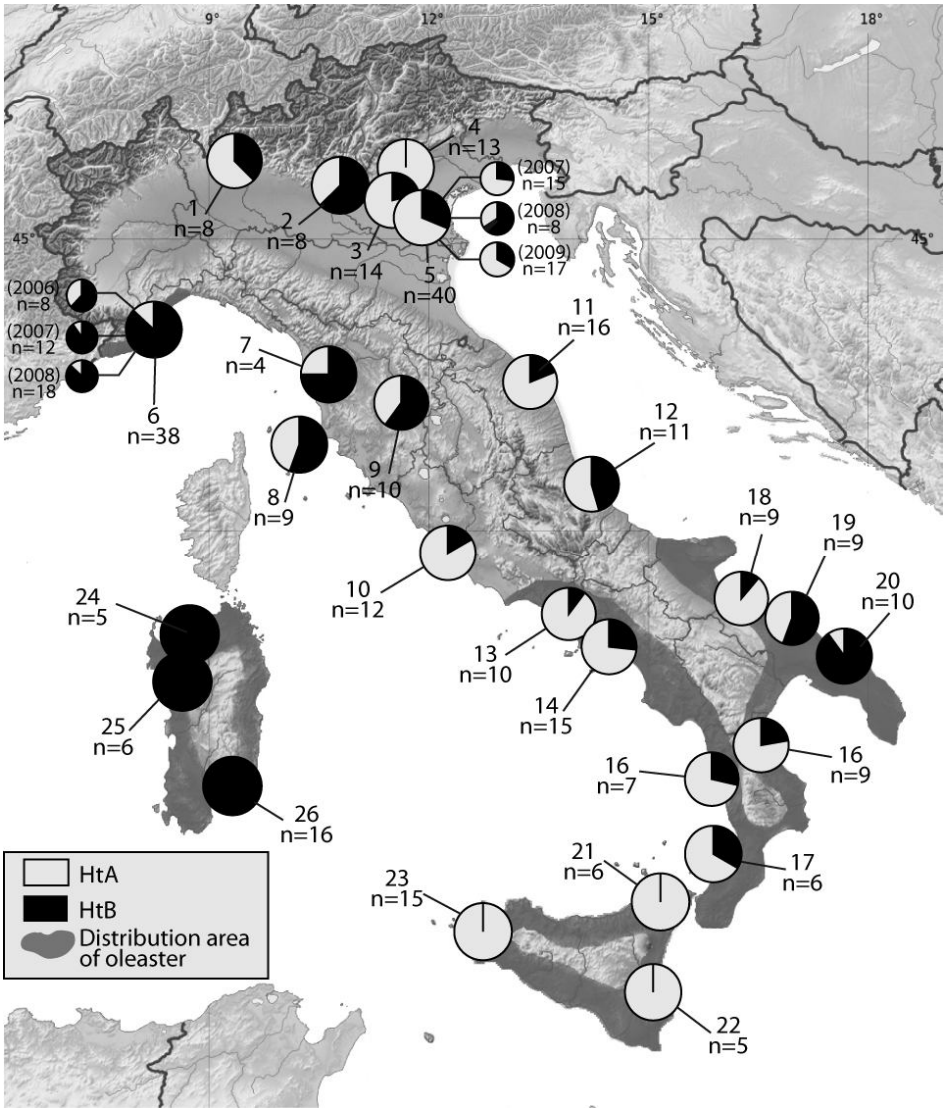

Fig. 2.1. Geographic distribution and proportion of the two "*Candidatus* Erwinia dacicola" haplotypes among the 26 *B. oleae* populations sampled in Italy. The first number indicates the location code, reported in Tab. 2.1. n= sample size. For locations 5 and 6 the three different years sampling are reported in details. Distribution area of oleaster by Carrion *et al*. (2010).

Structure	Source of variation	d.f.	Variance (%)	*F* indices	*P-value*
All populations	Among populations	25	39.17	F_{ST}=0.39168	<0.001
	Within populations	285	60.83		
Only peninsulae populations	Among populations	19	26.25	F_{ST}=0.26251	<0.001
	Within populations	238	73.75		
(a) Grouping by geographical region I	Among groups	2	38.01	F_{CT}=0.38011	<0.001
	Among populations within groups	23	16.98	F_{SC}=0.27397	<0.001
	Within populations	285	45.01	F_{ST}=0.54994	<0.001
(b) Grouping by geographical region II	Among groups	7	34.53	F_{CT} =0.34529	<0.001
	Among populations within groups	18	6.86	F_{SC}=0.10484	<0.001
	Within populations	285	58.61	F_{ST}=0.41393	<0.001
(c) Grouping by geographical region III	Among groups	5	19.42	F_{CT}=0.19417	<0.05
	Among populations within groups	14	8.89	F_{SC}=0.11035	<0.01
	Within populations	238	71.69	F_{ST}=0.28309	<0.001
(d) Grouping by wild olive areas	Among groups	3	51.15	F_{CT}=0.51155	<0.01
	Among populations within groups	11	12.61	F_{SC}=0.25820	<0.01
	Within populations	165	36.23	F_{ST}=0.63767	<0.001

Tab. 2.2. Analysis of molecular variance (AMOVA) of the symbiont "*Candidatus* Erwinia dacicola" based on 16S rRNA data among populations of Bactrocera oleae divided according to the phylogeographic hypotheses discussed in the text.
(a) Group 1: peninsulae (localities 1-20); group 2: Sicily (localities 21-23); group 3: Sardinia (localities 24-26).
(b) Group 1: north-east (localities 1-5); group 2: north-west (locality 6); group 3: central-west (localities 7-10); group 4: central-east (localities 11-12); group 5: south-west (localities 13-17); group 6: south-east (localities 18-20); group 7: Sicily (localities 21-23); group 8: Sardinia (localities 24-26).
(c) Group 1: north-east (localities 1-5); group 2: north-west (locality 6); group 3: central-west (localities 7-10); group 4: central-east (localities 11-12); group 5: south-west (localities 13-17); group 6: south-east (localities 18-20).
(d) group 1: Liguria region (locality 6); group 2: South (localities 13-20); group 3: Sicily (localities 21-23); group 4: Sardinia (localities 24-26).

Matching extent of symbiont and host haplotypes

A total of 776 bases corresponding to part of the NADH dehydrogenase subunit 1, the leucine tRNA and the 16S rDNA mitochondrial region, was sequenced for 80 olive flies representing 15 populations in Italy (Tab. 2.1). DNA sequences were aligned to identify polymorphisms. A total of 16 variant sequence forms were identified (no gaps were found) and given haplotype designations h1 to h16. These are shown in Table 2.1, which also provides details on their distribution by locality. The most common and widely spread haplotypes (h1 and h2) are shared respectively by 12 and 11 populations. Seven haplotypes instead resulted unique. Each of the Italian olive flies populations analyzed harboured on average 3.2 haplotypes. Sequences were deposited in GenBank under accession nos HQ667572 to HQ667587. Using the program TCS, a network of the haplotypes was constructed (Fig. 2.2). It shows that 15 of the 16 haplotypes identified are directly connected and only two haplotypes are missing (Fig. 2.2). NCPA performed with ANeCA showed that the null hypothesis of no association between network structure and geography cannot be rejected.

The overlay of the two *E. dacicola* haplotypes over the host mtDNA data, did not show any evident association (Fig. 2.2). In all the interior haplotypes (h1, h2, h3, h5,) that tend to be older and more frequent than tip haplotypes (Posada *et al*., 2006), both *E. dacicola* lineages are present and their proportions do not differ significantly (Fisher's exact test, df=3, P=0.577). The external mitochondrial haplotypes are not considered in this analysis because they are rare and thence forcedly correlated with either one of the two *E. dacicola* haplotypes. The lacking of association between symbiont and host lineages could be at least partly explained assuming an horizontal transfer of symbiont. Even if recently an intracellular existence has been hypothesized for *E. dacicola* during the larval stage (Estes *et al*., 2009), the olive fly symbiont can be at present, according to Petri (1909), considered extracellular. The extracellular condition offers more potential opportunity for contacts compared to endocellular symbionts lifestyle. Moreover, the olive fly lifestyle, especially during the larval stage, offers opportunities for symbiont losses and reacquisition to occur, making a different *E. dacicola* haplotype replacement possible.

As described by Girolami (1973) and Capuzzo *et al*. (2005), whereas in the adult flies symbiont are generally located in a close environment within the oesophageal bulb, in the

larval stage bacteria are located in the intestinal blind sacs in direct contact whit midgut lumen and thus exposed to possible replacements. Like the majority of the fruit flies, *B. oleae* is tightly associated with its host plant and it spends its larval life feeding on the olive fruit pulp. It is common to observe, especially during heavy infestations, the contemporary presence inside the same olive of more than one larva with occasionally crossing galleries; this could offer chance for possible replacement by a different haplotype of *E. dacicola* bypassing the model of strict vertical transmission. In such as scenario, the co-existence of both *E. dacicola* lineages could be postulated inside the same insect host but, in our analysis even though it was carried out solely on adults, this has never been observed (as confirmed by cloning). Therefore, the combined occurrence of the two *E. dacicola* variants in a single fly seems to be extremely rare or not possible at all, due to unknown intimate mechanisms to be investigated mainly at the larval stage.

It is well-known that symbiont losses are unavoidable in artificial substrates rearing conditions mainly during larval stages (Girolami & Cavalloro, 1972). The symbiont losses should therefore be considered as a frequent event during the olive fly life. A critical phase for this event could occur during the egg stage since the symbionts are borne externally. It is known that during the oviposition the mother infects the outer side of the eggs with masses of *E. dacicola* bacteria in order to ensure the vertical transmission to the progeny as reported by Petri (1909).

There are no clear indications about the colonization history of the olive fly and its geographical origin, which has been obscured by the long history of olive cultivation in the Mediterranean area (Daane & Johnson, 2010), but there are data about its African origin (Nardi *et al*., 2005, 2010; White, 2006). We think that this report could represent a first standpoint to understand the genetic variability of the olive fly in relation with that of its symbiont, and that further and more comprehensive studies could address the situation occurring in the entire colonization areas of *B. oleae* by extending the search over the whole Mediterranean basin. This line of research, besides being a useful tool to draw in details the olive fly colonization route, can be a model for many investigations pursuing the evolutionary interdependency between insects and their associated bacteria.

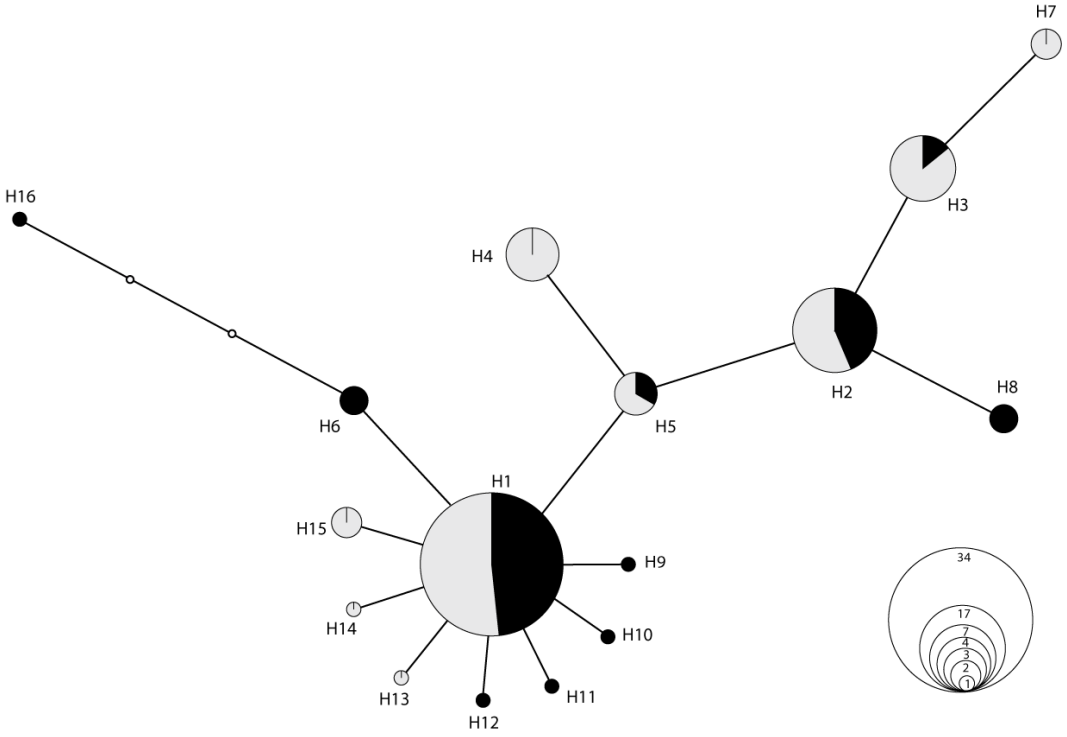

Fig. 2.2 Haplotype network of *Bactrocera oleae* realized by TCS 1.21 (Clement *et al*., 2000). Each haplotype is represented by a circle, with the area of the circle proportional to its frequency. The two tiny empty circles represent intermediate missing haplotypes. Inside each circle is reported the proportion of *Erwinia dacicola* lineage: haplotype A (black) and haplotype B (grey).

Acknowledgements

We thank A. Bozzolo, V. Caleca, S. Deliperi, C. Duso, M. Faccoli, D. Fornasiero, G. Frigimelica, M. Marzaro, B. Massa, N. Mori, M.E. Noce, G. Pellizzari, E. Petrucco Toffolo, C. Rapisarda, V. Vacante, C. Villari, L. Vivan, D. Zovi for their help in collecting samples.

This work is supported in part by the Italian Ministry of University (MIUR) under a Program for the Development of Research of National Interest (PRIN Grant No.2008BPLF7Y) and it is part of Savio C. PhD thesis, which is gratefully supported by "Veneto Agricoltura" scholarship.

Locality	Coordinates	Date of collection	Sample size*	*E. dacicola* haplotypes		*B. oleae* haplotypes[§]														
				htA	htB	h1	h2	h3	h4	h5	h6	h7	h8	h9	h10	h11	h12	h13	h14	h15	h16
1. Campione (Lombardy)	45°58'N-08°58'E	10/2009	8	3	5																
2. Bardolino (Veneto)	45.32'N-10°43'E	12/2008	8	5	3																
3. Dueville (Veneto)	45°38'N-11°32'E	10/2009	14	3	11																
4. Marostica (Veneto)	45°44'N-11°39'E	10/2007	13(10)	0	13	4_B	3_B	2_B											1_B		
5. Montegrotto terme (Veneto)	45°19'N-11°47'E	12/2007	15(5)	4	11	2_B	1_B		1_B											1_B	
		11/2008	8	4	4																
		12/2009	17	5	12																
6. Diano Marina (Liguria)	43°54'N-8°04'E	11/2006	8	6	2																
		09/2007	12(4)	11	1	1_B; 1_A	1_A	1_A													
		09/2008	18(6)	16	2	4_A							1_A								1_A
7. Firenze (Tuscany)	43°46'N-11°15'E	12/2007	4	3	1																
8. Capoliveri (Tuscany)	42°44'N-10°22'E	08/2009	9(5)	5	4	1_B; 2_A		1_B				1_B									
9. Arezzo (Tuscany)	43°28'N-11°51'E	11/2009	10	6	4																
10. Roma (Lazio)	41°53'N-12°29'E	12/2009	12(5)	2	10	1_A; 3_B	1_B														
11. Ancona (Marche)	43°37'N-13°31'E	09/2009	16(5)	3	13	2_B	3_B														
12. Moscufo (Abruzzo)	42°25'N-14°03'E	11/2009	11	5	6																
13. Capua (Campania)	41°06'N-14°13'E	09/2009	10	1	9																
14. Portici (Campania)	40°48'N-14°20'E	09/2009	15(5)	4	11	1_B			2_B			1_B								1_B	
15. Rossano Calabro (Calabria)	39°34'N-16°38'E	09/2009	9(4)	2	7	2_B	1_A; 1_B														
16. Rende (Calabria)	39°19'N-16°10'E	10/2009	7	2	5																
17. Gioia Tauro (Calabria)	38°25'N-15°54'E	11/2009	6(3)	2	4		1_B				1_A							1_B			
18. Bari (Apulia)	41°07'N-16°51'E	11/2002	9	1	8																
19. Castellana Grotte (Apulia)	40°53'N-17°09'E	10/2009	9	5	4																
20. Otranto (Apulia)	40°08'N-18°29'E	10/2009	10(5)	9	1	1_A;1_B	1_A			1_A							1_A				
21. Messina (Sicily)	38°11'N-15°33'E	08/2007	6(5)	0	6			2_B	1_B	2_B											
22. Siracusa (Sicily)	37°03'N-15°17'E	11/2008	5	0	5																
23. Marsala (Sicily)	37°47'N-12°26'E	07/2008	15(4)	0	15	2_B	1_B	1_B													
24. Bancali (Sardinia)	40°44'N-8°27'E	12/2009	5(3)	5	0	2_A										1_A					
25. Alghero (Sardinia)	40°33'N-8°19'E	10/2008	6(6)	6	0	4_A	1_A						1_A								
26. Cagliari (Sardinia)	39°12'N-9°06'E	09/2009	16(5)	16	0		2_A				1_A			1_A	1_A						
Total			311(80)	134	177	15_A; 19_B	6_A; 11_B	1_A; 6_B	4_B	1_A; 2_B	2_A	2_B	2_A	1_A	1_A	1_A	1_A	1_B	1_B	2_B	1_A

Tab. 2.1. Collection sites of the *Bactrocera oleae* populations, 16S rRNA haplotype frequencies of *Candidatus* Erwinia dacicola (htA and htB) and geographical distribution of the 16 *B. oleae* mitochondrial haplotypes.
* In brackets are indicated the number of olive flies whose mithochondrial region has been sequenced.
§ A and B indicate the symbiont haplotypes.

Chapter 3

Distribution of olive fly symbiont lineages in Mediterranean populations of *Bactrocera oleae* (Rossi)

Manuscript in preparation as:

Savio C., Martinez-Sañudo I., Mazzon L., Simonato M., Squartini A. and Girolami V. Distribution of olive fly symbiont lineages in Mediterranean populations of *Bactrocera oleae* (Rossi).

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

Introduction

The olive fly *Bactrocera oleae* (Rossi) is a pest of large economic importance, causing significant losses to oil quality and olives productions. Because of the monophagous behavior of the larvae that feed only on olive fruits, the fly is strictly related to the areas of its host, *Olea europea*. The history and the origin of the fly have long been studied and are still matter of debate. It is excluded that, the olive fly originated in Mediterranean region although nowadays it is found all over this area.

In the recent past several studies surveyed the genetic variability of the olive fly, using different molecular markers (Ochando & Reyes, 2000; Ochando *et al.*, 2003; Nardi *et al.*, 2003; 2005; 2006; 2010; Augustinos *et al.*, 2002; 2005; Segura *et al.*, 2008; Zygouridis *et al.*, 2009). According to Nardi *et al.* (2005, 2010) and White (2006), the olive fly probably originated in African countries, as well as its natural host, the wild olive; only in most recent times it invaded the Mediterranean region, attacking the *Olea* cultivated form (Nardi *et al.* 2005).

The presence of a hereditary symbiosis in the olive fly is well-known (Petri, 1909; Girolami, 1973; Capuzzo *et al.*, 2005; Sacchetti *et al.* 2008) and reported also in United States (Estes *et al.*, 2009), where *B. oleae* has recently been introduced (Rice, 1999). The olive fly symbiont, *Candidatus* Erwinia dacicola, hereafter *E. dacicola*, is harbored inside the oesophageal bulb (Petri, 1909), a specialized foregut *diverticulum* located in the olive fly head.

In our previous study (Savio *et al.*, submitted) we surveyed *E. dacicola* 16S rRNA variability across the Italian peninsula. As a results two symbiont haplotypes were identified, whose distribution was not random on the territory. Even though no significant correlation emerged between the two symbiont haplotypes and the host fly haplotypes observed, we think that a large scale investigation across the olive fly range can be useful to understand the evolution of symbiotic relationships between the host and its own symbiont.

In the light of these results we tried to investigate the *E. dacicola* genetic diversity together with its host fly variability in a wider range, extended to the Mediterranean area, South-Africa and Pakistan.

Materials and methods

Samples collection

Olive flies were collected as pupae or larvae from infested olives in 49 locations from 12 countries around the Mediterranean basin, plus one population from South Africa and one population from Pakistan (see Supplementary Table). This dataset includes the 26 Italian olive fly populations from Savio *et al*. (submitted). When adults' emergence occurred, flies were put into net rearing cages (10x10x10 cm) and fed with a sugar diet and water *ad libitum*. At least 2-3 days old adults were aseptically dissected under a laminar flow hood in physiological saline solution (0.9% NaCl) to extract the oesophageal bulb, that was gently transferred into Eppendorf tubes and kept at -20°C until processed.

DNA extraction, amplification and sequencing

A protocol originally developed for actinomycetes was used (Palmano *et al.*, 2000) to extract both the microbial DNA content of the oesophageal bulbs and the insect DNA.

Amplification of the 16s rRNA gene was performed using two couple of universal bacteria primers, fd1 and rp1 (Weisburg *et al.*, 1991) or 63f and 1389r (Osborn *et al*., 2000) as described by Savio *et al.,* (submitted). Moreover the content of 7 oesophageal bulbs after PCR was cloned into JM109 competent cells using the P-GEM-T Easy vectors (Promega), following the manufacturer's recommendations. Transformation was verified using PCR assays with the M13-T7 universal primers pair.

A region of the mitochondrial DNA of the olive fly comprising the 5' of NADH dehydrogenase subunit 1, the Leu-tRNA and the 16S 3' mitochondrial region was amplified using universal primer pair N1-J12261m- and LRN13398 (Simon *et al*., 1994, 2006) as described by Savio *et al.,* (submitted).

PCR products were checked by 1.0% agarose gel stained with SYBR® Safe (Invitrogen) and purified with the ExoSAP-IT kit (Amersham Biosciences) before sequencing.

Sequencing was performed at the BMR Genomics service (Padova, Italy) on automated DNA sequencers employing the primers used for PCR amplification and, in the case of *E. dacicola*, two additional primers for the 16S rRNA gene: fL2 (Capuzzo *et al*., 2005) and

fL4 (Mazzon *et al*., 2008). All sequences were aligned using MEGA 4.0.2 (Tamura *et al*., 2007), checked manually and blasted on Genbank (http://www.ncbi.nlm.nih.gov/BLAST/).

Candidatus Erwinia dacicola analysis

The *E. dacicola* sequences were analyzed using the program SAMOVA 1.0 (Dupanloup *et al*., 2002). We carried out this analysis despite the fact that sampling points of *B. oleae* are not geographically adjacent as it is assumed by this approach. SAMOVA can identify genetically distinct geographic groups of populations using a simulated annealing procedure that aims to maximise the proportion of total genetic variance due to differences between groups of populations (F_{CT}). After defining the number of groups (*K*) we ran 100 simulated annealing process for each possible *K*, ranging from $K = 2$ through $K = 8$, recording the progressive split of the populations according to their genetic variance. To select the optimal *K*, two criteria must be considered: F_{CT} values should reach a maximum and the configurations with one or more single-population groups should be excluded because this indicates that the group structure is disappearing (Magri *et al.*, 2006).

Fisher's exact test was applied to check for associations between haplotype of *E. dacicola* and haplotype of insect host.

Host data analysis

ARLEQUIN ver. 3.1 (Excoffier *et al*., 2005) was used to estimate the genetic variability of each *B. oleae* population, expressed as gene diversity (h), nucleotide diversity (π) and mean number of pairwise differences (k). Furthermore it was used to compute Tajima's *D* (Tajima, 1989) and Fu's F_S (Fu, 1997) neutrality tests among all sampling locations. Negative values of the neutrality indices could provide information on recent changes in demographic histories. F_S statistic should be considered significant if its *P*-value is below 0.02 (Fu, 1997). Significant negative *D* and *Fs* values can be interpreted as signatures of population expansion.

The haplotype network for the *B. oleae* mitochondrial data was constructed using the software TCS 1.21 (Clement *et al*., 2000). The TCS program creates a haplotype network using Statistical Parsimony (Templeton *et al.*, 1992), which outputs the 95% plausible set of most parsimonious linkages among haplotypes. Ambiguous linkages are depicted by

"loops" in the haplotype network. The network was used to perform a nested clade phylogeographic analysis (NCPA) using ANeCa ver. 1.2 (Panchal, 2007), which includes GEODIS ver. 2.0 (Posada *et al.*, 2000), to test the null hypothesis of lack of association between clades and geographic locations.

The analysis of molecular variance (AMOVA) was computed with the software package ARLEQUIN to infer population genetic structure. Mantel test, implemented in the same package software, was used to test the correlation between geographic and genetic distances (F_{ST} values) among the Mediterranean olive fly populations, using 1,000 permutations. For this analysis the pairwise F_{ST} values were transformed as log [F_{ST} /(1-F_{ST})].

Results

Erwinia dacicola *diversity and distribution*

A total of 524 olive fly oesophageal bulbs from 51 populations were analysed and in all cases the symbiont *E. dacicola* was detected. When aligned by MEGA three haplotypes were identified, here on called htA (accession no. HQ667588) htB (accession no. HQ667589) and htC (Tab. 3.1). The first two haplotypes were already reported in a previous study (Savio *et al.*, submitted) while htC is a new haplotype. This haplotype shows intermediate features between htA and htB, presenting the same nucleotides of htB at positions 11, 282, 943 and 945 (T, T, A, A respectively) and the nucleotide of htA at position 657 (G) (Tab. 3.1). Their distribution among sampled populations is reported in Tab. 3.2 and in the Fig. 3.1.

SAMOVA 1.0 was used to identify the optimal grouping of *E. dacicola* populations in the Mediterranean regions. The values for the *F* indices are shown in Fig. 3.2. The F_{CT} value reached the maximum at K=2 (F_{CT}= 0.6294). However, when K>2 the grouping started to disappear, that is, one group was composed of only one population. Thus we used K=2 as the best grouping scheme. Interestingly these two groups well correspond to the main geographical regions of the Mediterranean basin: The first group is mainly composed of populations from the west part of this area (locations 1-10, 17, 19, 20, 33, 37-39, 43, 44) encompassing also Crete populations, whereas the second group includes populations from

the east part of the Mediterranean basin (locations 11-16, 18, 21-32, 34-36, 40-42, 45-49). This geographical distribution is clearly shown in Fig. 3.1: the olive fly populations from Iberian peninsula, Morocco, France and Sardinia harbor (with the exception of few samples from location 6) exclusively htA, while the olive fly populations from Israel, Cyprus and Turkey harbor only htB. Populations located in the central Mediterranean area (Italy, for which the distribution has been already discussed in Savio *et al*. submitted, Slovenia and Greece) mostly harbor both *E. dacicola* haplotypes.

The South-Africa *B. oleae* population harbors htA: this result is confirmed for 2 consecutive years of sampling (data not shown). Interestingly, htC has been found only in the Pakistani *B. oleae* population.

In order to check the co-presence of different bacterial haplotypes in a single olive fly, the bacterial content of 7 oesophageal bulbs was cloned (1 from Valencia, 1 from Pula, 2 from Heraklion, 1 from Canakkale, 1 from Limassol and 1 from Stellenbosh). From 8 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 *E. dacicola* haplotype in each olive fly oesophageal bulb, as already shown by Savio *et al.* (submitted), suggesting that htA and htB do not coexist inside the same olive fly.

				Variable sites				
Haplotype		**Reference**	**Origin**	11	282	657	943	945
htA	HQ667588	Savio *et al*., submitted	Alghero (Italy)	C	C	G	C	T
	FM958430	Moret *et al*., unpublished	Valencia (Spain)	-	-	-	-	-
	FM958428	Moret *et al*., unpublished	Valencia (Spain)	-	-	-	-	-
htB	HQ667589	Savio *et al*., unpublished	Marsala (Italy)	T	T	A	A	A
	GQ478373	Estes *et al*., 2009	California (USA)	-	-	-	-	-
	GQ478377	Estes *et al*., 2009	California (USA)	-	-	-	-	-
	GQ478378	Estes *et al*., 2009	California (USA)	-	-	-	-	-
	FM958431	Moret *et al*., unpublished	Valencia (Spain)	-	-	-	-	-
	FM958429	Moret *et al*., unpublished	Valencia (Spain)	-	-	-	-	-
	AJ586620	Capuzzo *et al*., 2005	Bari (Italy)	-	-	-	-	-
htC		Present study	Swat (Pakistan)	T	T	G	A	A

Tab. 3.1 - Variable nucleotide positions of 16S rRNA sequences of "*Candidatus* Erwinia dacicola". Nucleotides identical to the top sequence are denoted by a dash.

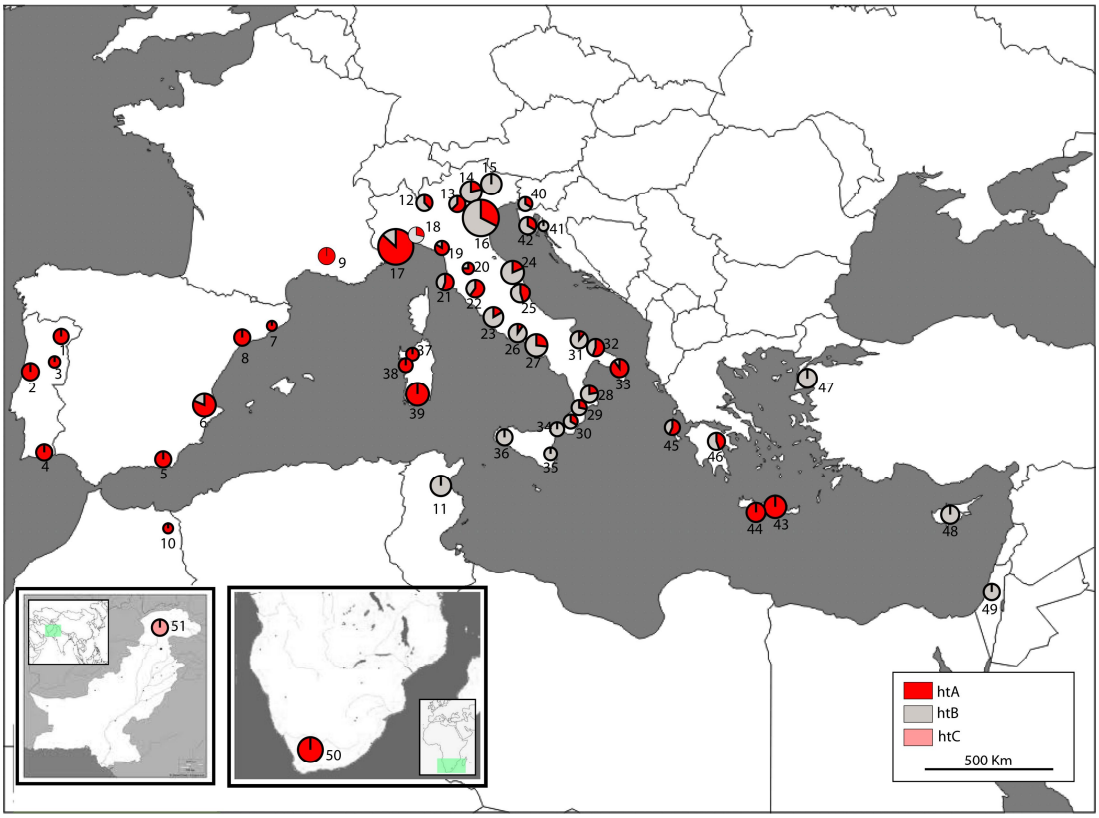

Fig. 3.1- Geographic distribution and proportion of the three "*Candidatus* Erwinia dacicola" haplotypes among the 51 *B. oleae* populations sampled. The number indicates the location code, reported in Tab. 3.2. The area of the circles is proportional to the sampling

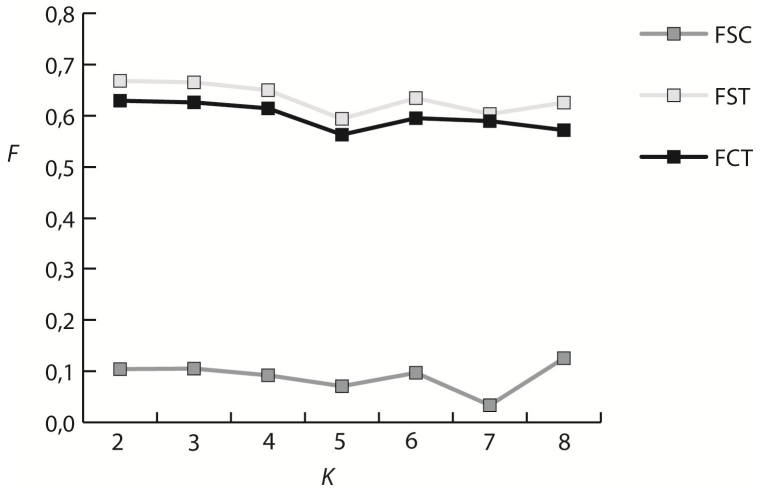

Fig. 3.2 - Values of fixation indices, *F*, obtained from SAMOVA for *Candidatus* Erwinia dacicola as a function of increasing number of groups (*K*). *F*CT: differentiation between groups; *F*ST: differentiation between populations among groups; *F*SC: differentiation between populations within groups.

	Locality	*N*	htA	htB	htC
1	Bragança (Portugal)	7	7	0	0
2	Lisbõa (Portugal)	9	9	0	0
3	SerraLeomil (Portugal)	4	4	0	0
4	Tavira (Portugal)	8	8	0	0
5	Almeria (Spain)	8	8	0	0
6	Valencia (Spain)	16	13	3	0
7	Tarragona (Spain)	3	3	0	0
8	Barcelona (Spain)	8	8	0	0
9	Avignon (France)	8	8	0	0
10	Oujda (Morocco)	3	3	0	0
11	Chaffar (Tunisia)	13	0	13	0
12	Campione (Italy)	8	3	5	0
13	Bardolino (Italy)	8	5	3	0
14	Dueville (Italy)	14	3	11	0
15	Marostica (Italy)	13	0	13	0
16	Montegrotto T. (Italy)	40	13	27	0
17	Diano Marina(Italy)	38	33	5	0
18	Camogli (Italy)	7	2	5	0
19	La Spezia (Italy)	6	5	1	0
20	Firenze (Italy)	4	3	1	0
21	Capoliveri (Italy)	9	5	4	0
22	Arezzo (Italy)	10	6	4	0
23	Roma (Italy)	12	2	10	0
24	Ancona (Italy)	16	3	13	0
25	Moscufo (Italy)	11	5	6	0
26	Capua (Italy)	10	1	9	0
27	Portici (Italy)	15	4	11	0
28	Rossano C. (Italy)	9	2	7	0
29	Rende (Italy)	7	2	5	0
30	Gioia T. (Italy)	6	2	4	0
31	Bari (Italy)	9	1	8	0
32	Castellana G. (Italy)	9	5	4	0
33	Otranto (Italy)	10	9	1	0
34	Messina (Italy)	6	0	6	0
35	Siracusa (Italy)	5	0	5	0
36	Marsala (Italy)	8	0	8	0
37	Bancali (Italy)	5	5	0	0
38	Alghero (Italy)	6	6	0	0
39	Cagliari (Italy)	16	16	0	0
40	Strunjan (Slovenia)	6	2	4	0
41	Pag (Croatia)	3	0	3	0
42	Pula (Croatia)	9	3	6	0
43	Heraklion (Greece)	15	15	0	0
44	Anopolis (Greece)	11	11	0	0
45	Cephalonia (Greece)	7	4	3	0
46	Athens (Greece)	9	4	5	0
47	Canakkale (Turkey)	11	0	11	0
48	Limassol (Cyprus)	10	0	10	0
49	Bet Dagan (Israel)	8	0	8	0
50	Stellenbosh (SouthAfrica)	19	19	0	0
51	Swat (Pakistan)	12	0	0	12
	TOTAL	524	270	242	12

Tab. 3.2 - 16S rRNA haplotype frequencies of *Candidatus* Erwinia dacicola (htA, htB and htC) among the sampled populations.

Bactrocera oleae *mitochondrial diversity*

A total of 236 individuals of *B. oleae* were analyzed for the mitochondrial region spanning from the 5' region of NADH dehydrogenase gene to the 3' region of 16S. The flies analyzed came from 39 of the 51 locations considered for bacteria, representative of the main geographical areas (Tab. 3.3 and Fig. 3.3). We detected a total of 38 mitochondrial haplotypes (H1-H38) whose 21 were identified in only one sample. The most common haplotype (H3) is shared by 70 individuals coming from 26 populations. The Anopolis population revealed the highest number of haplotypes (N_{HT}= 7), followed by Canakkale (N_{HT}= 6). Considering the populations with five or more samples, haplotype diversity (h) and nucleotide diversity (π) ranged from 0.22 to 1.00 and from 0.0000 to 0.0041, respectively (Tab. 3.3).

The neutrality tests yielded different patterns in the investigated populations. All Tajima's *D* values were non-significant, indicating that these populations are in mutation/drift equilibrium. The Fu's *Fs* test turned out to be much more sensitive to detect departures from neutrality, as it showed significant values for 3 populations, Marsala, Anopolis and Athens (see Tab. 3.3). These populations also showed high values for haplotype diversity. F_S tends to be negative under an excess of recent mutations, and a significant negative value can be taken as an evidence of population growth and/or selection (Fu, 1997).

The statistical parsimony analysis resulted in the network represented in Fig. 3.4; the nested clade design included all the haplotypes across two nesting levels. Considering the highest level, all the Mediterranean samples are distributed in 2-2, 2-3 and 2-4 clades; there were 2 missing or theoretical haplotypes, between H10/H17 and H15. On the other hand all Pakistan and South-Africa haplotypes are comprised in clade 2-1 and they are separated by at least three mutations.

No inference on population history could be made on all clades except for 1-1, 1-4, 1-6 and 2-1 clades, where NCPA suggested a restricted gene flow with isolation by distance.

Clade 1–1 contained a common, widespread haplotype (H3), which could be a presumed ancestor of the Mediterranean populations (Fig. 3.4). Sixteen haplotypes were connected to this ancestor by a single mutational step, most of which are relatively rare with a restricted distribution.

Haplotypes representative of East and Central-East Mediterranean basin are mainly grouped in clade 1-4, whereas the majority of the samples coming from the Iberian peninsula (West- Mediterranean), whose genetic variability is relatively restricted, are included in clade 1-6.

Clade 2-1, as reported above, included Pakistan and South-African populations and they shared the most common South-African haplotype (H19). However the two 1-step clades, embracing the Pakistani (clade 1-2) and the South-African (clade 1-7) populations, were not considered significant by the NCPA.

The null hypothesis of no geographical association was not rejected for all remaining clades which would indicate that populations behave as one panmictic unit.

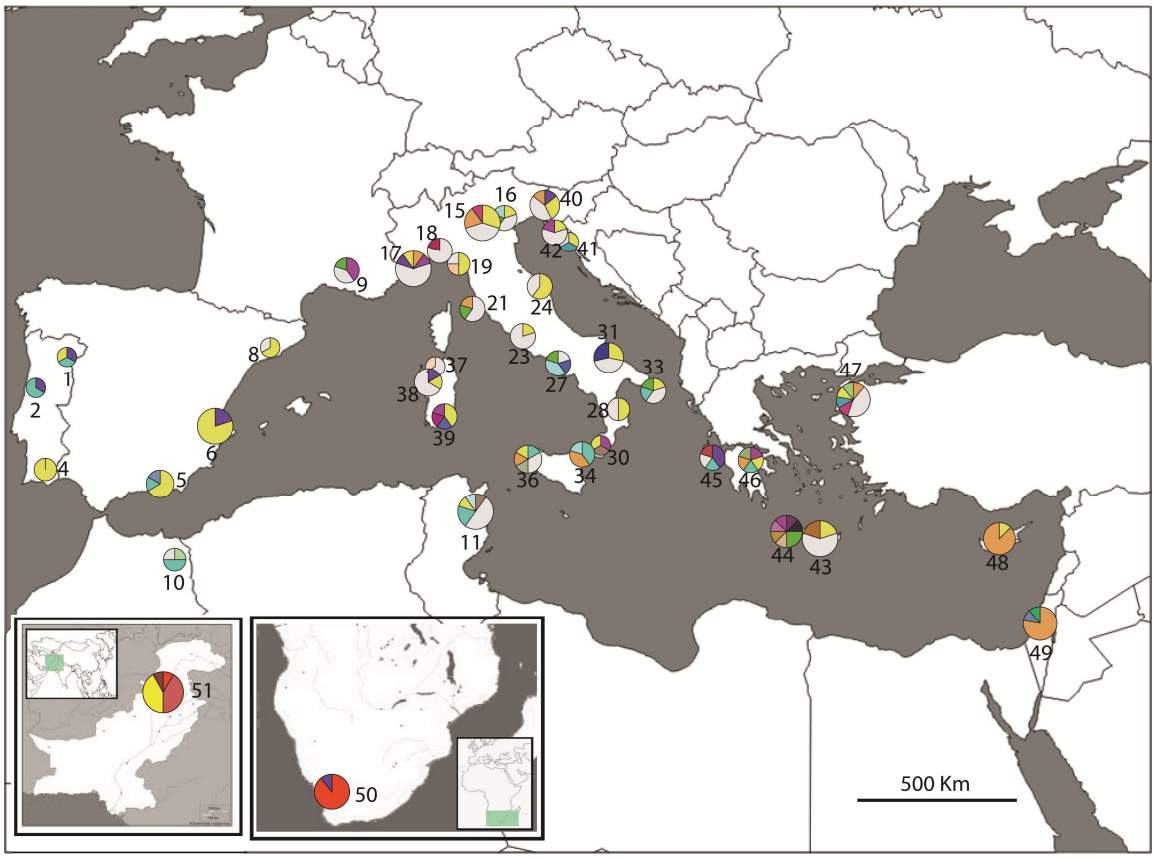

Fig. 3.3 – Distribution map of the 38 *Bactrocera. oleae* mitochondrial haplotypes among the populations. The number indicates the location code, reported in Tab. 3.2. The area of the circles is proportional to the sampling.

	Site of collection	Haplotypes frequencies	*N*	N_{HT}	h	π	k	Fu's F_S
1	Bragança (Portugal)	1 H1, 1 H2, 1 H7	3	3	1.00	0.0017	1.33	-1.2164
2	Lisbõa (Portugal)	1 H1, 2 H7	3	2	0.67	0.0017	1.33	1.0608
4	Tavira (Portugal)	4 H2	4	1	0.00	0.0000	0.00	0
5	Almeria (Spain)	4 H2, 1 H7, 1 H13	6	3	0.60	0.0008	0.66	-0.8584
6	Valencia (Spain)	2 H1, 8 H2	10	2	0.35	0.0004	0.35	0.4167
8	Barcelona (Spain)	2 H2, 1 H3	3	2	0.67	0.0017	1.33	1.0608
9	Avignon (France)	2 H3, 2 H10, 1 H24	5	3	0.80	0.0028	2.20	0.8036
10	Oujda (Morocco)	1 H3, 2 H7, 1 H26	4	3	0.83	0.0019	1.50	-0.2876
11	Chaffar (Tunisia)	1 H2, 5 H3, 2 H7, 1 H8, 1 H11	10	5	0.76	0.0014	1.13	-1.9439
15	Marostica (Italy)	3 H2, 4 H3, 2 H4, 1 H5	10	4	0.77	0.0021	1.66	0.1949
16	Montegrotto Terme (Italy)	1 H2, 2 H3, 1 H6, 1 H8	5	4	0.90	0.0023	1.80	-1.1952
17	Diano Marina (Italy)	1 H1, 1 H2, 6 H3, 1 H4, 1 H15	10	5	0.67	0.0027	2.13	-0.4553
18	Camogli (Italy)	4 H3, 1 H37	5	2	0.40	0.0005	0.40	0.0902
19	La Spezia (Italy)	2 H2, 1 H3, 1 H36	4	3	0.83	0.0025	2.00	0.1335
21	Capoliveri (Italy)	3 H3, 1 H4, 1 H24	5	3	0.70	0.0028	2.20	0.8036
23	Roma (Italy)	1 H2, 4 H3	5	2	0.40	0.0010	0.80	1.0404
24	Ancona (Italy)	3 H2, 2 H3	5	2	0.60	0.0015	1.20	1.6975
27	Portici (Italy)	1 H3, 1 H6, 2 H8, 1 H24	5	4	0.90	0.0036	2.80	-0.4448
28	Rossano Calabro (Italy)	2 H2, 2 H3	4	2	0.67	0.0017	1.33	1.5298
30	Gioia Tauro (Italy)	1 H2, 1 H10, 1 H11	3	3	1.00	0.0034	2.66	-0.3409
31	Bari (Italy)	2 H2, 3 H3, 2 H34	7	3	0.76	0.0020	1.61	0.9039
33	Otranto (Italy)	1 H2, 2 H3, 1 H7, 1 H23	5	4	0.90	0.0018	1.40	-1.6482
34	Messina (Italy)	2 H4, 2 H7, 1 H8	5	3	0.80	0.0020	1.60	0.2764
36	Marsala (Italy)	1 H2, 2 H3, 1 H4, 1 H7, 1 H35	6	5	0.93	0.0023	1.80	-2.3438 *
37	Bancali (Italy)	2 H3, 1 H27	3	2	0.67	0.0008	0.66	0.2006
38	Alghero (Italy)	1 H1, 1 H2, 4 H3	6	3	0.60	0.0018	1.40	0.3811
39	Cagliari (Italy)	2 H2, 1 H10, 1 H21, 1 H22	5	4	0.90	0.0030	2.40	-0.7011
40	Strunjan (Slovenia)	1 H1, 2 H2, 3 H3, 1 H4	7	4	0.81	0.0022	1.71	-0.4281
41	Pag (Croatia)	1 H2, 1 H3, 1 H9	3	3	1.00	0.0034	2.66	-0.3409
42	Pula (Croatia)	1 H2, 3 H3, 1 H10	5	3	0.70	0.0015	1.20	-0.1858
43	Heraklion (Greece)	2 H2, 6 H3, 2 H12	10	3	0.62	0.0013	1.06	0.6028
44	Anopolis (Greece)	1 H3, 1 H10, 2 H24, 1 H28, 1 H29, 1 H30, 1 H31	8	7	0.96	0.0041	3.17	-3.0822 *
45	Cephalonia (Greece)	2 H1, 1 H3, 1 H7, 1 H14	5	4	0.90	0.0031	2.40	-0.7011
46	Athens (Greece)	1 H2, 1 H4, 1 H7, 1 H10, 1 H17	5	5	1.00	0.0031	2.40	-2.6798 *
47	Canakkale (Turkey)	1 H2, 4 H3, 1 H4, 1 H5, 1 H9, 1 H25	9	6	0.83	0.0025	2.00	-2.1758
48	Limassol (Cyprus)	1 H2, 7 H4	8	2	0.25	0.0003	0.25	-1.18197
49	Bet Dagan (Israel)	7 H4, 1 H16, 1 H18	9	3	0.41	0.0007	0.61	-0.5321
50	Stellenbosh (South Africa)	8 H19, 1 H20	9	2	0.22	0.0002	0.22	-0.2634
51	Swat (Pakistan)	1 H19, 5 H32, 5 H33, 1 H38	12	4	0.69	0.0011	0.86	-0.9320
	Total		236					

Tab. 3.3 - Mitochondrial haplotypes found in each population of *Bactrocera oleae* and population parameters. *N*, sample size; N_{HT}, total number of haplotypes for each sampling location; h, gene diversity; π, nucleotide diversity; k, mean number of pairwise differences per sequence; Fu's F_s, Fu's *F*'s statistic. * indicates *P*< 0.02.

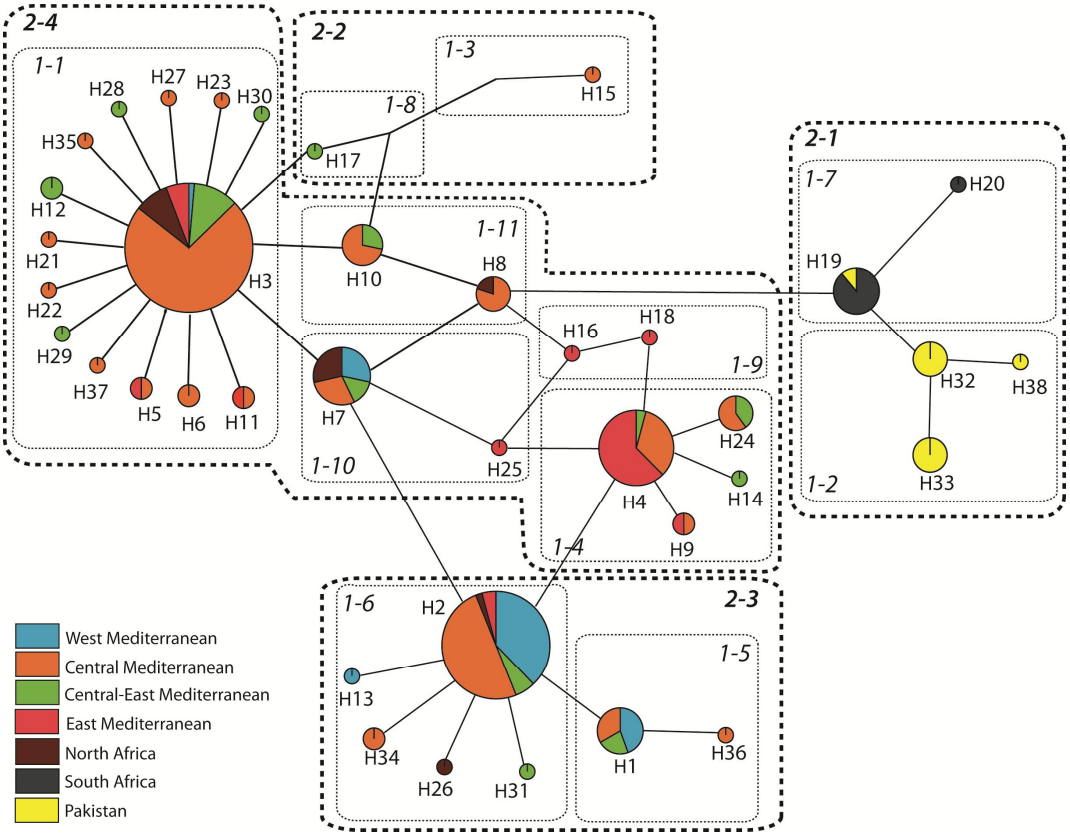

Fig. 3.4 - Nested cladogram of the 38 *Bactrocera oleae* mitochondrial haplotypes; one-step and two-step clades are shown. Sampling region of each haplotype is colour coded as in the legend.

The analysis of molecular variance (AMOVA) was conducted only on Mediterranean populations. Three different hypotheses were tested: populations were grouped on the basis of the SAMOVA results obtained from *E. dacicola*, by geographical areas and on the basis of remarks about host plant (original/recently introduced areas).

The hierarchical AMOVA data did not show significant structure between the two groups when assessed using the two groups found by SAMOVA analysis on *E. dacicola* (Tab. 3.4a). A significant difference among groups ($P < 0.001$) is instead observed when the

groups are those corresponding to the five geographical areas considered, that is Iberian peninsula, Italy and France, Greece, East Mediterranean populations and North-Africa (Tab. 3.4b).

In the last AMOVA test populations were grouped on the basis of the historical presence of wild olive trees on the territory (based on Carrión *et al.*, 2010). This grouping was inconsistent in describing the partitioning of genetic variability as in this case the variation within populations is much greater than the variation between groups or among populations (Tab 3.5c).

Regression analyses of F_{ST} versus geographic distances revealed positive relationship between genetic differentiation and geographic distance within the Mediterranean populations (r Mantel= 0.169, *P*= 0.032).

Structure	Source of variation	d.f.	Variance (%)	Fixation indices	*P-value*
All populations	Among populations	38	33	F_{ST}=0.3300	<0.001
	Within populations	197	67		
(a) Grouping by SAMOVA group	Among groups	1	0.77	F_{CT}=0.0077	0.2952
	Among populations within groups	35	17.12	F_{SC}=0.1725	<0.01
	Within populations	178	82.11	F_{ST}=0.1788	<0.001
(b) Grouping by geographical region	Among groups	4	18.08	F_{CT}=0.1807	<0.001
	Among populations within groups	32	4.48	F_{SC}=0.0546	<0.05
	Within populations	178	77.44	F_{ST}=0.2255	<0.001
(c) Wild-cultivated	Among groups	1	3.52	F_{CT}=0.0351	0.1065
	Among populations within groups	37	30.70	F_{SC}=0.3184	<0.001
	Within populations	197	65.79	F_{ST}=0.3421	<0.001

Tab. 3.4 - Analysis of molecular variance (AMOVA) of *Bactrocera oleae* populations, divided according to the hypotheses discussed in the text.

Testing the association between Erwinia dacicola and Bactrocera oleae haplotypes

Strict vertical transmission was tested checking the match between olive fly symbiont and its host haplotypes. The overlay of the three *E. dacicola* haplotypes over the host mtDNA data (Fig. 3.5) shows the presence of an association (Fisher's exact test, df=7, P <0.01). The analysis was conducted only on the interior mitochondrial haplotypes (H1, H2, H3, H4, H7, H8, H10, H19) because the external haplotypes are rare and thence forcedly correlated with one of the three *E. dacicola* haplotypes.

The most striking association concerned H4 and H8: almost all the samples having H4 (23/24) and all samples with H8 as mitochondrial haplotype shared htB as symbiont haplotype. It's noteworthy that H4, as well as the majority of the haplotypes connected to it with a single mutational step, is mainly spread in Eastern Mediterranean populations, whereas H8 is restricted in the central Mediterranean area and in Northern Africa. On the other hand samples sharing H1 and H19 are mainly found in the West-Central Mediterranean and South Africa respectively, where htA is the dominant haplotype. (Fig. 3.1 and 3.5). Excluding from the analysis these haplotypes with the highest proportion of association with a single symbiont haplotype, the Fisher exact test is no more significant.

Moreover we compared the proportion of symbiont lineages in the most common insect haplotypes with the proportion of symbiont lineages in all the derived insect haplotypes separated by only one mutational step and included in the clade identified by the nested clade analysis.(Fig. 3.4).

We found that the proportion of symbiont haplotypes did not vary between the most common haplotypes and directly connected haplotypes considered together (G-test not significant for all couples tested).

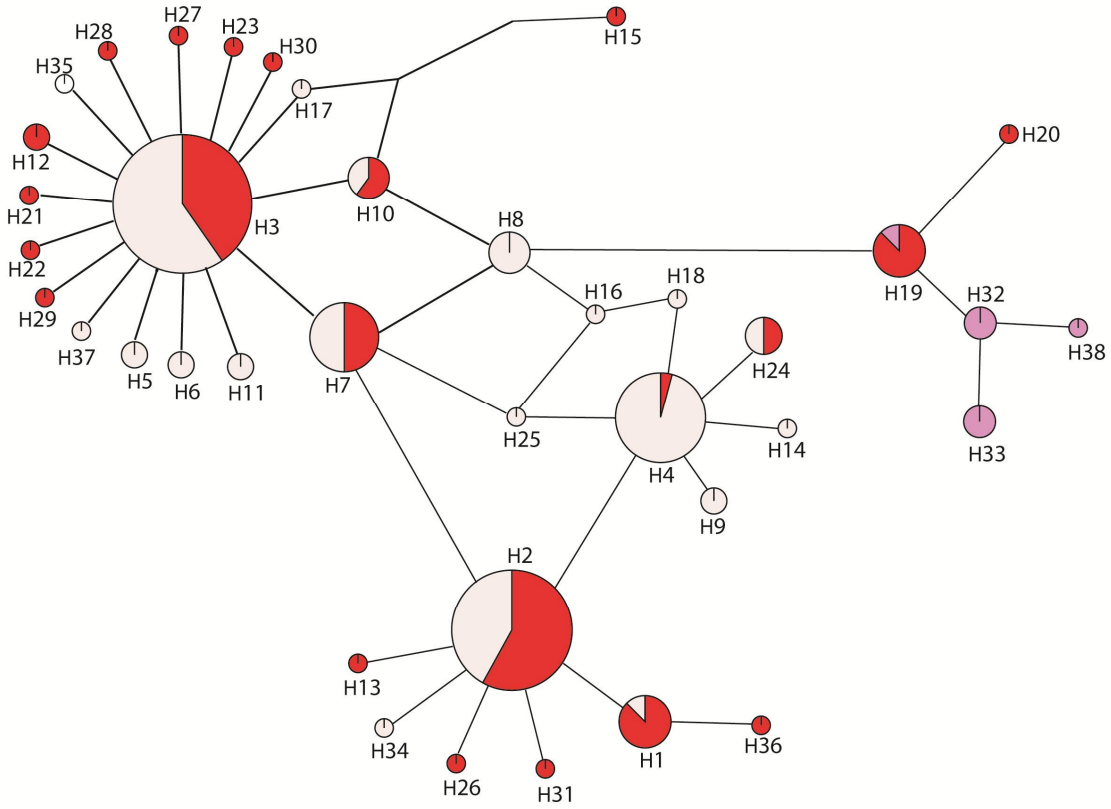

Fig. 3.5- . Haplotype network of *Bactrocera oleae* realized by TCS 1.21 (Clement *et al.*, 2000). Each haplotype is represented by a circle, with the area of the circle proportional to its frequency. Inside each circle is reported the proportion of *Erwinia dacicola* lineage: htA (red), htB (white) and htC (pink).

Discussion

In the present work we explored the genetic diversity of the olive fly symbiont *Erwinia dacicola* together with the genetic variability of the mitochondrial region of its insect host *Bactrocera oleae*.

Our results showed the presence of three different haplotypes of the olive fly symbiont: the Mediterranean and the South-African olive flies populations harbor htA and htB while htC is reported only in the Pakistani population. htA and htB where previously reported in Italian olive fly populations, where they seemed to be not randomly distributed (Savio *et al.*, submitted). The combined occurrence of different *E. dacicola* haplotypes in a single fly seems to be extremely rare or not possible at all, as shown also by Savio *et al.* (submitted), due to unknown intimate mechanisms that should be investigated.

When exploring the symbiont haplotypes distribution in the Mediterranean basin we found a strong correlation with the geography, as confirmed by the "a- posteriori" approach of the analysis of the molecular variance (SAMOVA): Western populations are mainly provided with htA while the Eastern populations harbor htB.

The two populations sampled in Crete are the most striking exception to this distribution. However the presence of htA in Crete could be explained by geographical isolation, thus the population could not experienced external gene flow. This hypothesis is also supported by the presence of unique mitochondrial haplotypes in Crete (H12, H28-31).

Considering the whole Mediterranean area a very high gene flow is present, with widespread haplotypes showing a star-shaped pattern of differentiation, suggesting a recent expansion with some local events of isolation by distance due to restricted gene flow. Nested clade analysis found little evidence of a geographical structure as the high nested clades are not significant, except for Pakistani haplotypes for which there is restricted gene flow with isolation by distance .(Fig. 3.4). Similar data were highlighted by previous phylogeographic studies (Nardi *et al*., 2005, 2010).

Despite the high level of gene flow across the Mediterranean, microsatellite analysis of Augustinos *et al.* (2005) supports the notion of a differentiation of three subpopulations and a gradual decrease of heterozygosity from the Eastern to the Western part of the Mediterranean.

Interestingly, this distribution partially mirrors the east-west subdivision found in our study among *Erwinia dacicola* haplotypes although AMOVA does not support this hypothesis in our insect mitochondrial data.

It is interesting to note that the insect haplotypes showing strict/significant association with symbiont haplotypes are geographically isolated. It could be hypothesized that this association is not the result of an effective linkage between the olive fly and its symbiont. It could be instead a consequence of the insect haplotypes typical of a given area, loosing and acquiring the predominant symbiont haplotype of the region. On the other hand the similar proportions of symbiont haplotypes found between ancestral and derived insect haplotypes confirmed the predominant presence of vertical transmission.

Interestingly the east-west differentiation genetic pattern in the Mediterranean basin found in *Erwinia dacicola* and *B. oleae* has been also observed in wild olives trees (Lumaret *et*

al., 2004; Besnard *et al.*, 2007), which was explained by a re-colonization from different glacial refugia (Besnard *et al.*, 2002). This hypothesis could be tested also for the olive fly; the combined analyses on the genetic variability of the insect host and its symbiont could help to shade light on the debated olive fly trade routes of colonization.

Acknowledgements

We thank Y. Ben-Dov, A. Bozzolo, V. Caleca, E. Cargnus, S. Deliperi, 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, N. Mori, 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, D. Zovi for their precious help in collecting samples.

Supplementary table

Region	Country		Locality	Coordinates	Date of collection
Mediterranean area	Portugal	1	Bragança	41°48'N 06°45'W	10/2007
		2	Lisboa	38°31'N 08°55'W	10/2007
		3	Serra de Leomil	40°57'N 07°40'W	10/2007
		4	Tavira	37°07'N 7°38'W	10/2009
	Spain	5	Almeria	36°50'N -2°28'W	10/2007
		6	Valencia	39°28'N 0°23'E	09/2008
		7	Tarragona	41°06'N 1°13'E	11/2009
		8	Barcelona	41°24'N 2°09'E	10/2010
	France	9	Avignon	43°55'N 4°48'E	11/2010
	Morocco	10	Oujda	34°37'N 1°56'E	12/2009
	Tunisia	11	Chaffar	34°34'N 10°33'E	01/2009
	Italy	12	Campione	45°58'N 08°58'E	10/2009
		13	Bardolino	45°32'N 10°43'E	12/2008
		14	Dueville	45°38'N 11°32'E	10/2009
		15	Marostica	45°44'N 11°39'E	10/2007
		16	Montegrotto Terme	45°19'N 11°47'E	12/2007
		17	Diano Marina	43°54'N 8°04'E	11/2006
		18	Camogli	44°20'N 9°09'E	11/2010
		19	La Spezia	44°04'N 10°02'E	11/2010
		20	Firenze	43°46'N 11°15'E	12/2007
		21	Capoliveri	42°44'N 10°22'E	08/2009
		22	Arezzo	43°28'N 11°51'E	11/2009
		23	Roma	41°53'N 12°29'E	12/2009
		24	Ancona	43°37'N 13°31'E	09/2009
		25	Moscufo	42°25'N 14°03'E	11/2009
		26	Capua	41°06'N 14°13'E	09/2009
		27	Portici	40°48'N 14°20'E	09/2009
		28	Rossano Calabro	39°34'N 16°38'E	09/2009
		29	Rende	39°19'N 16°10'E	10/2009
		30	Gioia Tauro	38°25'N 15°54'E	11/2009
		31	Bari	41°07'N 16°51'E	11/2002
		32	Castellana Grotte	40°53'N 17°09'E	10/2009
		33	Otranto	40°08'N 18°29'E	10/2009
		34	Messina	38°11'N 15°33'E	08/2007
		35	Siracusa	37°03'N 15°17'E	11/2008
		36	Marsala	37°47'N 12°26'E	07/2008
		37	Bancali	40°44'N 8°27'E	12/2009
		38	Alghero	40°33'N 8°19'E	10/2008

		39	Cagliari	39°12'N 9°06'E	09/2009
	Slovenia	40	Strunjan	45°31'N 13°34'E	09/2008
	Croatia	41	Pag	44°29'N 14°57'E	08/2008
		42	Pula	44°52'N 13°50'E	10/2008
	Greece	43	Heraklion	35°19'N 25°08'E	03/2009
		44	Anopolis	35°12'N 24°06'E	04/2010
		45	Cephalonia	38°10'N 20°34'E	10/2007
		46	Athens	37°56'N 23°01'E	12/2007
	Turkey	47	Canakkale	40°09'N 26°24'E	10/2008
	Cyprus	48	Limassol	32°57'N 34°45'E	11/2008
	Israel	49	Bet Dagan	31°59'N 34°49'E	12/2007
South-Africa	South-Africa	50	Stellenbosh	33°56'S 18°51'E	09/2009-07/2010
Pakistan	Pakistan	51	Swat	35°21'N 72°11'E	10/2010

Chapter 4

Bacterial communities associated with *Rhagoletis completa* Cresson and *Rhagoletis cerasi* (Linnaeus) (Diptera: Tephritidae)

Manuscript in preparation as:

Savio C., Martinez-Sañudo I., Mazzon L. and Girolami V. Bacterial communities associated with *Rhagoletis completa* Cresson and *Rhagoletis cerasi* (Linnaeus) (Diptera: Tephritidae).

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

Introduction

Insects may have various pathogenic and non-pathogenic associations with microorganisms (Buchner, 1965). The microorganism are present in different part of the insects body and sometime are present in specific organs that are adapted for the microorganism life.

Most of the data obtained on insect gut bacterial composition is derived from traditional culturable methods. However the availability in the recent past of culture-independent tools, such as PCR, provides an opportunity to detect and classify microorganisms that cannot be identified using traditional methods (Ward *et al.*, 1990). The use of PCR has indeed revealed the presence of symbiotic and the diversity of associated bacteria of a great number of insects (Munson *et al*., 1991; Baumann, 2005; Moran *et al*., 2005; Vasanthakumar *et al*., 2006; Mazzon *et al*., 2010).

The presence of a relationships between bacteria and Tephritid fruit flies has been well documented; some tephritid species harbor symbiotic bacteria, as in the case of olive fly *Bactrocera oleae* (Capuzzo *et al*., 2005) and *Tephritis spp.* (Mazzon *et al*., 2008), while other species live in associations with bacterial communities (Howard *et al*., 1985; Drew & Lloyd, 1987 Daser & Brandl, 1992; Marchini *et al*., 2002; Thaochan *et al*., 2010). One of the most fascinating aspects of these relationships are how the bacterial species communities are structured.

Rhagoletis completa, the walnut husk fly, and *R. cerasi*, the cherry fruit fly, are two oligophagous fruit flies pest attacking respectively *Juglans spp*. and *Prunus spp/ Lonicera spp.*. *R. completa* is a pest originating from North-America, detected in Europe for the first time in 1991 (Duso, 1991), while *R. cerasi* is a fruit fly native to Europe. As all the Tephritidae, they are provided with a specialized organ in the head, called “oesophageal bulb”, connected to pharynx, whose major function seems to be the housing of bacteria (Girolami, 1973, 1983; Ratner & Stoffolano, 1984). Nowadays there is little work exploring the oesophageal bulb microbiota in *R.completa* (Howard *et al*., 1985; Tsiropoulos, 1976) and only a single information is available for *R. cerasi* (Blümel & Masalmeh, 1989). These authors report the presence of bacteria in the oesophageal bulb of *R. completa* as well as in the rest of developmental stages. However, in *R. cerasi* bacteria has been found only in

adults and pupae and not in the preimaginal stages. It should be emphasized that these few microbial studies have depended mainly on traditional culture-based techniques.
The study of bacteria colonizing these two Tephritids, together with the knowledge of their physiological role would expand the possibility of biological control of insects pest (Marchini *et al*., 2002).

Here we report the characterization of the *R. completa* and *R. cerasi* oesophageal bulb bacteria, using traditional culturing and culture independent methods. Furthermore we discuss, as preliminary data, the permanency of these associations investigating the presence of bacteria in eggs, larvae and pupa instars in *R. cerasi*.

Materials and methods

Rhagoletis cerasi*: host collection and immature stages handling*

Host collection and handling

Cherries fruits were collected in 2009 and 2010 from three *Prunus avium* orchards located in North-East Italy: Ponte San Nicolò (Padova), Montegrotto Terme (Padova) and Schio (Vicenza). Ponte San Nicolò orchard is constituted of heterogeneous trees, coming from ungrafted seeds. Field cherries were transported to the laboratory, avoiding any contact with wetted bags and were visually inspected for insect attacks under a stereomicroscope.

Eggs collection

Cherries were checked under steromicroscope to find eggs. Usually *R. cerasi* deposits only one egg into each fruit. When a egg was detected, a slice of the skin upper it was cut out under a laminar flow hood, without removing the pulp around the egg. Eggs were picked up with some fruit pulp using sterile tweezers, avoiding any contact with the external cherry skin, based on the hypothesis that bacteria are numerous on the fruit surface and absent in the pulp. Each egg was placed on a PCA plate agar and tweezers were streaked separately for checking accidentally contaminations for at least three days.

Larvae and pupae collection

During the eggs collection, larvae (1st and 2nd-3rd instars) were also collected. Larvae were gently picked up from the fruit pulp under a stereomicroscope using sterile tweezers under a laminar flow hood. First instar larvae are easily distinguishable because only one pair of spiracles are present.

In order to obtain mature larvae and pupae, tree branches carrying mature cherries were transported to the laboratory and kept suspended over boxes containing sterile paper or vermiculite to allow pupation of mature larvae.

New formed pupae found in the boxes were surface sterilized using sodium hypochlorite 1% p.a. for 10 minutes, rinsed twice in sterile water and put on vermiculite. This treatment was used to reduce the presence of bacteria and fungi on the surface of the samples.

Part of the insects (larvae and pupae) were immediately put on 60 mm diameter Petri dishes containing PCA agar, while the remaining were put in sterile vials containing GET buffer (50 mM glucose, 10mM EDTA, 25 mM Tris-HCl, pH 8) and homogenated with sterile pestles. The whole body of these latter samples were used to extract total DNA from microbial content.

Eggs, larvae and pupae were left in the Petri dishes for about one week, where larvae could feed on the agar substrate. If after this period no bacterial colonies emerged on the plates, eggs/larvae were aseptically picked up from the substrate and prepared for the molecular analysis, while pupae were squeezed on PCA plates. If bacterial colonies appeared on the plates, some morphotype were selected, and brought into pure culture.

R. cerasi *and* R. completa*: adults stages collections and handling*

Flies (*R. completa* and *R. cerasi*) were collected from the three above mentioned orchards where immature stages were collected and three additional ones throughout North Italy, using yellow sticky traps from 2007 to 2010. Moreover, in order to asses the stability of the oesophageal bulb microflora in both *Rhagoletis* species, flies coming from Ponte San Nicolò and Schio flies were sampled twice, once in mid- May and once in mid-June 2010. 16 flies were captured during each sampling, except for Mid-May in Schio, where we could not captured more than 10 flies, due to low infestation.

Sticky traps were checked daily and the living flies were removed and immediately processed. Flies were aseptically dissected in physiological saline solution removing the oesophageal bulb and its content was analyzed by cultivation-dependent and independent methods. In the first case the content was streaked individually on plate count agar (PCA); after incubation at room temperature for 72 hours individual bacterial colonies of some morphotype were selected, and brought into pure culture. In the case of cultivation-independent method, oesophageal bulb was homogenated in sterile eppendorf tubes containing GET buffer and preserved at -20 °C until the DNA extraction.

In some cases the dissected oesophageal bulb was prepared to be viewed under an optical microscope to check the presence of bacterial content. The glass cover slip of the sample was then streaked on PCA plates, and the slide was washed with 0.5 ml of sterile water and used as template to extract DNA.

R. cerasi *rearing in microbiologically controlled conditions*

Part of *R. cerasi* pupae pupating in the sterile paper or vermiculite were surface-sterilized immediately after pupation by a 10 min immersion in 1% sodium hypochlorite solution. Then they were rinsed twice, air-dried in sterile conditions and kept under siliceous sand during the whole winter season, until emergence of adults. After the winter, before emerging, flies were surface sterilized again as described above, rinsed twice with sterile water and kept into sterile vials. When adult emerged, they were immediately aseptically transferred into larger vials containing a layer of plate count agar (PCA) on the bottom as a sterility check and sealed with a sterilized transparent gas-permeable cellulose membrane for dialysis (Sigma-Aldrich). A drop of sterile glucose solution was placed on the internal side of the membrane to allow insect feeding. The drop of solution was rewetted, whenever necessary, by spraying water onto the cellular membrane. For more details on insect rearing see Mazzon *et al.* (2008). After one week specimens were aseptically dissected and the oesophageal bulbs were analysed as described above.

DNA extraction, amplification and data analysis

DNA was extracted from bacterial colonies and homogenated flies samples, using the extraction protocol described in Palmano *et al.* (2000).

The 16S rRNA was partially amplified using two couples of universal bacterial primers fd1-rp1 (Weisburg *et al.*, 1991) or 63f-1389r (Osborn *et al*., 2000) as described in Savio *et al.*, (submitted). Sequencing was provided by BMR Genomics service (Padova, Italy).

Sequences were aligned with MEGA 4.0 (Tamura *et al.*, 2007) and their identity was assessed with the Sequence Match function implemented in Ribosomal Database Project II (RDP-II) (Cole *et al*., 2009).

Prior to further analyses each sequence was checked as possible chimera with Bellerophon (Huber *et al.*, 2004) using the Huber–Hugenholtz correction. All chimeric sequences and poor-quality sequences were excluded from subsequent analysis. The software package DOTUR was used to assign sequences to operational taxonomic units (OTUs) for the bacterial identities found in the adult stages. This program accurately assigns sequences to OTUs based on sequence data by using values that are less than the cutoff level (Rani *et al.*, 2009). 16S rRNA sequences were grouped into same OTUs by using 99% identity threshold.

Results and discussion

Rhagoletis cerasi *immature stages*

The analysis conducted on *R. cerasi* eggs showed that in 115/116 cases they did not developed bacterial colonies on the surface of agar plates (Tab. 4.1 and Fig. 4.1). Hence almost all eggs do not contain bacteria on the surface, in contact with the pulp or around the wounded pulp.

Furthermore the 25% of the collected eggs eclosed on agar plates and galleries created by 1st instar larvae on agar were devoid of bacterial colonies. The result was confirmed by the culture-independent analysis, as the PCR showed negative results for 21/21 tested eggs. Since no disinfectant has been used, these results may be considered as a strong suggestion that no bacteria are transmitted by the mother to eggs.

The 88% of 1st instar larvae collected inside the cherry showed no microorganisms when transferred on agar plates, as the galleries did not develop bacterial colonies (the data are available only for 2009) (Fig. 4.2). The amount of bacteria increased in 2nd- 3rd instars

(32.4%) and mature (39.1%) larvae. The pupae contained similar proportions of bacteria (37.5%). There were no differences among the different sampled localities (data not shown) and results are similar in the two sampled years: Fisher's exact test 2009-2010: eggs, P=0.573; 2^{nd}- 3^{rd} instars P= 0.683; mature larvae P= 0.411.

It must be noticed that immature larvae has been collected directly inside the fruit pulp while mature larvae were picked up from paper/vermiculite in laboratory conditions. Hence the presence of bacteria in these latter samples could be underestimated, because in natural conditions a possible source of bacteria acquirement could come from drops of rain and dewdrops on the surface of the fruits. In facts larvae perforate the fruit skin with their hooks, probably to breath. The percentage of infected pupae could be underestimated too, since they were twice sterilized to exclude external contaminants.

When analysing the homogenated insects by PCR with the primer set 63f-1389r in most of the cases we found *Wolbachia sp*., an intracellular bacterium whose presence in *R. cerasi* is well known, causing cytoplasmically incompatibility (Riegler & Stauffer, 2002). Thus the analysis were conducted only with the universal primers fd1-rp1. Sequencing of bacteria coming from both culturable and unculturable methods revealed the prevailing presence of Enterobacteriaceae species in the midgut of larvae and within the pupae. The most representative genera, from a RDP match analysis, were found to be *Erwinia*, *Rhanella* and *Klebsiella*. No unculturable bacteria were identified, a part *Wolbachia,* as both methods gave the same result in terms of genera diversity.

To summarize, from the present preliminary analyses it seems that the female does not transmit bacteria to the eggs and bacteria are likely to be acquired from larvae and mature larvae when they reach the ground to pupate. This fact opens new perspectives as this possibility has never been taken into account in others Trypetinae. Interestingly, as reported above, Blümel & Masalmeh in 1989 reported, using traditional methods, the absence of bacteria in eggs and larvae of the cherry fruit fly.

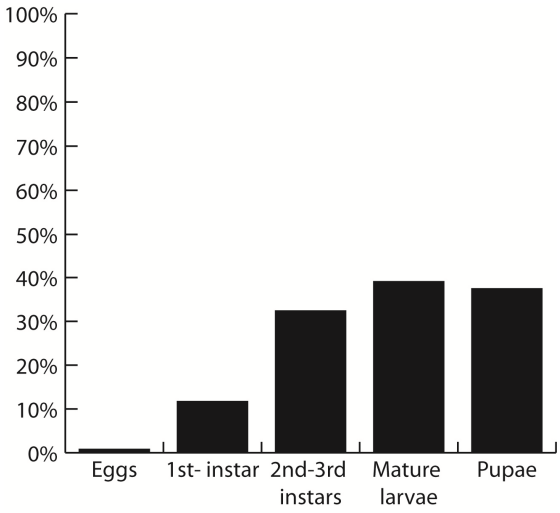

Fig. 4.1 - Percentage of infected *Rhagoletis cerasi* in different life stages.

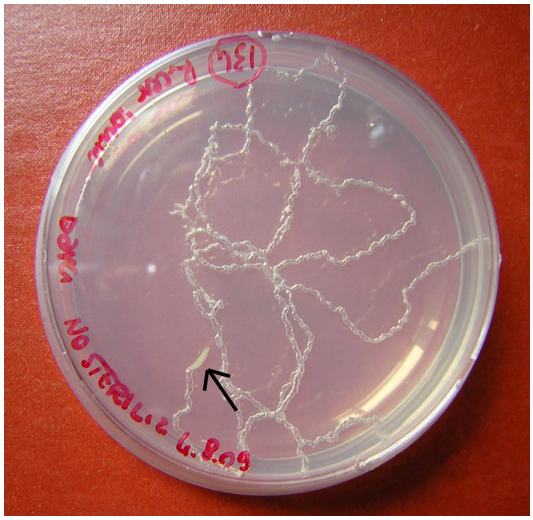

Fig. 4.2 – Galleries created on agar substrate by a 2nd instar larva of *Rhagoletis cerasi*. The black arrow indicates the larva position.

	N 2009	Infected 2009	*N* 2010	Infected 2010	*N* tot.	Tot. infected
Eggs	28	0	88	1	116	1
1st- instar	33	4	1	0	34	4
2nd-3rd instars	24	7	13	5	37	12
Mature larvae[1]	47	16	22	11	69	27
Pupae[2]	16	6	0	0	16	6

Tab. 4.1 – Infection status in preimaginal stages of *Rhagoletis cerasi* collected in 2009 and 2010. *N*= sample size. [1]-Picked up from vermiculite or sterile paper in laboratory conditions. [2]- Surface sterilized.

Wild flies

All the wild captured *R. cerasi* and *R. completa* flies harboured bacteria in their oesophageal bulbs. Since sequences coming from the direct PCR of the oesophageal content and from the isolated colonies on plates gave, in most of the cases, the same results, they were considered together in the following analyses. As a consequence we can also exclude the presence of non-culturable bacteria.

In the tables 4.2 and 4.3 we reported separately the results obtained from the match analysis using only "type" sequences deposited in RDP database and using also "non type" sequences, as they gave different results.

Rhagoletis completa

The analysis of the sequences using DOTUR showed that when an OTU was defined as a group of sequences with 1% difference, the composite sequence collection contained 7 OTUs (Tab. 4.2). Allowing higher levels of variation, five (2% difference) or three (3% difference) OTUs were predicted (data not shown).

Considering the matching with RDP "Type" sequences the 92% bacterial strains detected within oesophageal bulbs of wild *R. completa* is affiliated with 2 OTUs (OTU 1 and OTU 2), whose identity refers to *Raoultella planticola* (39/103), *R. ornithinolityca* (15/103) and *Enterobacter asburiae* (40/103) (Tab. 4.2).

On the other hand when conducting the matching with "non type" RDP sequences we obtained a dominance of bacteria affiliated to *Raoultella*/*Klebsiella* genera. In this case the bacterial sequences previously defined as *Enterobacter asburiae* are approximately referred to *Klebsiella oxytoca*. This result highlight how the assignation of bacterial species could be controversial, even with the available molecular techiques.

These data indicate that the bacteria harboured in the *R. completa* oesophageal bulb share a high similarity among them. Further studies will be useful to clarify how the bacteria are chosen and transmitted during the different fly life stages.

Rhagoletis cerasi

In the cherry fruit fly the bacterial content of the oesophageal bulb was found to correspond to 6 OTUs at the 99% of similarity (Tab. 4.3).

Considering the first OTU, 20/109 sequences corresponded to *Enterobacter asburiae*, the same bacterial strain also found in *R. completa* microflora, where it represented about the 40% of the totality of the bacteria. In the same OTU are included *Raoultella* and *Erwinia* genera; the latter incorporated *E. persicina* and *E. rhapontici*. The entirety of these groups constituted around the 50% of the bacterial isolates in *R. cerasi* oesophageal bulb.

The second OTU includes *Serratia* genus and, considering "non-type" matching, *Rhanella* genus. *Serratia fonticola* and *Rhanella acquatilis* are usually found in fresh water, hence their presence could suggest that *R. cerasi* could acquire them feeding on dewdrop from the surface of plants. Therefore bacterial acquirement is not strictly related to the fruit host.

Preliminary studies conducted in 2010 indicate that the microflora composition inside the oesophageal bulbs changes significantly over the time in Ponte San Nicolò *R. cerasi* population. In fact the presence of *Rhanella spp*. decreased significantly over the season, that is flies emerged in mid-May harboured *Rhanella spp*. while those emerged one month later did not. On the contrary, the presence of *Klebsiella* increased significantly (Fig. 4.3 and Tab. 4.4). Even if the population sampled in Schio showed some differences in the bacterial genera composition they were not significant (Fig. 4.3 and Tab. 4.4).

Rhagoletis cerasi *reared in microbialogically controlled conditions*

We reared 13 *R. cerasi* flies in microbial controlled conditions from surface sterilized pupae and only in one case we observed bacteria on the PCA on the bottom of the vials. 7/13 flies were aseptically dissected and PCR on the oesophageal bulb content confirmed the absence of bacteria. *Wolbachia* was instead identified in all samples.

OTU	No. Seq	Closest RDP match with Type sequences	Closest RDP match with Non Type sequences
1	54	(39) 99-96% *Raoultella planticola*; AF129443 (15) 100%-97 *Raoultella ornithinolytica*; U78182	(39) 100-98,1% *Raoultella planticola;* X93215 (15) 100-98% *Raoultella ornithinolytica*; AF129441
2	41	(40) 99-96% *Enterobacter asburiae*; JCM6051; (1) 98% *Klebsiella pneumoniae* ; Y17657	(38) 100-97,4% *Klebsiella oxytoca*; AJ871855 (2) 99.1% *Enterobacter aerogenes*; FJ976587 (1) 99.1% *Klebsiella pneumoniae*; DQ470485
3	3	(1) 97,4 % *Citrobacter murliniae*; AF025369 (1) 96.5% *Raoultella terrigena* ; Y17658 (1) 98,8 % *Buttiauxella agrestis* ; AJ233400	(1) 98.8 % *Enterobacter aerogenes*; GQ337696 (1) 98.8 % *Buttiauxella agrestis*; AJ293684 (1) 97.9 % *Raoultella terrigena*; AY292873
4	2	(2) 99,8-98,7% *Enterobacter cowanii*; AJ508303	(2) 99,8% *Enterobacter cowanii*; EU629163
5	1	100% *Serratia nematodiphila*; EU036987	100% *Serratia marcescens*; AJ297950
6	1	89.2% *Serratia entomophila*; AJ233427	99.5 % *Euscelidius variegatus*; Z14096
7	1	100 %*Providencia rettgeri*; AM040492	100% *Providencia rettgeri*; AM040492

Tab. 4.2 – Bacterial species identified in the oesophageal bulb of *Rhagoletis completa.* In brackets number of specimens matching with the indicated species from RDP database.

OTU	No. Seq	Closest RDP match with Type sequences	Closest RDP match with non type sequences
1	73	(20) 99-96% *Enterobacter asburiae*; JCM6051 (10) 100-97% *Enterobacter amnigenus*; AB004749 (2) 94.1% *Enterobacter amnigenus*; AB004749 (1) 97.8% *Enterobacter cloacae*; Z96079 (1) 97.0% *Enterobacter hormaechei*; AJ508302 (1) 96.1% *Enterobacter aerogenes*; AB004750 (4) 97.9 % *Enterobacter ludwigii*; AJ853891 (3) 100-98.9% *Buttiauxella agrestis*; AJ233400 (2) 99-97% *Kluyvera intermedia*; AF310217 (7) 98-94.5% *Erwinia persicina*; U80205 (1) 98.9% *Erwinia rhapontici*; AJ233417 (1) 98% *Klebsiella pneumoniae*; AF130982 (3) 98.9-95.9%*Raoultella terrigena*; Y17658 (7) 95.1-92.2% *Raoultella planticola*; F129443 (7) 100-95% *Klebsiella oxytoca*; AF129440 (3)100-96% *Raoultella ornithinolytica*;U78182 (1) 97.1-96.8% *Citrobacter freundii* (1) 100% *Citrobacter gillenii*; AF025367	(1) 100-98.9% *Enterobacter asburiae*; EU221358 (1) 100% *Enterobacter aerogenes*; FJ976587 (11) 100-97%*Enterobacter amnigenus* EU275356 (4) 100% *Enterobacter cloacae*; AY787819 (4) 99.5-98% *Enterobacter ludwigii*; GQ915080 (1) 98.7% *Enterobacter hormaechei*; GQ260075 (1) 97.4% *Enterobacter sp.* EF489444 (3) 100% *Buttiauxella agrestis*; AJ293684 (1) 99.4% *Kluyvera sp.*; GQ915083 (2) 99.7-97% *Kluyvera intermedia*; AB004747 (2) 98.1-97.3% *Erwinia persicina*; AJ001190 (1) 100% *Erwinia rhapontici*; U80206 (1) 98.9% *Klebsiella pneumoniae*; AY291290 (2) 98.2-97.3% *Raoultella terrigena*; Y292874 (1)98.9% *Raoultella terrigena*; AF129442 (6) 94.4-92% *Raoultella planticola*; AF129444 (20)100-97.5% *Klebsiella oxytoca*; AJ871855 (4) 99% *Klebsiella oxytoca*; AJ871857 (3)100-96% *Raoultella ornithinolytica*; Y17662 (1) 97.1% *Klebsiella sp.*; EU870375 (1) 97.1% *Citrobacter freundii*; AM184281 (1) 100% *Citrobacter freundii*; AB548828 (1) 100% Endosym. of *S.s levis*; FJ626247 (1) 96.3% *Leclercia sp.*; GQ856079 (1)100% *Endophytic bacterium*; FJ205671
2	28	(3) 100-98.9% *Serratia liquefaciens*; AJ306725 (2) 95.7-94.6% *Serratia grimesii*; AJ233430 (5) 98.9% *Serratia fonticola*; AJ233429 (15)96.0 -90.9% *Serratia fonticola*; AJ233429 (3)98.9- 95.2% *Yersinia kristensenii*;AF366381	(3) 100-99% *Serratia liquefaciens*; DQ123840 (2) 100-98.9% *Serratia grimesii*; DQ086780 (1)99% *Serratia fonticola*; AJ279002 (4) 100% *Serratia fonticola*; AY236502 (2) 93.7-91.4% *Serratia fonticola*; AY236502 (1) 98.9% *Rahnella aquatilis*; FJ811856 (2) 98.9-98% *Rahnella aquatilis*; AY253920 (1) 100% *Rahnella genomosp.*; GQ148969 (3) 100%-99.2-94.0 *Rahnella sp.*; U88435 (7) 100% *Rahnella sp.*; U88434 (2) 98.9-95% *Yersinia kristensenii*; EF179126
3	4	(4)99-98% *Pantoea agglomerans*; AJ233423	(2)100% *Pantoea agglomerans*; FJ357810 (2)100-99% *Pantoea agglomerans*; AJ583011
4	2	(2) 98.9% *Pseudomonas monteilii*; AF064458	(2)100-99% *Pseudomonas putida*; AY952321
5	1	(1) 97.8% *Cedecea davisae*; AF493976	(1) 95.8% bacterium; FJ816030
6	1	(1)100% *Pseudomonas libanensis*; AF057645	(1)100% *Pseudomonas libanensis*; DQ071559

Tab. 4.3 – Bacterial species identified in the oesophageal bulb of *Rhagoletis completa.* In brackets number of specimens matching with the indicated species from RDP database.

	Ponte San Nicolò		Schio	
	Mid-May	Mid-June	Mid-May	Mid-June
Citrobacter	1	0	1	0
Enterobacter	6	4	1	3
Erwinia	0	0	0	1
Hafnia	1	0	0	0
Klebsiella	2	11	2	7
Kluyvera	0	0	2	2
Raultella	2	1	0	1
Rhanella	4	0	4	2
Total	16	16	10	16

Tab. 4.4- Composition of the microflora found in *R. cerasi* oesophageal bulbs of Ponte San Nicolò and Schio during two sampling collections in 2010.

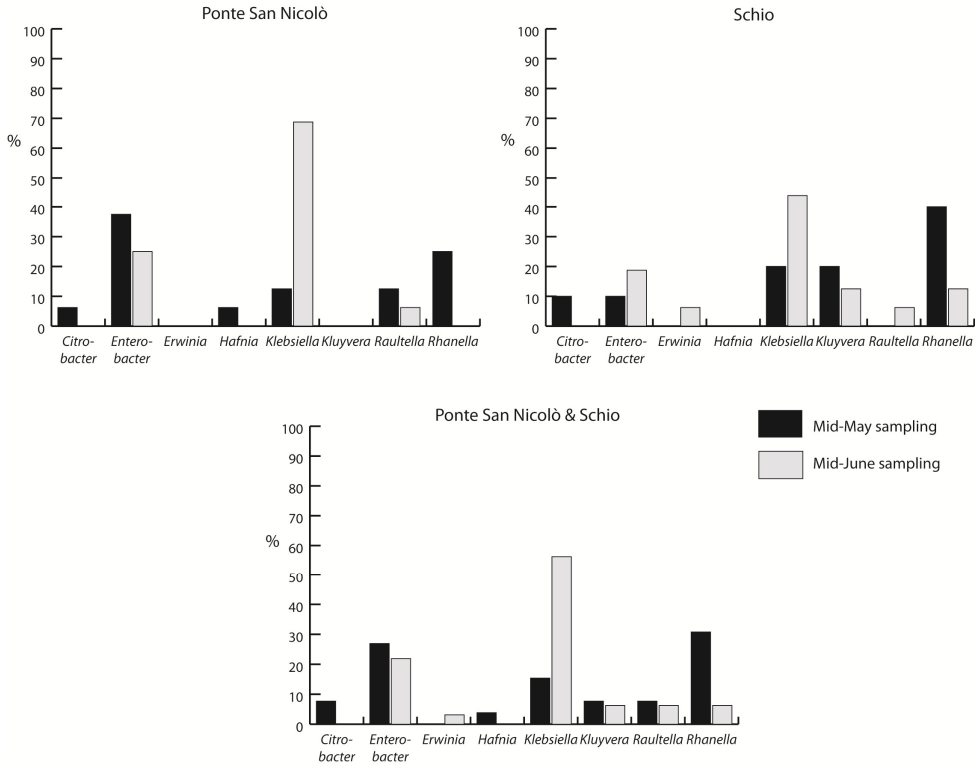

Fig. 4.3 – Percentage of bacterial genera found in *Rhagoletis cerasi* oesophageal bulbs during two sampling collections in 2010 in Ponte San Nicolò and Schio.

Conclusions

The present study surveyed the presence of the microflora and its composition in *R. cerasi* and *R. completa*, revealing that, unlike *B. oleae*, a real symbiosis is not established in their oesophageal bulbs. The microflora found could instead be defined as "associated". Furthermore, analysing the bacterial infection status in preimaginal stages of *R. cerasi*, we could exclude the presence of vertical transmission of bacteria from the female to the offspring during the oviposition, as eggs are uninfected. From our results it seems that bacteria are not essential for larval survival and mature larvae could acquire microorganisms from the environment and eventually transmit them to the adults, even if the mechanisms are still unknown.

Differently from what observed in pre-imaginal stages, where the presence of bacteria could be facultative, adults in the totality of the analyzed cases harboured bacteria in their oesophageal bulb, both *R. cerasi* and *R. completa*, belonging to a limited variety of Enterobacteriaceae species.

Flies can acquire bacteria from the larval/pupa stages but the most suitable hypothesis is they get them from the environment. It seems likely that flies can select the most suitable species to be multiplied inside the oesophageal bulb. Here we usually found few bacterial species and, as a rule, a single species is dominant, easily detectable with direct PCR on the oesophageal bulb content or by colonies isolation on plates.

The role of bacteria is not yet clear but it should be remembered that inside the oesophageal bulb of Tephritinae subfamily there is a continuously production of "membranous masses" filled with bacteria (Girolami, 1973). This imply a metabolic cost for the flies, hence the bacteria must be an utility in the fly survival. Otherwise it is certain that flies can survive without bacteria as shown by the flies reared in microbial controlled conditions.

Concluding *R. cerasi* develops within cherries not contaminated with bacteria, so we can infer that the fruit pulp offers to the larvae all the nutritional elements required, even if poor in proteins. As a future analysis it could be interestingly investigate the presence of antibacterial peptides on the laid eggs chorion as found in *Ceratitis capitata* by Marchini *et al*. (1997).

Chapter 5

Green fluorescent protein (GFP)-labeling of indigenous enterobacteria and persistence in their natural host *Rhagoletis completa* Cresson (Diptera, Tephritidae)

Submitted as:

Martinez-Sañudo I., Savio C., Mazzon L., Girolami V., Ciolfi S. and Marri L. Green fluorescent protein (GFP)-labeling of indigenous enterobacteria and persistence in their natural host *Rhagoletis completa* Cresson (Diptera, Tephritidae). Letters in Applied Microbiology.

I contributed to part of the experimental work, data analysis and to the paper writing.

Abstract

Strains of enterobacteria naturally occurring in the oesophageal bulb of fruit flies *Ceratitis capitata* and *Rhagoletis completa* were labeled by conjugation with the *gfp* gene carried by the suicide plasmid pTn*5*gfpmut1. The *gfp* gene was stably maintained in two tested transgenic strains. In one case, *gfp*-tagged bacterial cells were used to feed larvae of the original host *R. completa*. Engineered bacteria were able to colonize the gut of larvae and persisted through all larval instars to pupal stage. Flies continued to harbor viable *gfp*-tagged bacteria for at least one month after the ingestion.

Key words: green fluorescent protein, Enterobacteriaceae, Tephritidae

Fruit flies (Diptera: Tephritidae) are a highly successful, widespread group of insects causing economic damage in agriculture. Insects, like all other animals are colonised by a wide diversity of microbes resulting in many transient and some persistent relationships, then a comprehensive understanding of their biology requires that it must be studied in an ecological context with microorganisms as important component of the system (Steinhaus, 1960). The relationship between bacteria and fruit flies has been known since the beginning of the 1900 when Petri reported the presence of an hereditary symbiosis in the olive fly *Bactrocera oleae* (Rossi) (Ratner & Stoffolano, 1984). Whereas vertical transmission of symbiotic bacteria has been ascertained in the olive fly and in other non-frugivorus fruit flies belonging to the subfamily Tephritinae (Capuzzo *et al.*, 2005; Mazzon *et al.*, 2008), it is still a matter of debate in other tephritids even if it was hypothesized for the mediterranean fruit fly (medfly) *Ceratitis capitata* (Wiedemann) (Girolami, 1986; Lauzon *et al.*, 2009). Previous studies based on culture dependent methods of the bacterial community of *C. capitata* and *Rhagoletis* spp. digestive tract showed that members of the family Enterobacteriaceae, mainly *Klebsiella* spp. and *Enterobacter* spp. were the most often isolated bacteria (Drew & Lloyd, 1987; Howard *et al.*, 1985; Lauzon *et al.*, 1998; Marchini *et al.*, 2002; Rossiter *et al.*, 1983). Similarly, a systematic study of the structure and diversity of microbial communities in eggs, larvae, pupae, adult medflies and host fruit,

using a polyphasic approach revealed that Enterobacteriaceae constitute the dominant populations in the medfly's gut (Behar *et al*., 2005, 2008a, 2008b).

The use of green fluorescent protein (GFP) gene and derivatives to monitor the fate of microorganisms in an insect system has been mainly documented in host-pathogen interactions (Favia, 2008). Neverthenless, strains of *E. agglomerans* and *K. pneumoniae* originally isolated from the gut of walnut husk flies, *R. completa,* were transformed with a multicopy plasmid encoding a GFP derivative and bleomycin resistance. Both genes were stably expressed in the transformed bacteria and re-isolated 24 hours after ingestion from the gut of flies (Peloquin *et al.*, 2000, 2002).

Mini-Tn5 transposon derivatives have been widely used for the construction, manipulation and analysis of complex phenotypes in a wide range of Gram-negative bacteria (Suarez *et al.*, 1997). The mini-transposon is located on a suicide delivery plasmid for generation of stable chromosomal fusions. The stability of the engineered phenotype, being the desired trait encoded and maintained on the chromosome, is imperative for strains destined for use in an environmental study (de Lorenzo *et al.*, 1990, Eberl *et al.*, 1997). Thus, the chromosomal tagging of bacterial strains using GFP, allows to use the derived transgenic bacteria in studies about the spatial and functional interactions of microbes within the insect system and also to introduce detrimental genes for future pest and disease control strategies (Husseneder & Grace, 2005).

In this paper we examined the ability of enterobacterial strains isolated from the oesophageal bulb, a foregut diverticulum of adult Tephritids (Girolami, 1983; Ratner & Stoffolano, 1984), of *C. capitata* and *R. completa* to incorporate GFP encoding gene and, therefore, the colonizing ability of *gfp*-labeled bacteria in the original host.

Walnut husk flies were caught using yellow sticky traps in a English walnut orchard (*Juglans regia*) near Padua (45°20'N, 11°57'E) and in a black walnut orchard (*Juglans nigra*) near Pordenone (46°03'N, 12°47'E). Traps (20 x 25 cm) were checked daily, living flies were collected and immediately processed. Oesophageal bulbs were aseptically removed from flies and analyzed by cultivation-dependent and independent methods. In the first case they were streaked individually on plate count agar (PCA). After incubation at room temperature for 72 h single colonies for each distinct morphotype observed were

brought into pure culture. DNA was extracted either from pure bacterial cultures or from bulb homogenates and further characterized by 16S rRNA gene sequencing as previously described (Mazzon *et al*., 2008). Six new bacterial isolates as well as 12 *K. oxytoca* strains previously isolated from the oesophageal bulb of *C. capitata* (Marchini *et al.*, 2002) were checked for phenotypic traits in order to be used as recipients for conjugation of the shuttle plasmid pTn*5*gfpmut1 from *E. coli* S17-1 *λpir* (Tanaka *et al.*, 2006). The suitable recipient strains, listed in Table 5.1, were conjugated with *E. coli* S17-1 *λpir* (pTn5gfpmut1) as essentially described by Tanaka *et al.* (2006). Colony forming units were selected and enumerated on M9 medium (Miller, 1972) supplemented with kanamycin (50 µg/ml). The colonies were examined for green fluorescence under a Leica MZ FLIII stereomicroscope equipped with a GFP2 filter, and the frequency of transfer was calculated in relation to the total number of recipient cells counterselected by auxotrophy. Transconjugants were further isolated and labeled FR12 *gfp*-1 and FR12 *gfp*-2, etc. for each different recipient culture. Stability of the recombinant *gfp*-labeled strains derived by the two brightest fluorescing colonies was evaluated either by the shake flask method (Simon *et al.*, 1983), and on solid medium (Peloquin *et al*., 2000). For the latter transformed cells were grown on LB plates without antibiotics at room temperature for nine days, every three days, cells were re-streaked onto fresh plates, and examined for green fluorescence. *Gfp*-tagged bacteria were provided to insects in order to verify their fate and their host colonization ability. For this purpose, larvae of *R. completa* feeding on walnut husk tissues were collected as first and second instars; after immersion for two-three minutes in a solution of 1% sodium hypochlorite they were rinsed in sterile distilled water and then transferred in an artificial larval medium (Girolami, 1986), containing formaldehyde 1‰ w/w. After 48 hours each larva was aseptically transferred into a PCA plate and after three-four days (only if the medium remained axenic) were transferred for about 72 hours to PCA plates previously inoculated with the *gfp*-tagged bacteria. Larvae were then gently immersed in the solution of sodium hypochlorite 1% and rinsed twice in sterile distilled water. Part of the larvae (n=15) belonging to different instars, were sacrificed before pupation under a stereomicroscope to verify bacterial acquisition while the remaining larvae (n=15) were transferred to the artificial larval medium, lacking formaldehyde, to reach the mature stage and to allow pupation. Pupae were dissected 24 hours, 72 hours, seven days and 30 days

after pupation. All samples were prepared to be viewed under the fluorescence microscope. The acquisition of engineered bacteria was also tested in the adult stage. Flies (n=15) emerging from pupae collected in 2009 were fed *ad libitum* on SOC liquid medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 20 mM glucose) containing transgenic bacteria for two days. After 24, 72 hours, and five, 12, 30 days the oesophageal bulb and the midgut were removed from flies to detect the presence of acquired bacteria.

Based on the analysis of 16S rRNA sequences *Raoultella* (formerly *Klebsiella*) was the most representative genus among the bacterial isolates from oesophageal bulbs of *R. completa*. No differences were found at genus level between 16S rRNA sequences obtained from culturable and non-culturable dependent methods (data not shown). Among the total of 18 prototrophic strains tested, only four resulted sensitive to kanamycin, one of the antibiotic resistance markers of plasmid pTn5gfpmut1 (Tanaka *et al.*, 2006), which was then used for counterselection. The plasmid was successfully transferred into all four strains, and 100% of transconjugants expressed GFP as observed by fluorescence microscopy. However, the frequency of transfer showed different quantitative variation in recipient ability. In fact, *Raoultella* sp. FR53 resulted from 79 times to about 1,000 more efficient than other recipients (data not shown). The brightest GFP recombinants, FR12 gfp-2 and FR53 gfp-1, were tested for stability as reported above, 99% of all colonies examined ($\geq$ 500 from the shake flask method, and $\geq$ 1000 on solid medium) continued to express green fluorescence under non-selecting conditions after the longest period of observation (72 hours and nine days, respectively). Based on its longer-lasting fluorescent signal, the recombinant strain FR53 *gfp*-1 was selected in order to analyse its fate when introduced into larvae, pupae and adults of *R. completa*. Colonization by *gfp*-tagged bacteria appeared in larvae already 12 hours after ingestion, mainly in their midguts. The fluorescence was clearly visible under the epifluorescence microscope in all larval instars, also ten days after bacteria acquisition, just before pupation (data not shown). The ability of *gfp*-tagged bacteria to be transmitted across different life stages of the fly was verified in all dissected pupae developed from larvae fed with the transgenic strain . Regardless of the age of the pupae (24 and 72 hours, 30 days), a strong fluorescent signal was always present (Fig. 5.1). Similar results were observed in adults where fluorescence was detected into

oesophageal bulbs and midguts one month after the ingestion of engineered bacteria (Fig. 5.2). In addition, green fluorescent colonies were easily isolated on PCA plates from each life stage.

Data reported above show that *Raoultella* sp. FR53 *gfp*-1 cells were successfully ingested by *R. completa,* efficiently transmitted from larval to pupal stage and established a stable population in the fly gut. Flies harbored living *gfp*-tagged bacteria for up 30 days following ingestion. Our findings highlight the potential of the system used in this study either to monitor the bacterial transmission throughout different life stages, and to serve as a vector of detrimental genes for the biological control of the insect pest.

The 16S rRNA gene sequences of the three *Raoultella* spp. strains object of this study have been submitted to the DDBJ/EMBL/GenBank databases under accession numbers HQ833321, HQ833322, and HQ833323.

Acknowledgements

We thank Y. Saeki for providing us with strain *E. coli* S17-1 *λpir* (pTn5gfpmut1). This work was supported by the Italian Ministry of University (MIUR) under a Program for the Development of Research of National Interest (PRIN Grant No.2008BPLF7Y). The work is part of Savio C. PhD thesis, which is supported by “Veneto Agricoltura” scholarship.

Organism or plasmid	Phenotype, genotype or source	Reference
Strains		
Escherichia coli S17-1 λ*pir*	Tmp^R, Spc^R, Str^R *recA pro thi hsdR RP4-2-Tc::Mu-Km::Tn7* λ*pir*	Simon *et al.*, 1983
Klebsiella oxytoca KL26	Amp^r Km^r	Marchini *et al.*, 2002
Raoultella sp. FR12	Amp^r	
Raoultella sp. FR19	Amp^r	This study
Raoultella sp. FR53	Amp^r Km^r	
FR12 *gfp*-2	GFP recombinant of FR12	This study
FR53 *gfp*-1	GFP recombinant of FR53	This study
Plasmid		
pTn5Kmgfpmut1	Amp^r Km^r; *gfpmut*1	Tanaka *et al.*, 2006

Tab. 5.1- Bacterial strains and plasmid. **Amp**, ampicillin (50 µg/ml); **Km**, kanamicin (50 µg/ml); **Str**, streptomycin (50 µg/ml); **Gm**, gentamycin (30 µg/ml); **Spc**, spectinomicyn (100 µg/ml).

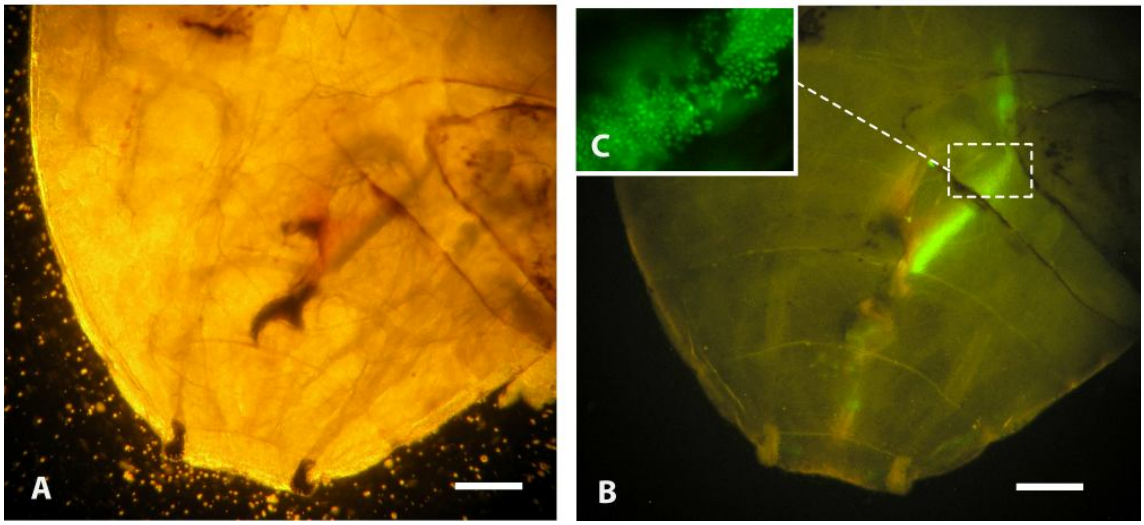

Fig. 5.1 - Colonization of a pupa of *R. completa* with *gfp*-tagged bacteria, 48 hours after pupation viewed by phase contrast (A) and fluorescence microscopy (B) 100x magnification (C). Scale bar, 350 µm. Pupa was fed as larva with *Raoultella* sp. FR53 *gfp*-1.

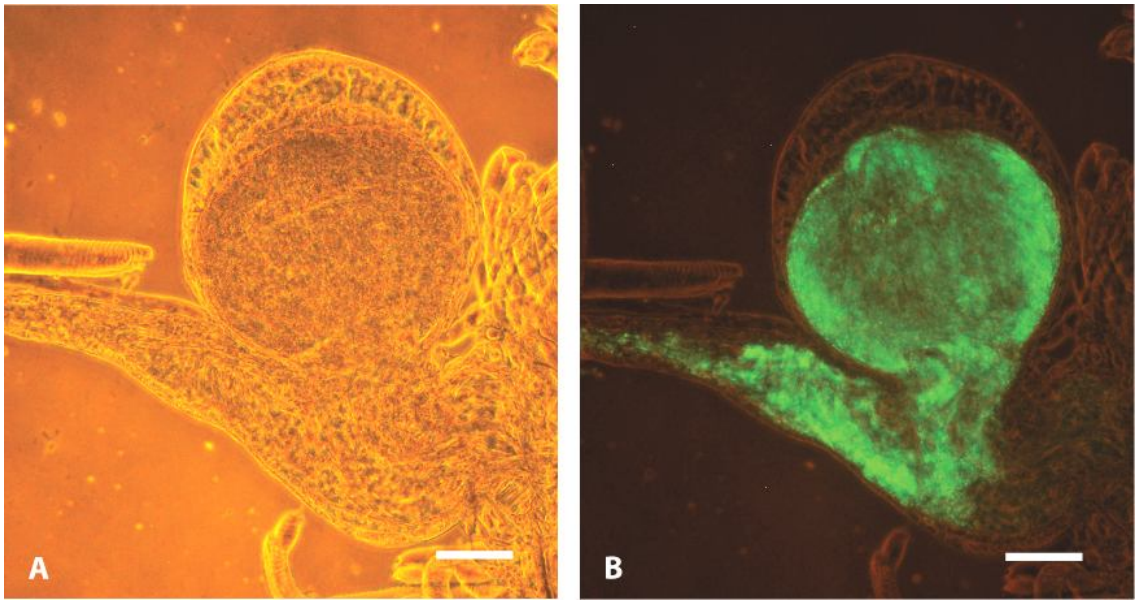

Fig. 5.2 - *Gfp*-tagged bacteria in the oesophageal bulb of *R. completa* 48 hours after the ingestion viewed by phase contrast (A) and fluorescence microscopy (B). Scale bar, 50 µm.

Chapter 6

Phylogenetic relationships between flies of the Tephritinae subfamily (Diptera, Tephritidae) and their symbiotic bacteria

---

Published as:

Mazzon L., Martinez-Sañudo I., Simonato M., Squartini A. Savio C. and Girolami V. 2010. Phylogenetic relationships between flies of the Tephritinae subfamily (Diptera, Tephritidae) and their symbiotic bacteria. Molecular Phylogenetics and Evolution **56:**312–326.

I contributed to part of the experimental work and to the paper writing.

Abstract

The Tephritinae is considered the most specialized subfamily of fruit flies, predominantly infesting flowerheads of Asteraceae. Some species are known to host specific non-culturable symbiont bacteria ("Candidatus *Stammerula* spp.") in the midgut. In this work we (i) examined the phylogenetic relationships among the insect hosts, (ii) investigated the presence of bacteria in other hitherto unexamined species, and (iii) evaluated the phylogenetic congruence between insects and symbionts.

A total of 33 Tephritinae species in 17 different genera were analyzed. Two regions of the mitochondrial DNA (16S rDNA and COI-tRNALeu-COII) were examined in the insect host, while the 16S was analyzed in the bacteria. From the phylogenetic trees, four of the five tribes considered were statistically supported by each of the clustering methods used. Species belonging to the tribe Noeetini never clustered at significant levels. The phylogenetic COI-tRNALeu-COII tree showed internal nodes more highly supported than the 16S phylogeny. The analysis of the distribution of symbiosis across the subfamily has highlighted the presence of bacteria only in the tribe Tephritini and in the genus *Noeeta* from the tribe Noeetini.

A cophylogenetic analysis revealed a substantial congruence between hosts and symbionts. The interesting exceptions can be justified by events like losses, duplications and hosts switching opportunities, which are likely to arise during the biological cycle of the fly in consideration of the extracellular status of these symbionts.

Keywords: Endosymbiosis; "*Candidatus* Stammerula"; Cophylogeny; Molecular phylogeny; Tephritinae.

Introduction

Several kinds of intimate mutual associations between bacteria and insects are known (e.g. Buchner, 1965; Baumann & Moran, 1997). Extensive studies have been devoted in particular to those in which symbiotic bacteria are endocellular and housed in the cytoplasm of specialized cells called bacteriocytes or mycetocytes. In these cases bacteria are bound to

live in a close environment with limited possibilities of gene exchange (Buchner, 1965; Moran & Baumann, 2000; Gil *et al*., 2004; Wernegreen, 2004). These symbionts are vertically transmitted to the next generation and are not able to live outside of their host. Technical progress in the fields of molecular phylogeny and bioinformatics has allowed to trace the evolutionary routes of these non culturable bacteria. Several examples agree on the evidence that host phylogeny mirrors symbiont phylogeny (Chen *et al.*, 1999; Clark *et al.*, 2000; Moran, 2001; Moran *et al.*, 2003; Gruwell *et al.*, 2007). A congruence between insect's and endocellular symbiont's phylogenies implies a single bacterial infection of the host ancestor, followed by co-evolution of both partners (Gil *et al.*, 2004). Moreover, endocellular bacteria, in contrast to their closest free-living bacteria, display distinctive genetic properties including AT-biased base composition, accelerated molecular evolution, and, in some cases, a small genome size; these features suggest an increased genetic drift (Moran & Baumann, 2000).

The presence of extracellular symbiotic bacteria has been described for insects belonging to different systematic groups. These symbiotic bacteria are harboured in the gut cavity and some are known to play substantial biological roles for their hosts (Dillon & Dillon, 2004). For some the presence of a vertical transmission has been reported. For example stinkbugs of the family Platasipidae harbour a bacterial symbiont in the midgut which is transmitted to the new generation orally, by a capsule containing the bacteria placed under the egg mass (Fukatsu & Hosokawa, 2002; Hosokawa *et al.*, 2006).

Such extracellular associations are thought to be evolutionarily more occasional than the endocellular associations, on the grounds that the symbionts are not isolated in the body cavity and thus prone to invasion and replacement by foreign microbes. Therefore, in these cases, a phylogenetic congruence between extracellular bacteria and host is assumed to be uncommon (Donovan *et al.*, 2004).

The family Tephritidae is commonly known as fruit flies and includes about 4300 described species worldwide in almost 500 genera (White, 2006). Tephritid larvae develop mainly in fruits, leaves, or within the flower heads of Asteraceae (White, 1988). Many species, mostly the carpophagous ones, are considered a notorious group of agricultural pests. Tephritinae, considered the most specialized subfamily of tephritids and containing

203 genera and 1847 species from all zoogeographical regions (Norrbom *et al.*, 1999; Korneyev, 1999), predominantly infest flower heads of Asteraceae.

As regards the bacterial relationships of tephritid flies there is extensive literature describing the presence of "associated bacteria" mostly belonging to genera *Enterobacter*, *Klebsiella* and *Pantoea* (Lloyd *et al.*, 1986; Drew & Lloyd, 1987; Daser & Brandl, 1992; Marchini *et al.*, 2002; Lauzon, 2003). These associations, despite their importance during the life of the insect, can be considered facultative (Drew & Lloyd, 1991).

The first hereditary symbiosis in the Tephritidae family has been described in the olive fly *Bactrocera oleae* (Rossi) by Petri (1909). The mother transmits symbiotic bacteria to the new generation, smearing the surface of its eggs with bacteria. The symbiont is extracellular but multiplies 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, the obligate symbiont multiplies within a foregut diverticulum (Petri, 1909), called oesophageal bulb, that is present in all adult tephritid flies (Girolami, 1973). Recently, the presence of a hereditary, host-specific symbiotic bacterium, designated as '*Candidatus* Erwinia dacicola' (hereafter *Erwinia dacicola)*, has been confirmed in all wild specimens and in adults emerging from previously surface-sterilized pupae (Capuzzo *et al.*, 2005).

Besides *B. oleae*, the presence of symbiotic bacteria has been long since reported in some species belonging to the subfamily Tephritinae (Stammer, 1929). In this case the oesophageal bulb is smaller than the one of *B. oleae* and is devoid of bacteria. Symbiotic bacteria are instead located in the first tract of the midgut, in contact with the epithelium but, as later reported by Girolami (1983), outside the peritrophic membrane. Thus they are not in direct contact with the food bolus and with the general gut transit. Upon examining the cases originally pointed out by Stammer, the presence of symbiotic bacteria has been described in flies belonging to several genera of Tephritinae, delineating their phylogeny. The biomolecular analyses carried out in 25 insect species showed a correspondence between the species of host fly and the sequences of bacterial 16S rDNA (Mazzon *et al.*, 2008). The phylogenetic analyses recognized three main groups; members of one of them are monophyletic and homogeneous and now designated under a novel genus and species

as '*Candidatus* Stammerula tephritidis' (Mazzon *et al.*, 2008) (hereafter *Stammerula tephritidis*).

The phylogeny of the Tephritidae family, primarily based on morphological data, has been recently abridged using molecular techniques based on nucleotide sequence data. These recent studies have suggested several previously unknown relationships and have offered new possibilities for tephritid classification (Han & McPheron, 1997; 1999; Han and Ro, 2009). Concerning the subfamily Tephritinae, to date only the report from Han *et al.* (2006) was available, in which the authors examined 16S rDNA of species originating in various zoogeographical regions.

In the present work, we have sequenced a region of the 16S rDNA and the cytochrome oxidase genes COI and COII of Tephritinae species, in which the presence of symbiotic bacteria has been found in Mazzon *et al.* (2008). Moreover, the analysis has been expanded to other tribes of Tephritinae in order to verify the presence of symbiosis also in species in which Stammer (1929), using the traditional microscopic techniques, did not detect it. In total, over 65% of the genera belonging to the Tephritinae subfamily present in the Italian fauna checklist (Belcari *et al.*, 1995) have been analyzed.

The approach enabled: a) to study the molecular relationship among the tribes and the species of the Tephritinae; b) to verify in which species and tribes the symbiotic bacteria are present; c) to test the phylogenetic congruence between tephritid flies and their symbiotic bacteria for the species harboring them.

Materials and methods

Origin of the biological material

In the present work, 36 species belonging to three tephritid subfamilies were analyzed, encompassing the western palaeartic range for the Tephritinae subfamily (Tab. 6.1). For this subfamily, besides the 16 species of the 25 in which symbiotic bacteria have been previously reported (Mazzon *et al.*, 2008), 17 other species were studied. Therefore, a total of 33 species and 17 different genera belonging to the largest five of the nine Palaeartic tribes (Norrbom *et al.*, 1999) have been examined. We also selected one species from the

family Platystomatidae as remote outgroup and chose four species belonging to Dacinae and Trypetinae subfamilies as more closely related outgroups (Tab. 6.1).

Fly specimens were obtained from infested flower heads of Asteraceae collected in the field in Northern Italy mostly in the same locations reported in the previous research (Mazzon *et al.*, 2008).

Flower heads, inspected for the presence of larvae or pupae, were kept in the laboratory for some days to allow larvae to maturate and enter the pupal stage. Pupae were subjected to surface sterilization and emerged adults were reared in aseptic conditions as described below.

Tephritid higher-level classification and specific names follow Merz (1994; 1999), Norrbom *et al.* (1999) and Korneyev (1999). A set of samples from each of the studied species were dried, pinned and deposited as reference in the Department of Environmental Agronomy and Crop Sciences – Entomology (DAAPV), University of Padua, Italy.

Insects

DNA extraction, amplification and sequencing of insect genes

DNA from the insect body, was extracted with the salting-out protocol (Patwary *et al.*, 1994). At least two samples for each species 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 ethidium bromide or SYBR® Safe DNA gel stain (Invitrogen).

Two regions of the mitochondrial DNA, namely a fragment of 16S rDNA and a fragment including the 3' region of cytochrome oxidase sub-unit I, tRNA-Leu and the 5' region of cytochrome oxidase sub-unit II, were amplified.

PCR reaction was carried out in a 20 µl volume containing 2 µl of DNA extract , 4 µl PCR Buffer of 5X colorless GoTaq Flexi Buffer (Promega), 2.5 mM MgCl, 100 mM dNTPs, 0.5 µM of each primer and 1U of GoTaq Flexi DNA polymerase (Promega).

A combination of universal and specific primers was used for PCR amplification and sequencing (Tab. 6.2). The cycling program was carried out in an Eppendorf Mastercycler Gradient and consisted of a first step at 96°C for 5 min followed by 35 cycles with a denaturation step of 96 °C for 1 min., an annealing step ranging between 52°C and 62°C for

1 min and an extension step of 72°C from 1-2 min followed by a final extension at 72°C for 5 min. The amplified products were separated in a 1% agarose gel and viewed under UV with an ethidium bromide or SYBR® Safe (Invitrogen) staining**.** PCR products were purified with the ExoSAP-IT kit (Amersham Biosciences) and directly sequenced. Sequencing was performed at the BMR Genomics service (Padova, Italy).

DNA sequence alignment

Sequences of 16S rDNA and COI-tRNALeu-COII were inspected and aligned using MEGA 4.0 (Tamura *et al.*, 2007).

During the alignment of 16S sequences some ambiguous portions appeared. In order to maximize information from those portions that are usually excluded from an analysis, alignment was performed considering the rDNA secondary structure, avoiding the deletion of the ambiguous stretches. However, as a comparison, we also aligned 16S sequences in the straight fashion by removing the ambiguous portions from the dataset with Gblocks 0.91b software (Castresana, 2000). Considering the secondary structure of rDNA in the alignment, we took into account the correlation between nucleotide sites in the stem regions; this information, when ignored, could lead to a bias in the supporting confidence of the clades. Moreover, a phylogenetic study based on the secondary structure seems to also have an advantage with closely related species (Buckley *et al.*, 2000). Secondary structures of 16S rDNA were manually aligned according to the method outlined by Yoder and Gillespie, 2004. We used Mega 4.0 software to construct a mask for the structural alignment based on the rDNA secondary structure of *Drosophila melanogaster* (Cannone *et al.*, 2002, http://www.rna.ccbb.utexas.edu). For the stems we considered both the Watson-Crick and the GU-UG pairs. Stems regions were carefully checked to allow at least 75% of the sequences to match the secondary structure guide. When 75% of a sequence did not follow the secondary structure guide a consensus secondary structure was created for that stem using the "secondary struct consensus" option available in the software *PHASE* Version 2.0 (Jow *et al.*, 2005).

All alignments are available online as supplementary data (see Appendix). All sequences were deposited in GenBank. Accession numbers for each sequence are show in Tab. 6.1

(GenBank Accession Nos.: GQ175790-GQ175826 for COI-tRNALeu-COII and GQ175827-GQ175863 for 16S rDNA).

Phylogenetic analysis

In DNA sequence datasets, multiple substitutions that saturate sites are the background noise disturbing evolutionary signal (Swofford *et al.*, 1996; Jesus *et al.*, 2007; Chen *et al.*, 2009). To test for saturation, observed pairwise proportions of transitions and transversions were plotted against pairwise sequence divergence calculated through a K80 model (Kimura, 1980) using DAMBE v4.2.13 (Xia & Xie, 2001).

Phylogenetic relationships among sequences of insects, for the 16S data set, the COI-tRNALeu-COII data set and the combined 16S and COI-tRNALeu-COII data sets, were estimated using two methods: approximate maximum-likelihood (ML) and Bayesian inference (BI) analysis.

For ML analysis the best-fit model of sequence evolution was selected by MODEL TEST v3.06 (Posada & Crandall, 1998) using Akaike Information Criterion tests (Posada & Buckley, 2004) for both 16S dataset and COI-tRNALeu-COII dataset. The best model found was used for approximate ML, using PHYML v2.4.4 software (Guindon & Gascuel, 2003), with neighbour-joining starting trees and 100 bootstrap replications. The software GARLI v0.951 (Zwickl, 2006) with 100 replicates of bootstrap was also used to obtain trees for the following analysis. Based on previous fruit flies phylogenetic studies (Han & McPheron, 1997, 1999; Han *et al.*, 2006) clades were considered statistically significant with bootstrap probabilities (Bp) values of 70%. For the BI analysis of the host data set MrBayes 3.1.2 (Ronquist & Huelsenbeck, 2003) was used. For the 16S rRNA dataset a doublet model to the stem regions and a standard 4by4 nucleotide model for the loop regions were applied**;** for the COI-tRNALeu-COII dataset a codon site partioning scheme, with a 4by4 nucleotide model for each codon position we used. Two independent iterations were run for 5,000000 generations and sampled every 100 generations.

The 50% majority rule consensus tree and Bayesian posterior probability of support were obtained discarding the first 25% of sampled generations (burn in of 12500). Posterior probabilities (Pp) values of 95% were considered statistically significant for clades to be supported in the phylogenetic tree as suggested by Huelsenbeck & Rannala (2004).

As shown in several studies, topologies recovered from combined data are generally better resolved than those based on individual partitions (Baker & DeSalle, 1997).

TREEFINDER software version October 2008 (www.treefinder.de) (Jobb *et al.*, 2004) was used in order to confirm the possibility of combining molecular data sets for the host analysis. This software compares topologies of two or more trees from different datasets using some paired-sites tests such as SH test (Shimodaira & Hasegawa, 1999). As in the null hypothesis all the topologies are good explanations of the data, a not significant p-value ($P > 0.05$) allows to combine the tested datasets.

Symbionts

Symbiotic bacteria detection

In order to extend the search for bacterial symbionts to those species where Stammer, using traditional microscopy-based approaches, had not detected them, we used the culture-independent methodology adopted by Capuzzo *et al.* (2005) and Mazzon *et al.* (2008).

Each pupa, following a routine lab procedure, was extracted by dissecting the flower head under a stereomicroscope and surface-sterilized by a 5 min immersion in 1% sodium hypochlorite, rinsed in sterile water for two times, air-dried in sterile conditions and kept in sterile vials until adult emergence. Resulting flies were kept under microbiologically controlled conditions and then aseptically transferred, under a laminar flow hood, into larger vials, containing a layer of Plate Count Agar on the bottom as a sterility check. Vials were sealed with a sterilized transparent gas-permeable cellulose membrane for dialysis (Sigma-Aldrich chemical co. S. Louis MO, USA). A drop of sterile glucose solution was placed on the internal side of the membrane to allow insect feeding. The drop of solution was re-wetted, whenever necessary, by spraying water on the external surface of the cellulose membrane at the top of the vials. After a week of rearing, adult flies were dissected in order to draw the midgut and verify the presence of bacteria. Then, samples were routinely preserved in 95% ethanol and stored at -80°C until processed.

At least the midgut of two flies per species were analyzed (Tab. 6.1). Each midgut of the insects reared under microbiologically-controlled conditions was analyzed by physiological staining using the LIVE/DEAD® BacLight™ Bacterial Viability Kits and plating on different standard microbiological media, as previously reported by Capuzzo *et al.* (2005),

in order to test the presence of culturable bacteria. Subsequently, from the midgut of the most abundant and representative species of each tribe (Tab. 6.1) not yielding culturable bacteria, we performed biomolecular analyses with universal bacterial 16S rDNA primers fD1 (59-AGAGTTTGATCCTGGCTCAG-39) and
rP1 (59-ACGGTTACCTTGTTACGACTT-39) (Weisburg *et al.*, 1991).
Moreover, midguts of two samples collected in the field belonging to two species (*T. formosa* and *T. matricariae*) in which symbiotic bacteria had been detected previously (Mazzon *et al.*, 2008), were cloned using the QIAGEN pDrive cloning vector, and 20 positive colonies per ligation were sequenced.

Bacterial sequence alignment and analysis

Symbiont sequences were aligned following the same methodology used for the insect host sequences. In this case the rRNA secondary structure of *Escherichia coli* was used as a guide (Cannone *et al.*, 2002, http://www.rna.icmb.utexas.edu/).
The phylogenetic trees were realized with the methods reported above for the insect host (ML and BI). The best-fit model of sequence evolution selected by MODEL TEST v3.06 (Posada & Crandall, 1998) using Akaike Information Criterion tests (Posada & Buckley, 2004) was (TrN+I+G). Accession numbers for each of the sequences are shown in Tab.6.1.

Coevolutionary analysis

To analyze the coevolutionary history between insect host and symbiotic bacteria three different methods based on the trees, on the distances and on the likelihood were employed as follows using, respectively, TreeMap software, ParaFit software and the S-H test.
Reconciliation analysis was performed using TreeMap v2.02β, available at http://www.it.usyd.edu.au/~mcharles/software/software.html (Charleston & Page, 2002). TreeMap allows to estimate the level of congruence between host and symbiont by calculating the number of codivergences, duplications, losses and host switches (the program considers the host switch event as a possible explanation for incongruent host-parasite phylogenies).
TreeMap computes all optimal solutions by exhaustive search represented by Jungles algorithm (Charleston, 1998). As the number of possible reconstructions for the history of a

host–symbiont assemblage can be very large, finding all possible solutions can be computationally prohibitive in terms of both time and memory (Charleston, 1998; Page *et al.*, 2004). Therefore, the set of potential solutions was imposed to no more than six hosts switches. Default settings were used for evolutionary events (assigning a cost of zero for codivergence events and a cost of one for host switches, duplication and losses).

The statistical significance was evaluated by testing the null hypothesis that the observed number of codivergence events was not larger than the expected number of codivergence events between the observed host tree and 1000 randomly generated trees.

ParaFit software (Legendre *et al.*, 2002), available online at http://www.bio.umontreal.ca/casgrain/en/labo/parafit.html, was used to assess the null hypothesis of independent evolution of hosts and symbionts and to test the significance of each host-symbiont link. This leads to the identification of the species involved in cospeciation (Legendre, 2001). ParaFit software is a useful tool and has the advantage over tree-based methods, because it can accommodate uncertainty in tree topologies, multiple hosts per symbiont lineage and, as in our case, multiple symbionts per host lineage (Light & Hafner, 2007).

ParaFit software can compute this statistical test using phylogenetic distances, obtained using Mega 4.0 software, with ML estimates of pairwise genetic distances. Distances matrices were transformed into a rectangular matrix by principal coordinate analysis, using DistPCoA software (Legendre & Anderson, 1998) before being used in ParaFit. These matrices allow the software to calculate the probability of host-symbiont coevolution.

The congruence of host and symbiont was also assessed with Shimodaira-Hasegawa likelihood-based test (S-H test) which assess if a set of given trees equally explain sequences data (Shimodaira & Hasegawa, 1999; Goldman *et al.*, 2000). Thus we tested each dataset against both insect and symbiont trees with a run of 10000 RELL (re-estimation of likelihoods) bootstrap replicates using PAUP*4.0b2.

Results

Insects

Sequence analysis

From the DNA amplification and sequencing with the primers reported above, we obtained fragments ranging from 989 to 1036 bp for the 16S rDNA gene, with an average of 1027 bp, and from 1407 to 1553 for COI-tRNALeu-COII genes with an average of 1414 bp.

For 32 of the species considered, this constitutes the first sequencing of the cytochrome oxidase gene and for 25, the first sequencing of 16S rDNA gene.

Correlation between number of substitutions per site and genetic distances for transitions and transversion of the 16S and the COI-tRNALeu-COII fragments are reported in Fig. 6.1. Transitional substitution in 16S rDNA showed a heavy degree of saturation, therefore suggesting high convergence in this region among taxa.

For the 16S rDNA fragment the average proportion of T:C:A:G was 45:6:38:11. To avoid the loss of information due to the removal of ambiguous portions, an alignment of 1094 bp sites was realized on the basis of the secondary structure of *Drosophila melanogaster*.

For COI-tRNALeu-COII fragment the average proportion of T:C:A:G was 39:14:34:12 with an aligned number of sites of 1568 bp.

The best trees obtained from GARLI and PHYML for both genes were compared using TREEFINDER. For 16S dataset we considered the GARLI output as it yielded the tree with the best likelihood (log-likelihood= -8969 against -9237 of PHYML). The comparison between this tree and COI-tRNALeu-COII trees showed that the 16S dataset is consistent with COI-tRNALeu-COII dataset (P=0.21). On the other hand for the COI-tRNALeu-COII dataset the best tree was obtained with PHYML (log-likelihood= -14473 against -14869 of GARLI). In this case the COI-tRNALeu-COII tree was not in agreement with the 16S trees (P< 0.01). However, as both datasets were not mutually incompatible, we considered 16S and COI-tRNALeu-COII suitable to be combined.

For the combined data set a fragment ranging from 2399 to 2582 bp with an average of 2441 bp was obtained with a T:C:A:G proportion of 41:11:36:12. The alignment produced a data set with a total of 2662 bp.

Phylogenetic analysis

The best-fit evolutionary model for the ML-based phylogenetic analysis (as determined by Modeltest) was a general time-reversible model (GTR+I+G), for the 16S, the COI-tRNALeu-COII and the combined dataset.

In each dataset both ML and BI methods yielded similar topologies with different statistical supports for branch nodes.

For the 16S dataset aligned through secondary structure, 19 nodes for ML, 18 for BI and 17 for both methods were highly statistically supported; although secondary structure alignment did not increase the number of supported nodes in the deeper branches, compared to Gblocks analysis (data not shown), it improved supported values both for ML and BI methods.

For the COI-tRNALeu-COII dataset, 22 nodes for both ML and BI, and 20 for both methods resulted highly statistically supported. In the combined dataset numbers of supported nodes were the same as COI-tRNALeu-COII except for BI with 24 highly statistical supported nodes.

All the phylogenetic trees showed that each of the 33 analyzed species, belonging to the subfamily Tephritinae, formed a strongly monophyletic group highly supported in both Bp and Pp analyses (Fig. 6.2, 6.3 and 6.4). Four of the five tribes analyzed (Tephritini, Terelliini, Myopitini, Xyphosiini) were recognized as monophyletic and produced highly supported clades by each of the clustering methods used and in all datasets. Species belonging to the tribe Noeetini were instead never clustered together (Fig. 6.2, 6.3 and 6.4).

Some disagreement appeared in the relationship among tribes considering the different phylogenies. While in the 16S tree the five tribes did not show a further clustering among them, in COI-tRNALeu-COII the tribes presented supported relationships at the deeper node branches (Fig. 6.2 and 6.3). As the 16S dataset features the lowest number of supported tree nodes (in particular the nodes at the deeper branches important to resolve the relationships among the tribes) and it presents a high saturation for transitions (Fig. 6.1) it appears to be less informative. It is noteworthy that the support values in the deeper branches of COI-tRNALeu-COII disappear when both dataset are combined (Fig. 6.4). Although diverse factors could explain this lack of signal such as different evolutionary constrains among genes (Naylor and Brown, 1998; Rokas *et al.*, 2003), if one considers the

evidence shown in Fig.6.1, saturation could be considered one the likely sources of conflict in this combined tree (Fig. 6.4). Moreover, even if combined analyses are usually considered more meaningful, in some cases, due to the contradictory signal of the single dataset, they may become ineffective (Leicointre and Deleporte, 2005; Chen *et al.*, 2009). Therefore, we chose to examine our results mainly considering the COI-tRNALeu-COII data set.

The main relationships among clusters in the COI-tRNALeu-COII dataset showed that clade 1 (tribe Tephritini) and clade 2 (tribe Myopitini) grouped with high statistical support values (Bp/Pp=74/97). Moreover, the species of clade 5 (tribe Terelliini) resulted the sister group of the other four tribes with statistical support values of 61 for Bp and 98 for Pp (Fig. 6.2).

The Tephritini tribe is subdivided into three highly supported monophyletic subgroups: 1A (Bp/Pp=100/100), 1B (Bp/Pp=92/100) and 1C (Bp/Pp=91/100) (Fig. 6.2). The first one (1A), represented by the species of genera *Tephritis*, *Capitites*, *Trupanea* and *Acanthiophilus*, corresponds to the "*Tephritis* Group" defined by Merz (1999), on the basis of a morphological approach. The second one (1B) includes species of the genera *Campiglossa*, *Oxyna* and *Dioxyna* belonging to "*Campiglossa* group" defined by Norrbom *et al.* (1999) and Korneyev (1999). The third one (1C) is formed by *Campiglossa* group together with *Sphenella marginata*. In the 16S tree, the clade 1A shows a low support (Bp/Pp=61/*) while 1B and 1C is not recognized (Fig. 6.3).

The clade 3 (Tribe Xyphosiini) is grouped to the clade 4 (Tribe Noeetini) with statistical support value for the BI analysis (Pp=100). This group is not confirmed by the other dataset (Fig. 6.3 and 6.4).

The clade 2 (tribe Myopitini), the clade 3 (tribe Xyphosiini), the clade 4 (tribe Noeetini) and the clade 5 (tribe Terelliini) are all included in the "Lower Tephritinae" (Korneyev, 1999).

Symbionts

Distribution of the symbiosis across the subfamily Tephritinae

The presence of uncultured inheritable symbiotic bacteria, firstly proposed by Stammer (1929) in the midgut of some species of the subfamily, was already confirmed by our previous report (Mazzon *et al.*, 2008). In the same account the presence of symbiotic bacteria in species of the genus *Noeeta*, belonging to the tribe Noeetini, was reported. In the present work, we verified, using modern approaches, the absence of symbiotic bacteria in the midgut of the remaining tribes as postulated by Stammer.

The analyses were carried out with adults emerging from surface-sterilized pupae, subsequently reared under microbiologically controlled conditions into vials containing a layer of Plate Count Agar on the bottom. Sporadic cases (<5%) in which microbial colonies had developed on the PCA agar on the bottom of the vials were not considered.

Using this method, the midgut content of a total of 54 specimens was analyzed (at least two samples per species) (Tab. 6.1). Neither by phase contrast microscopy nor by the LIVE/DEAD BacLight bacterial viability test could we detect the presence of bacteria in the midgut of dissected adults. No bacterial growth was observed by plating the midgut in different media either. Moreover, the amplification of DNA extracted from the midgut, using universal bacterial primers targeting a fragment of 16S rDNA, consistently gave negative results (Tab. 6.1).

In addition, 90% of the clones obtained from the midgut of *T. formosa* and *T. matricariae* corresponded to *Stammerula tephritidis*. This fact supports the presence of a single or numerically-prevailing symbiotic bacterium in the midgut and confirms the validity of approaches based on direct sequencing of amplicons raised directly from the insect tissues without the cloning step.

It appears that symbiotic bacteria residing in the midgut are present in all the examined species of the tribe Tephritini and in the Tribe Noeetini in the genus *Noeeta* but not in *Ensina*. In the analysis of tribes Terelliini, Myopitini and Xyphosiini the presence of bacteria in the midgut of individuals reared in microbiologically controlled conditions, was never observed (Fig. 6.5). The list of studied species in which symbiotic bacteria have been found is shown in Table 6.1.

Coevolutionary analysis

To assess the coevolution between hosts and symbionts, the 17 tephritid species containing symbiotic bacteria and their corresponding prokaryotes, whose sequences were previously published (Mazzon *et al.*, 2008), were analyzed. As for the previous phylogenetic considerations on the host, the phylogeny of COI-tRNALeu-COII was chosen for the coevolutionary analysis.

For the TreeMap analysis, since this program requires fully resolved trees, we used ML instead of BI diagrams as they showed to be more resolved for both hosts and symbionts. Comparing the host and endosymbiont tanglegram (Fig. 6.6a) using the jungle algorithm implemented in TreeMap v2.02β (Charleston & Page, 2002), 21 optimal reconstructions were found. The optimal solutions postulated a maximum of 20 codivergence events (=10 cospeciations), from 6 to 17 losses, 1 to 6 switches and 12 to 14 duplication events (Fig. 6.6). The number of cospeciating nodes resulting (20) divided by the total number of nodes (32) and multiplied by 100 gave a percentage of nodes with cospeciation equal to 62.5%. The randomization of 1000 associated host-symbiont trees, suggests that the fit between the host and the symbiont trees is statistically significant ($P< 0.001$) (Fig. 6.6c).

The global test of cophylogeny, assessed through ParaFit, resulted in the rejection of the null hypothesis that the evolution of the host and symbiont was independent ($P=0.007$, 999 permutations), thus supporting the global association between hosts and symbionts. To assess the significance of each host–symbiont association ParaFit calculates two statistics, parafitLink1 and parafitLink2. Although no fundamental differences were found between the two statistic tests, we chose to use parafitLink1 because it is more conservative, especially in situations in which portions of the two trees are coevolutionary while other portions are not (Legendre *et al.*, 2002). Using parafitLink1 we found that 4 of the 17 host-symbiont links were significant ($P<0.05$) (Tab. 6.3). Legendre *et al.* (2002) report that in a situation of a significant global test and no significant value for some individual links, these host-symbiont associations are not coevolved links. However, because the test of individual links has less power than the global tests, in some situations this individual test could not be significant, due to this lack of power.

The results of the SH-test indicated that there is a significant disagreement in the most likely topology supported by insects and symbionts dataset. This result suggests that

the observed incongruence found (differences between host and symbiont phylogenies) cannot be explained by sampling error but as the result of other historical events such as host switching, losses or duplications.

Discussion

Molecular relationships among the tribes of the subfamily Tephritinae

A single molecular phylogeny study of the subfamily Tephritinae was available up now (Han *et al.*, 2006). It has been obtained analyzing the 16S rDNA of species coming from several zoogeographic regions.

We have focused our study on a single geographic area (western Palaeartic) considering the 16S rDNA and the COI-tRNALeu-COII loci. When possible our phylogenetic results have been compared to the phylogenetic reconstructions obtained by Han *et al.* (2006) and with conventional classification by Norrbom *et al.* (1999), Friedberg & Norrbom (1999), Korneyev (1999), Merz (1999) (Tab. 6.4).

Our molecular phylogenetic analysis confirms, in agreement with the molecular analysis of Han *et al.* (2006), the monophyly of the subfamily with high statistical support value (Fig. 6.2, 6.3 and 6.4). Moreover, the strict affinity among species of the subfamily has been proposed by several authors on the basis of a morphological and biological approach (Zwölfer, 1983; Foote *et al.*, 1993; Norrbom *et al.*, 1999; Korneyev, 1999).

The molecular results of the three data sets studied (COI-tRNALeu-COII, 16S rDNA and combined dataset) showed four monophyletic and highly supported clades corresponding to recognized tribes (Tephritini, Terelliini, Myopitini, Xyphosiini) on the basis of conventional classification. The remaining tribe studied (Noeetini) was not statistically supported by any of our phylogenies or by Han *et al.* (2006) (Fig. 6.2, 6.3, 6.4 and Tab. 6.4). The relationships among the tribes have been discussed in the literature from a morphological (Korneyev, 1999; Norrbom *et al.*, 1999) as well as a molecular point of view (Han *et al.*, 2006). In this respect, our phylogeny of the COI-tRNALeu-COII provides useful updated information since all the internal nodes are highly supported.

The clade 1 corresponds to the large tribe of Tephritini. This tribe is, on the basis of three possible synapomorphies, conventionally included with other five tribes, in the "Higher Tephritinae" (Korneyev, 1999). Inside this tribe, the COI-tRNALeu-COII phylogeny mirrors the conventional accepted groups; the *Tephritis* group (clade 1A) by Merz (1999) and *Campiglossa* group (clade 1B) by Norrbom *et al.* (1999) and Korneyev (1999) are clearly recognized and highly supported (100/100 and 92/100 respectively) (Fig. 6.2).

Interestingly, in the phylogenetic analysis presented by Han *et al.* (2006) species belonging to the tribe Tephritini do not form a monophyletic group; it can be inferred that this could be probably due to the different geographical origin of the analyzed species (Tab. 6.4).

The clades 2, 3, 4 and 5 cluster tribes belongs to the "Lower Tephritinae" of Korneyev (1999). Our data indicate that, in agreement with Korneyev, the "Lower Tephritinae" form a paraphyletic group of tribes in contrast to the monophyletic "Higher Tephritinae".

Clade 2 corresponds to the tribe Myopitini. This is formed by species showing a close affinity both under morphological and biological aspects (Freidberg & Norrbom, 1999; Korneyev, 1999). In the phylogenetic tree inferred from the COI-tRNALeu-COII analysis, the tribe Myopitini is a sister group of the tribe Tephritini with a high statistical support (74/97) (Fig. 6.2 and Tab. 6.4). In the analyses of the 16S rDNA carried out by Han, this tribe resulted in a grouping with Noeetini (except *Ensina sonchi*) and Xyphosiini in a supported clade, but the relationships among the three tribes were not defined. From the morphological standpoint the exact relationship of the Myopitini with the rest of the subfamily remains uncertain (Freidberg & Norrbom, 1999; Korneyev, 1999). The position of the tribe Myopitini as a sister group of the tribe Tephritini offers an interesting matter of discussion deserving future speculation.

Clade 3 corresponds with the tribe Xyphosiini, which is a very restricted tribe formed by only six genera, only three of which are Palaeartic (Norrbom *et al.*, 1999) and only one of which was available for the present analysis. Recently, the inclusion in this tribe of other genera of neartic origin (genus *Naespilota*) is supported by both the morphological (Korneyev, 1999) and the molecular point of view (Han *et al.*, 2006).

Clade 4 corresponds to the tribe Noeetini (Norrbom *et al.*, 1999). This is the only tribe that was not recovered as a monophyletic group if we consider *Ensina sonchi* as a

member of this tribe. None of the inferred trees supports this clade (Fig. 6.2, 6.3 and 6.4). It is remarkable that traditionally *E. sonchi* is considered as a member of the Tribe Noeetini based only on a single synapomorphic trait (Norrbom *et al.*, 1999). In our analysis *E. sonchi* is not grouped with the rest of Noeetini as in the biomolecular analysis of Han *et al.* (2006). However, they suggest an additional molecular study as the single synapomorphic character found in *Ensina* and Noeetini (*sensu stricto*) is considered a result of a converged evolution.

Clade 5 corresponds to the Tribe Terelliini. The phylogenetic tree of the COI-tRNALeu-COII placed tribe Terelliini as a sister group of the remaining tribes with a statistically supported value (61/98) (Fig. 6.2). However in the analysis of the single 16S rDNA produced by Han *et al.* (2006) the tribe Terelliini appeared as a sister group to the "Higher Tephritinae" although this relationship was statistically poorly supported. The tribe Terelliini corresponds to a sister group of the remaining tribes also under the cladistic analysis of Korneyev (1999) which considers Terelliini as the most primitive tribe of the subfamily, resting on the presence of plesiomorphic characters. It is relevant to recall that all adult fruit flies possess a foregut diverticulum called oesophageal bulb (Girolami, 1973; 1983; Capuzzo *et al.*, 2005). Interestingly, Terelliini show a specific kind of oesophageal bulb that is morphologically intermediate between the oesophageal bulb of the remaining tribes of the subfamily and the one typical of the Trypetinae subfamily (e.g. *Rhagoletis* sp.) (Girolami, 1973).

The more satisfactory phylogeny of the COI-tRNALeu-COII gene in both clustering methods used could be a worthy indication for future phylogenetic studies of the tephritid flies, based on mitochondrial DNA. In summary, the present data, added to those of Han *et al.* (2006), could provide an important contribution to the increase of knowledge of the phylogenetic relationships among the taxa of this subfamily.

Distribution of the symbiosis among the subfamily Tephritinae

The analysis carried out in the previous work (Mazzon *et al.*, 2008) and in the present work has revealed the presence of symbiotic bacteria only in all the species of the tribe Tephritini and in species of genus *Noeeta* (tribe Noeetini). On the contrary, in all the species of Tribes Myopitini, Xyphosiini and Terelliini analyzed no symbiotic bacteria could be detected (Fig.

6.5). Thus, as postulated by Stammer (1929), and confirmed by our data, many Tephritinae flies seem not to harbor symbiotic bacteria in their midgut. It is nonetheless important to underline that the presence of uncultured bacteria in other structures different from the midgut can not be excluded.

It is remarkable that the presence of symbiotic bacteria seems clearly connected to tribe Tephritini which resulted, as concerns the west palaeartic species investigated here, a monophyletic and highly supported group (Fig. 6.2, 6.3 and 6.4). This tribe is the largest and most widespread of all tribes of the subfamily Tephritinae (Korneyev, 1999).
Conventional classification includes the tribe Tephritini in the "Higher Tephritinae" with the tribes Dithrycini, Eutretini, Schistopterini, Acrotaeniini and Tephrellini (Korneyev, 1999). The "Higher Tephritinae" is a monophyletic complex of the advanced tribes considered by Korneyev a derived clade of the "Lower Tephritinae". Future studies should ideally extend the analysis to other tribes included in the "Higher Tephritinae" also to other zoogeographical regions, to ascertain whether the presence of symbiotic bacteria is a peculiarity of the tribe Tephritini or of all the tribes of the "Higher Tephritinae".

The other remainder of the four tribes studied are included in "Lower Tephritinae". It is an aggregation of tribes which retains many primitive characters and seems paraphyletic. When analyzing the "Lower Tephritinae" we have detected symbiotic bacteria only in the tribe Noeetini and in particular in the genus *Noeeta* while no symbionts were found in *Ensina*. Symbiotic bacteria of the *Noeeta* genus are phylogenetically far from the symbiotic bacteria present in the tribe Tephritini thus denoting a different history (Mazzon *et al.*, 2008). It is known from the morphological point of view that the tribe Noeetini has not firmly established limits (Korneyev, 1999; Norrbom *et al.*, 1999). Moreover, the phylogeny of the 16S rDNA presented by Han *et al.* (2006) showed little affinity between *E. sonchi* and the *Noeeta* genus. Also in our molecular phylogenetic analyses, *E. sonchi* did not appear closely related to the *Noeeta* genus nor were any symbiotic bacteria found in *E. sonchi*. Ascertaining the presence of the symbiosis in this tribe would however require further analyses.

It follows from our results that the distribution of bacterial symbioses inside tephritid flies can be suggested as a useful tool for clarifying some debated aspects of the phylogenetic relationships among the tribes in the subfamily Tephritinae. Moreover our

study suggests that evolutionary events have apparently led the tribe Tephritinae towards a marked attitude to acquire and maintain the bacterial symbiosis. Under this aspect an interesting contribution can arise in our understanding of the biological significance of such mutualisms. The scenario can be ideally evaluated upon comparison with the situation in other tribes, in which, despite the common trait of insect development within Asteraceae flower heads, the evolution does not seem to have selected the establishment of a bacterial interaction.

Concordant evolution of subfamily Tephritinae with its symbiotic bacteria

Many host-symbiont associations represent a combination of cospeciation and host-switching (Roy, 2001; Weiblen & Bush, 2002; Ricklefs *et al.*, 2004). When studying the possibility of a coevolution between Tephritinae and their symbiotic bacteria, TreeMap software suggested a significant fit between host and symbiont trees (Fig. 6.6). However, a rather low percentage of the nodes (62.5%) of both host and symbionts was congruent. ParaFit software also suggests the presence of a global coevolution between host and symbionts, although it indicates that some symbiotic bacterial species did not cospeciate with their hosts. This would agree with the fact that, besides cospeciations, other independent events such as duplications, sorting events and host-switching could be occurred.

The TreeMap reconstructions, chosen among those with a higher number of cospeciation and a lower number of costs, suggested the presence of two main events in the subfamily with a different history (Fig. 6.6b). The most important event implies that the common ancestor of tribe Tephritini acquired its symbiotic bacteria (probably an ancestral Gammaproteobacteria close to the current *Erwinia* group), coevolving over time, suffering also some losses and acquisitions (Fig. 6.6b). These monophyletic and unculturable symbiotic bacteria have been designated "*Candidatus* Stammerula spp.", and "*Candidatus* Stammerula tephritidis" for the genus *Tephritis* spp. (Mazzon *et al.*, 2008). Of particular interest is the case arising in two species belonging to the *Campiglossa* group (*Campiglossa guttella* and *Dioxyna bidentis*) which have symbiotic bacteria, different from *Stammerula* (Fig 6.6a), but related to the free-living *Erwinia persicina* (Mazzon *et al.*, 2008). It can be

hypothesized that the presence of this kind of symbiotic bacteria could result from losses of *Stammerula* and subsequent acquisitions from free-living bacteria.

A minor event is represented by an independent occurrence which concerns the acquisition of a different symbiotic bacterium, close to the current free-living *Ewingella*, in flies of the genus *Noeeta* (Mazzon *et al.*, 2008).

A different consideration is necessary for the relationships between the symbiotic bacterium *Erwinia dacicola* (also closely related to free-living *Erwinia* sp.) and its host *Bactrocera oleae*. *B. oleae*, commonly known as olive fly and considered a major agricultural pest, belongs to the subfamily Dacinae. Interestingly, all the reconstructions suggest that olive fly has subsequently acquired its symbiont, via an horizontal transfer (Fig. 6.6b). In our view it can be inferred that the acquisition of an ancestral *Erwinia* also by *B. oleae* could represent another instance of an independent accession event. This would underline once again an attitude of these proteobacteria to undertake symbiotic relationships with fruit flies. We foresee that these aspects will deserve thorough investigation in future works.

Many bacterial species belonging to the genus *Erwinia* are considered weak pathogens that specifically invade plant tissues through lesions (Beattie, 2006). The majority of non-frugivorous tephritids spend most of their life cycle on, or within host plant tissues. This tight association with the host could explain the recurring and independent events of acquisition.

As currently known, the symbionts found in these flies are all extracellular and their host insect lifestyle offers more potential opportunities for contacts with the outer environment compared to endocellular symbionts. As described by Girolami (1973) and Mazzon *et al.* (2008), whereas in the adult flies symbiotic bacteria are located in a close environment between the peritrophic membrane and the midgut epithelium, in the larval stage bacteria are located in the intestinal blind sacs (Petri, 1909; Stammer, 1929) without the protection of the peritrophic membrane. Here symbiotic bacteria could be temporarily in contact with generic free living bacteria present in the intestinal lumen and are therefore exposed to invasion and replacement by foreign microbes. Another critical phase for possible contact with the outer environment could occur during the oviposition when the mother infects the outer side of the egg's chorion with bacteria in order to ensure the vertical transmission to the offspring as described by Petri (1909) for the olive fly and by

Stammer (1929) for the Tephritinae subfamily. In addition the feeding behavior of the tephritid flies creates other opportunities. In Tephritini there are examples of multiple species visiting or developing in flower heads of the same host plant; this could offer further chance for possible invasion and replacement by foreign bacteria and bypass the model of strict vertical transmission.

Our studies, with the support of the best available and most used coevolution analysis programs, notwithstanding the experimental and statistical limitations, showed a robust history of tandem diversification, characterized by cospeciations, losses, acquisitions and host-switching events. Factors such as insect host-symbiotic bacteria physiological compatibility, should also be taken into account for a more thorough interpretation.

As aforementioned, the ecology and biology of the hosts render them potentially susceptible to accidental horizontal transfers of the symbiont. Nevertheless, the presence of extensive congruence between insects and bacteria, indicates that also for these extracellular symbiotic partners of tephritid flies vertical transmission is the primary basis of the Tephritinae-symbiont concordant phylogeny.

Acknowledgments

We thank Bernhard Merz for his assistance with tephritids identifications. We are grateful to Enrico Negrisolo, M. Alejandra Perotti and Henk R. Braig for critical insights. We appreciated the helpful comments of two anonymous reviewers.

Taxon	Host Plants	Origin	GenBank Accession		
			Insect: COI-COII	Insect: 16S rRNA	Symbiont: 16S rRNA
OUTGROUPS					
FAMILY PLATYSTOMATIDAE					
Platystoma sp.		ITALY, Veneto, Cogollo del Cengio	GQ175826	GQ175863	-
FAMILY TEPHRITIDAE					
Subfamily Trypetinae					
Tribe Carpomyni					
Rhagoletis cerasi L.	*Prunus avium* L.	ITALY, Veneto, Torreglia	GQ175823	GQ175860	-
Rhagoletis completa Cresson	*Juglans regia* L.	ITALY, Veneto, Este	GQ175824	GQ175861	-
Subfamily Dacinae					
Tribe Ceratitidini					
Ceratitis capitata (Widemann)	-	GREECE	AJ242872†		-
Tribe Dacini					
Bactrocera oleae (Rossi)	*Olea europaea* L.	ITALY, Liguria, Imperia*	GQ175825	GQ175862	AJ586620^{+}
INGROUPS (Tephritidae)					
Subfamily Tephritinae					
Tribe Myopitini					
Myopites inulaedyssentericae Blot	*Inula crithmoides* L.	CROATIA, Istria, Rovigno (2)(~)	GQ175815	GQ175852	-
Urophora congrua Loew	*Cirsium erisithales* (Jacq.) Scop.	ITALY, Veneto, Belluno (2)	GQ175819	GQ175856	-
Urophora cuspidata (Meigen)	*Centaurea scabiosa* L.	ITALY, Friuli V.G, Fanna (5)(~)	GQ175818	GQ175855	-
Urophora quadrifasciata (Meigen)	*Centaurea jacea* L.	ITALY, Friuli V.G, Fanna (3)(~)	GQ175816	GQ175853	-
Urophora terebrans (Loew)	*Cirsium eriophorum* (L.) Scop.	ITALY, Piemonte (2)	GQ175820	GQ175857	-
Urophora stylata (Fabricius)	*Cirsium arvense* (L.) Scop.	CROATIA, Istria, Rovigno (3)(~)	GQ175817	GQ175854	-
Tribe Noeetini					
Ensina sonchi (Linnaeus)	*Sonchus* sp.	ITALY, Veneto, Legnaro (3)(~)	GQ175812	GQ175849	-
Noeeta bisetosa Merz	*Hieracium piloselloides* Vill.	ITALY, Friuli V.G, Fanna	GQ175814	GQ175851	EF469632§
Noeeta pupillata (Fallén)	*Hieracium murorum* L.	ITALY, Veneto, Cogollo del C.	GQ175813	GQ175850	EF469635§
Tribe Tephritini/Campiglossa Group					
Campiglossa doronici (Loew)	*Doronicum austriacum* Jacq.	ITALY, Veneto, Cogollo del C.	GQ175802	GQ175839	EF469636§
Campiglossa guttella Rondani	*Hieracium murorum* L.	ITALY, Veneto, Asiago	GQ175801	GQ175838	EF469637§
Dioxyna bidentis (R.-D.)	*Bidens tripartita* L.	ITALY, Friuli V.G, Fanna	GQ175800	GQ175837	EF469631§
Oxyna flavipennis (Loew)	*Achillea millefolium* L.	SLOVENIA, Kranjska Gora (2)(~)*	GQ175798	GQ175835	EF469630§
Tribe Tephritini/Sphenella Group					
Sphenella marginata (Fallén)	*Senecio alpinus* Auct.	ITALY, Veneto, Asiago	GQ175799	GQ175836	EF469629§
Tribe Tephritini/Tephritis Group					
Acanthiophilus helianthi (Rossi)	*Centaurea jacea* L.	ITALY, Friuli V.G, Fanna	GQ175795	GQ175832	EF469625§
Capitites ramulosa (Loew)	*Phagnalon saxatile* (L.) Cass.	ITALY, Liguria, Imperia	GQ175803	GQ175840	EF469628§
Tephritis arnicae (Linnaeus)	*Arnica montana* L.	ITALY, Veneto, Asiago	GQ175793	GQ175830	EF469616§
Tephritis bardanae (Schrank)	*Arctium lappa* L.	ITALY, Veneto, Foza	GQ175794	GQ175831	EF469617§
Tephritis cometa (Loew)	*Cirsium arvense* (L.) Scop.	CROATIA, Istria, Rovigno	GQ175792	GQ175829	EF469615§
Tephritis divisa Rondani	*Picris echioides* L.	ITALY, Liguria, Imperia	GQ175790	GQ175827	EF469619§
Tephritis matricariae (Loew)	*Crepis vesicaria* L.	ITALY, Veneto, Legnaro	GQ175791	GQ175828	EF469623§
Trupanea amoena (Frauenfeld)	*Reichardia picroides* (L.) Roth	ITALY, Liguria, Imperia (2)(~)*	GQ175796	GQ175833	EF469626§
Trupanea stellata (Fuessly)	*Erigeron annuus* (L.) Pers.	ITALY, Veneto, Verona	GQ175797	GQ175834	EF469627§
Tribe Terelliini					
Chaetorellia jaceae (R.-D.)	*Centaurea jacea* L.	ITALY, Friuli V.G, Fanna (2)	GQ175805	GQ175842	-
Chaetostomella cylindrica (R.-D.)	*Centaurea triumfetti* All.	ITALY, Friuli V.G, Fanna (4)(~)	GQ175804	GQ175841	-
Orellia falcata (Scopoli)	*Tragopogon orientalis* (L.) Celak	ITALY, Friuli V.G, Fanna (2)	GQ175807	GQ175844	-
Terellia colon (Meigen)	*Centaurea scabiosa* L.	ITALY, Friuli V.G, Fanna (6)(~)	GQ175806	GQ175843	-
Terellia ruficauda (Fabricius)	*Cirsium arvense* (L.) Scop.	ITALY, Friuli V.G, Fanna (2)	GQ175811	GQ175848	-
Terellia serratulae (Linnaeus)	*Cirsium pannonicum* (L. fil.) Link.	ITALY, Friuli V.G, Fanna (2)(~)	GQ175810	GQ175847	-
Terellia tussilaginis (Fabricius)	*Arctium tomentosum* Miller	ITALY, Friuli V.G, Tarvisio (2)	GQ175808	GQ175845	-
Terellia virens (Loew)	*Centaurea maculosa* Lam.	ITALY, Veneto, Rovolon (2)(~)	GQ175809	GQ175846	-
Tribe Xyphosiini					
Xyphosia laticauda (Meigen)	*Centaurea triumfetti* All.	ITALY, Friuli V.G, Fanna (2)	GQ175821	GQ175858	-
Xyphosia miliaria (Schrank)	*Carduus nutans* L.	CROATIA, Istria, Rovigno (6)(~)	GQ175822	GQ175859	-

Table 6.1 - List of taxa examined in this study, their assignments to subfamilies and tribes according with Merz (1999) and Korneyev (1999), host plant, locality and data of collection. Accession number for insect host and symbiotic bacteria when present (dash [-] indicates that no symbiotic bacteria have been detected) are reported.
The number in parentheses following insect origin is the number of individuals that were tested in order to detect the presence of unculturable symbiotic bacteria.
(~) indicates those cases whose midguts were the objects of DNA analyses.
* indicates those cases in which symbiont bacterial DNA has been extracted from samples of different origin or year with respect to Mazzon *et al.* (2008).
† Sequence from Spanos *et al.* (2000); [+]Sequences from Capuzzo *et al.* (2005); §Sequences from Mazzon *et al.* (2008). All other sequences were obtained in the current study.

Target gene	Primer name	Sequence 5'→3'	Source
16S	LR-J -12883	CTCCGGTTTGAACTCAGATC	(Xiong and Kocher, 1991)
	TV-N-14202	AGCATTTCATTTACATTGAA	(Han and McPheron, 1997)
	DFI	CATTGGGCAGGTYARACT	(this study)
	DFI2	GATTTATAGGGTCTTCTCGTC	(this study)
	DR*	GATGTACCGGAAGGTGTATCT	(this study)
	DRI*	GTTATTCGTTTATAAAGRTATC	(this study)
	LRN13398	CGCCTGTTTAACAAAAACAT	(Simon *et al.*, 1994; 2006)
	SR – N14220*	ATATG(CT)ACA(CT)ATTGCCCGTC	(Simon *et al.*, 1994; 2006)
	N1 – J12261m	TACTTCGTAAGAAATTGTTTGAGC	(Simon *et al.*, 1994; 2006)
	SR-J-13342	CCTTTGTAC (AG)GT CAA AAT AC(CT) GC	(Simon *et al.*, 1994; 2006)
	SR-N14745*	GTGCCAGCAG(CT)(CT)GCGGTTA(AGCT)AC	(Simon *et al.*, 1994; 2006)
	SR-N14588*	AAACTAGGATTAGATACCCTATTAT	(Simon *et al.*, 1994; 2006)
COI-LeutRNA-COII	C1-J-2195	TTGATTTTTTGGTCATCCAGAAGT	(Simon *et al.*, 1994; 2006)
	TKN3796	ACTATAAAATGGTTTAAGAG	(Simon *et al.*, 1994; 2006)
	LCO1490	GGTCAACAAATCATAAAGATATTGG	(Simon *et al.*, 1994; 2006)
	C1-J-2183	CAACATTTATTTTGATTTTTTGG	(Simon *et al.*, 1994; 2006)
	C1 -J -2792	ATACCTCGACGTTATTCAGA	(Simon *et al.*, 1994; 2006)
	TL2-N-3014	TCCATTGCACTAATCTGCCATATTA	(Simon *et al.*, 1994; 2006)
	C1-J-2441	CCTACAGGAATTAAAATTTTTAGATGATTA	(Simon *et al.*, 1994; 2006)

Table 6.2 - Oligonucleotide primer sequences used in the polymerase chain and sequencing reactions for 16S rDNA and COI-tRNA-Leu-COII. Primers indicated with an asterisk (*) were used only in polymerase chain reaction.

Symbiont – Host	*ParaFitLink1*	*ParaFitLink2*
Stammerula tephritidis - *Tephritis divisa*	0.05*	0.05*
Stammerula tephritidis - *Tephritis matricariae*	0.09	0.08
Stammerula tephritidis - *Tephritis cometa*	0.09	0.08
Stammerula tephritidis - *Tephritis arnicae*	0.13	0.11
Stammerula tephritidis - *Tephritis bardanae*	0.04*	0.04*
Stammerula sp. - *Acanthiophilus helianthi*	0.16	0.14
Stammerula sp. - *Trupanea amena*	0.10	0.09
Stammerula sp. - *Oxyna flavipennis*	0.06	0.06
Stammerula sp. - *Sphenella marginata*	0.75	0.74
Stammerula sp. - *Campiglossa doronici*	0.18	0.16
Stammerula sp. - *Capitites ramulosa* and *Trupanea stellata*	0.09	0.08
Stammerula sp. - *Capitites ramulosa* and *Trupanea stellata*	0.06	0.06
Dioxyna bidentis symbiont - *Dioxyna bidentis*	0.57	0.55
Campiglossa guttella symbiont - *Campiglossa guttella*	0.37	0.35
Noeeta bisetosa symbiont - *Noeeta bisetosa*	0.03*	0.02*
Noeeta pupillata symbiont - *Noeeta pupillata*	0.03*	0.03*
Erwinia dacicola - *Bactrocera oleae*	0.02*	0.02*
Global Test	0.01*	

Table 6.3 - Results from the ParaFit analysis: the test was used to asses the null hypothesis of independent evolution of hosts and symbionts and to test the significance of each host-symbiont link (17 host-symbiont links) using ParaFitLink1 and ParaFitLink2 statistics. Probabilities are computed after 999 random permutations. *Significant association ($P \leq 0.05$).

Clade n.	COI-tRNALeu-COII	16S	COI-tRNALeu-COII +16S	16S (Han *et al.*, 2006)
(1) Tephritini tribe[a]	74/100	94/100	100/100	85/100/99 (with other tribes)
(2) Myopitini tribe[a,b]	100/100	100/100	100/100	99/100/99
(3) Xiphosiini tribe[a,b]	100/100	100/100	100/100	95/100/99 (with *Naespilota)*
(4) Noeetini tribe[c]	-/-	nr	-/-	nr
(5) Terelliini tribe[a]	97/100	100/100	100/100	100/100/99 (except *Naespilota)*
(1A) *Tephritis* group[d]	100/100	61/-	100/100	99/100/99
(1B) *Campiglossa* group[a,c]	92/100	nr	59/99	nr
(1C) *Sphenella*+*Campiglossa* group	91/100	nr	100/100	-/94/- (low support)
(1+2) Tep+Myo	74/97	nr	-/-	nr
(1+2+3+4) Tep+Myo+Xip+Noe[a]	61/98	nr	nr	nr
(3+4) Xip+Noe	-/100	nr	-/-	nr
Tephritinae subfamily[a,c]	100/100	87/100	100/100	90/100/99

Table 6.4 - Statistical support values for the nodes recognized in the trees inferred from COI-tRNALeu-COII, 16S rDNA and combined data set (COI-tRNALeu-COII+16S rDNA). Clade numbers are reported in Fig. 6.2, 6.3 and 6.4. The first number indicates the Bp from the bootstrap test of the ML analysis and the second number indicates the Pp from Bayesian analysis. The last column reports statistical support values obtained by Han *et al.* (2006): the first number indicates Pp, the second one indicates Bp and the third one indicates Pc (ME analysis). Superscript letters indicate clades recognized by traditional classification: [a](Korneyev, 1999); [b](Friedberg and Norrbom, 1999); [c](Norrbom *et al.*, 1999); [d](Merz, 1999).

Fig. 1

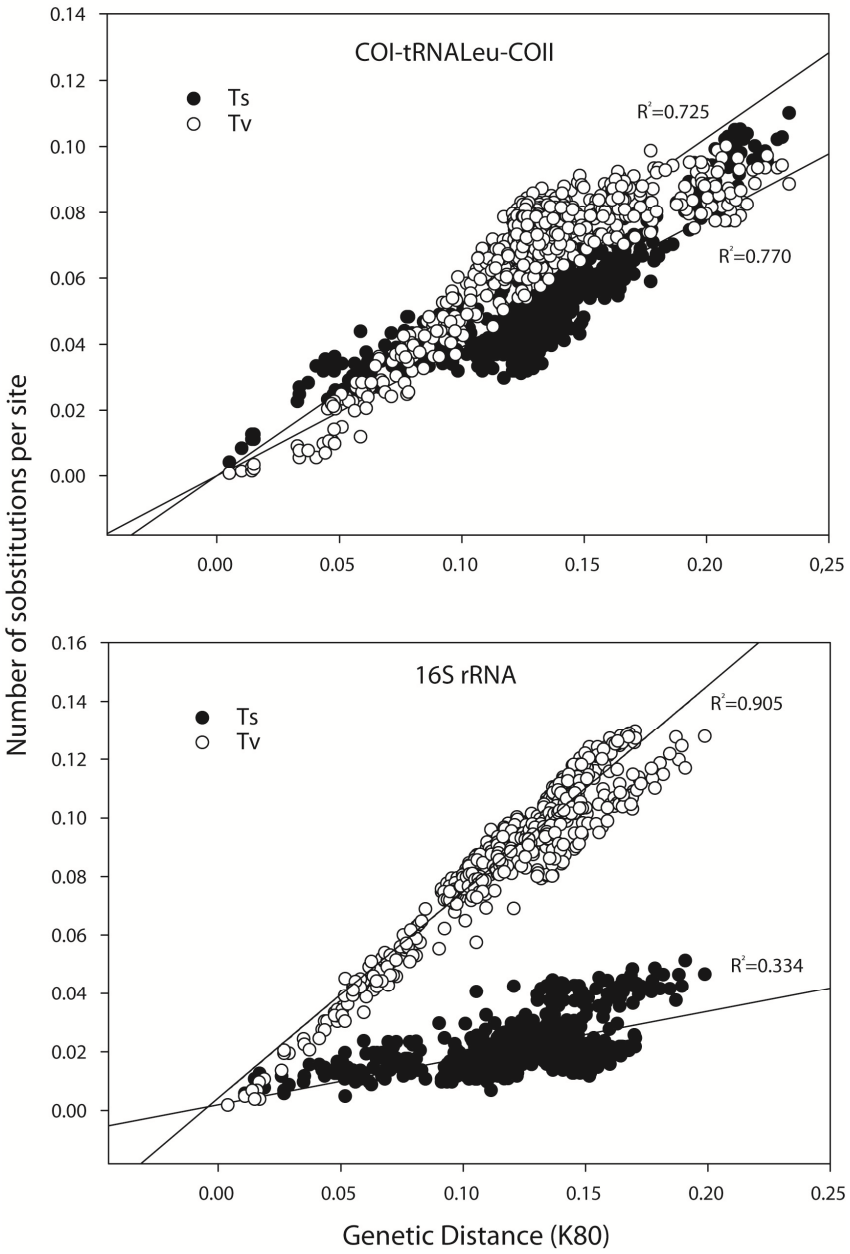

Fig. 6.1 - Saturation plots for COI-tRNALeu-COII and 16S rDNA gene fragments. The genetic distance (Kimura two-parameter model; K80) was plotted against the number of transitions (Ts) and transvertions (Tv).

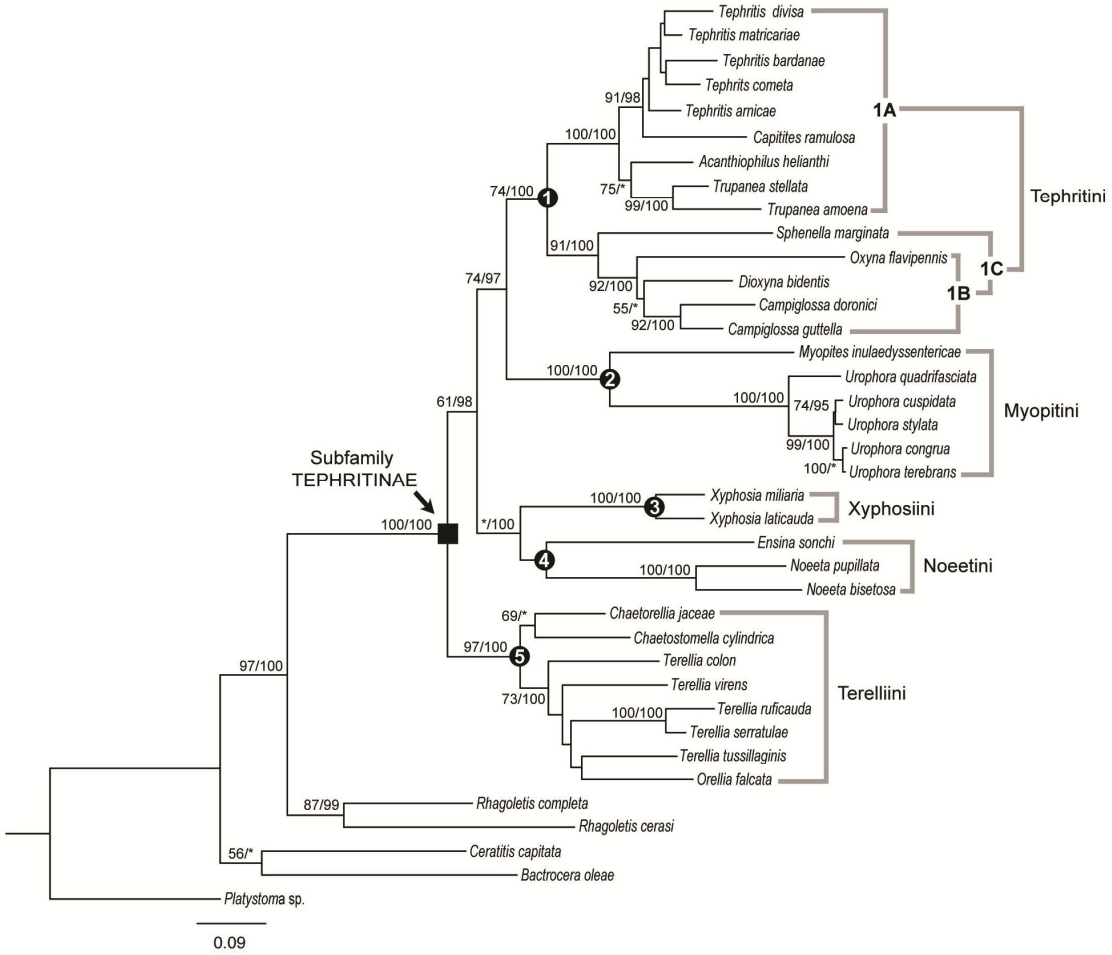

Fig. 6.2 - Phylogenetic reconstruction on the basis of the COI-tRNALeu-COII gene sequences of the subfamily Tephritinae. The nodes representing taxonomic groups are indicated with a number. The first number on each branch indicates the value of bootstrap probability (Bp) from the bootstrap test of the ML analysis and the second number indicates the value of posterior probability (Pp) from Bayesian analysis. Asterisks indicate bootstrap probabilities lower than 50% and posterior probabilities lower than 95%.

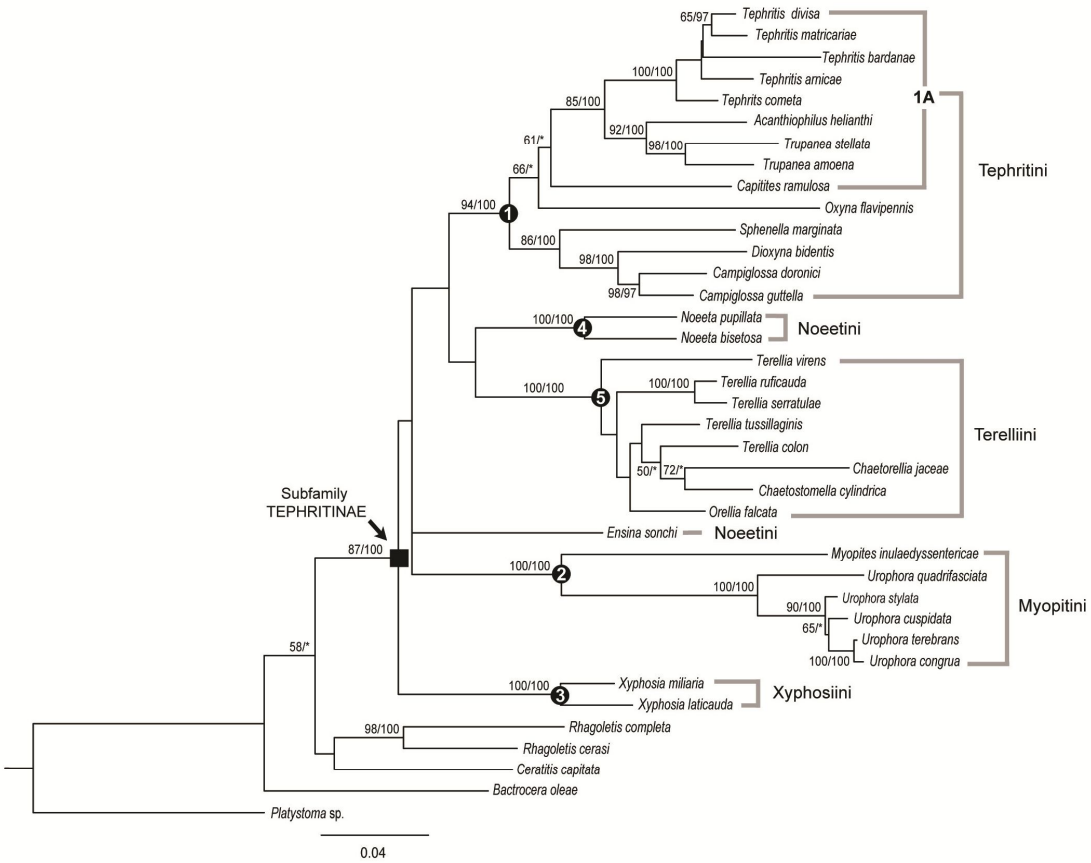

Fig. 6.3 - Phylogenetic reconstruction on the basis of the 16S rDNA gene sequences of the subfamily Tephritinae. The nodes representing taxonomic groups are indicated with a number. The first number on each branch indicates the value of bootstrap probability (Bp) from the bootstrap test of the ML analysis and the second number indicates the value of posterior probability (Pp) from Bayesian analysis. Asterisks indicate bootstrap probabilities lower than 50% and posterior probabilities lower than 95%.

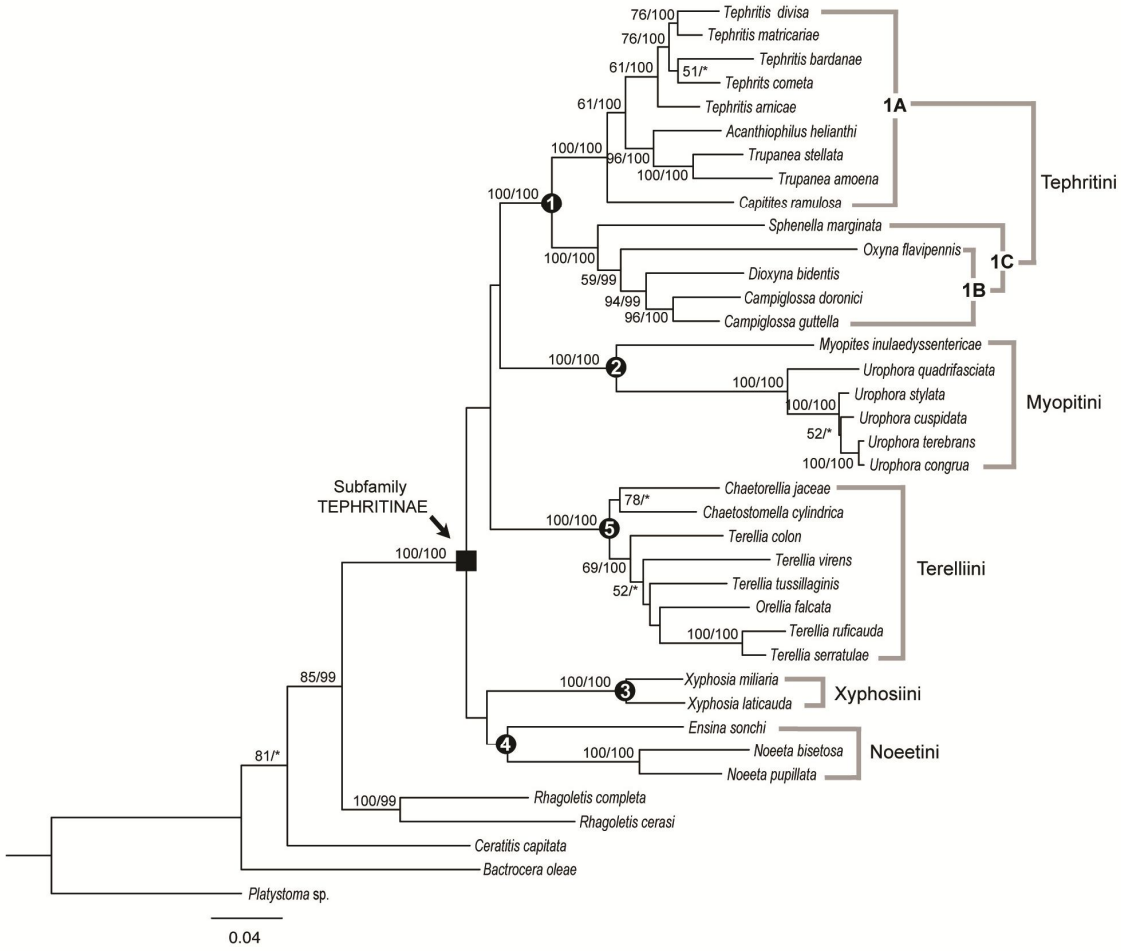

Fig. 64 - Phylogenetic reconstruction on the basis of the combined data set (16S rDNA + COI-tRNALeu-COII) of the subfamily Tephritinae. The nodes representing taxonomic groups are indicated with a number. The first number on each branch indicates the value of bootstrap probability (Bp) from the bootstrap test of the ML analysis and the second number indicates the value of posterior probability (Pp) from Bayesian analysis. Asterisks indicate bootstrap probabilities lower than 50% and posterior probabilities lower than 95%.

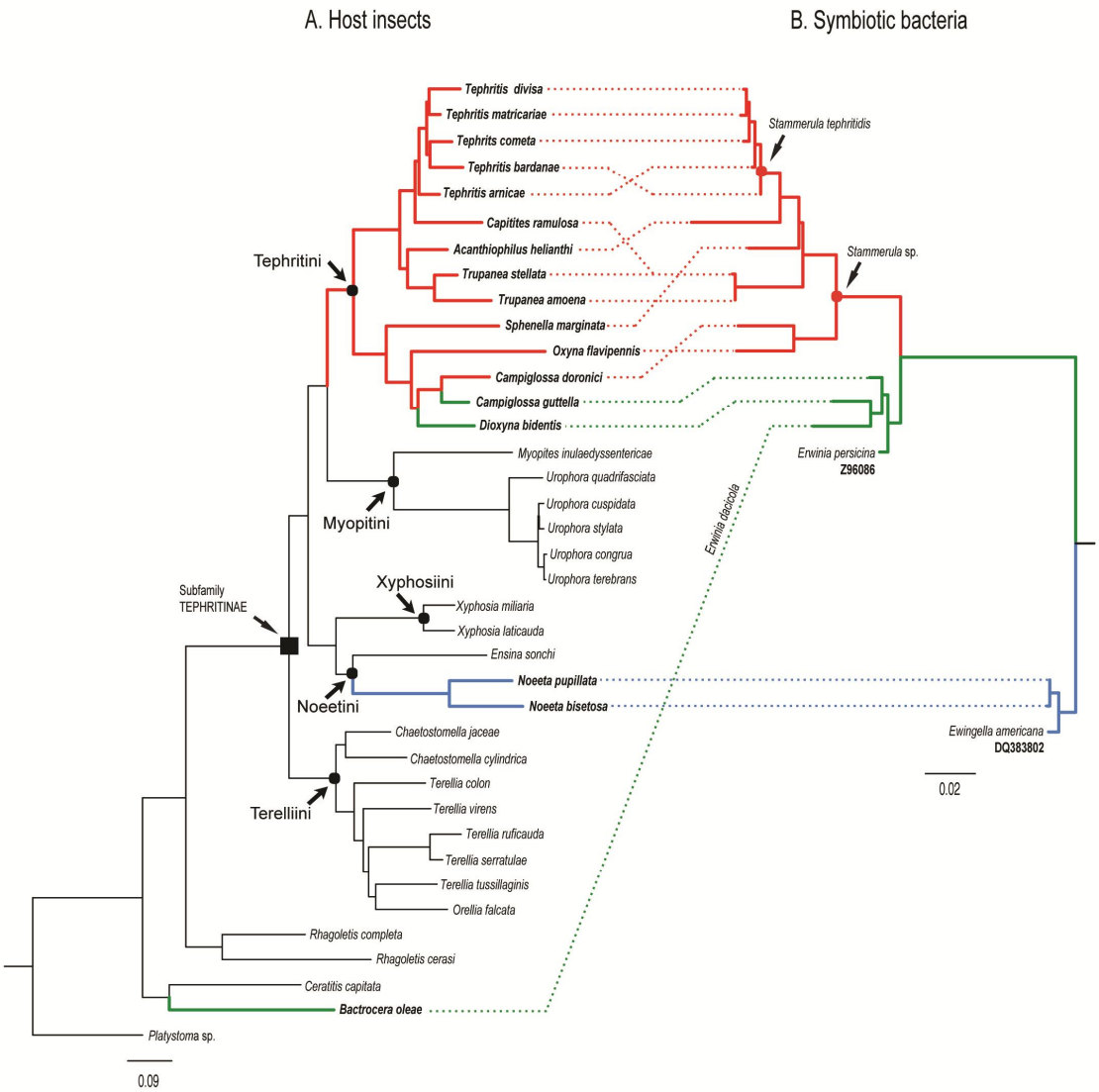

Fig. 6.5 - Phylogenetic reconstruction on the basis of the COI-tRNALeu-COII gene sequences of the subfamily Tephritinae (left) and their symbiotic bacteria (right). The symbionts tree was obtained using Maximum-Likelihood method on 16S rDNA dataset through PHYML. Colored branches corresponds to taxa where symbiotic bacteria has been detected. Dashed lines depict the observed host-symbiont associations.

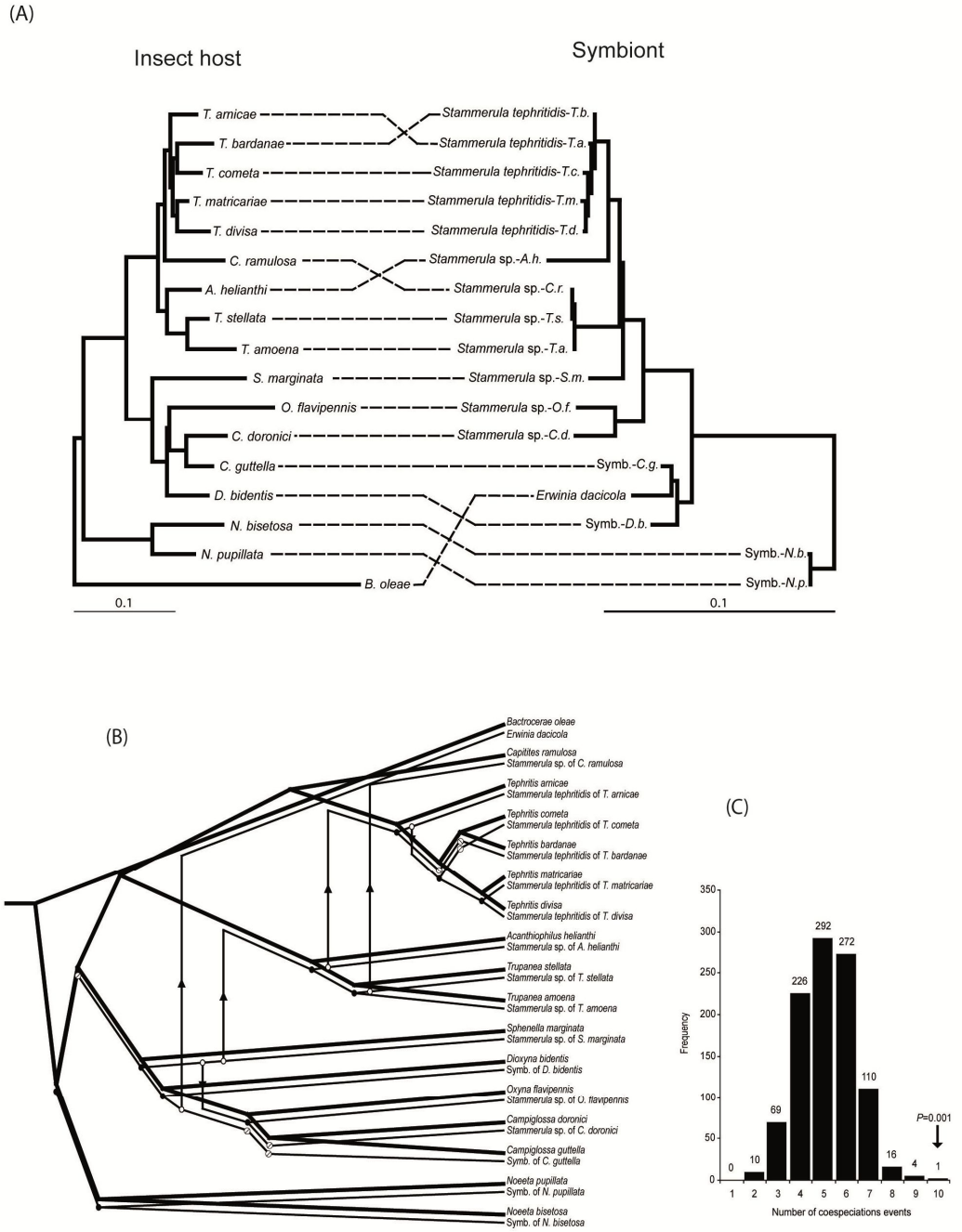

Fig. 6.6 - (A) Tanglegram for tephritid host insects and their symbiotic bacteria. Each symbiont is connected to its host by a line. Maximum likelihood trees based on the COI-tRNALeu-COII for the Tephritinae host (left) and the 16S rDNA of their symbiotic bacteria (right).
(B) One of the 21 possible coevolutionary reconstructions (maximizing the number of cospeciations) for the two trees showed in A processed by the program TreeMap v2.02β.
Legend: Bold lines have been used for hosts tree and thin lines for symbionts tree; (●) cospeciation event; (○) duplication event; (⊗) sorting event (losses); (→) host switch.

(C) Randomization test, generate for TreeMap, for the phylogenetic congruence between host and symbiont. Frequency distribution of the number of cospeciations events in random associations (sample size 1000).

Conclusion

The present thesis surveyed the bacterial relationships within some fruit flies species. Different aspects have been taken into account: the kind of relationship (facultative/obligate), the microflora composition and its transmission, the genetic variability of hereditable symbiont.

It is important to do not forget that the original purpose of the fruit flies-bacteria interaction studies was mostly economical. In fact the presence of bacteria directly affects the quality and the performance of adult reared on artificial media to obtain sterile male (SIT), used as a pest control methods. The current studies together with further researches on the biological relationships could help the improving of this technique.

In an overall vision the fruit flies-bacteria interaction can be divided into obligatory hereditable or associated relationships. To date the presence of obligated and vertically transmitted symbiotic bacteria is restricted to the olive fly and some tribes of the Tephritinae subfamily. In the rest of the cases, bacteria are facultative within their host and not hereditable and are usually called “associated” (Drew & Lloyd, 1987). Evidence for this latter pattern has been reported in this thesis for *Rhagoletis cerasi* and probably for *R. completa.* According with traditional culturable-dependent and independent results it seems that the female does not transmit bacteria to the eggs, and bacteria are likely to be acquired from the environment by larvae and mature larvae before pupation.

From a phylogenetic point of view the results of this work rise a question about the meaning of vertical bacterial transmission confined within some tephritid groups. For example up today in the subfamily Tephritinae the presence of an obligatory hereditable symbiosis is confined to Tephritini and Noeetini tribes (Mazzon *et al*., 2008, 2010). It will be interesting to investigate the symbiosis biological implications by mean of the compared biology and morphology.

The mechanisms and the way implied in the vertical bacterial transmission from larval to adult stages have been questioned and remain undefined to date. In order to achieve this goal the use of *gfp*-tagged bacteria in this study turned out to be an useful tool to detect the fate and the exact structures harbouring bacteria, mainly in the preimaginal stages. The first preliminary analyses on *R. completa* are encouraging, and the experimentations could be

extended to other tephritid species, as they also serve as a vector of detrimental genes for the biological control of the insect pest.

In addition to the above mentioned purposes, the study of the bacterial interactions could be useful within the study of population genetics of the insect hosts, as showed in this thesis for the olive fly symbiont, *Candidatus* Erwinia dacicola. *Bactrocera oleae* is a pest characterize by extensive gene flow, mainly in the Mediterranean populations (Augustinos *et al.,* 2005; Nardi *et al.,* 2005, 2010; Zygouridis *et al.,* 2009) and our analyses on mitochondrial data confirmed this hypothesis. On the other hand its symbiont haplotypes revealed a completely different scenario: considering only the Mediterranean basin only two *E. dacicola* haplotypes have been found and they are not randomly distributed. Moreover the Pakistani olive fly population harboured a third haplotype, with mixed features between the Mediterranean ones.

We think that the combined analyses of host and symbiont variability have the potential to provide new helpful information on the colonization routes of the olive fly and especially to understand the long co-evolutionary life history with its symbiont. Thus, this kind of approach could be extended to other arthropod-symbiont interactions having similar pattern of symbiotic relationship.

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.

Baker R.H. & DeSalle R. 1997. Multiple sources of character information and the phylogeny of Hawaiian drosophilids. *Syst. Biol.* **46:**654-73.

Baumann P. & Moran N.A. 1997. Non-cultivable microorganisms from symbiotic associations of insects and other hosts. *Anton. Leeuw. Int. J. G.* **72:**39–48.

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

Beattie G.A. 2006. Plant-associated bacteria: survey, molecular phylogeny, genomics and recent avances. Pp. 1-56. In: Plant-Associated Bacteria. S.S. Gnanamanickam (ed.), Springer. Printed in the Netherlands.

Behar A., Yuval B., Jurkevitch E. 2005. Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly *Ceratitis capitata*. *Mol. Ecol.* **14:**2637–2643.

Behar A., Jurkevitch E., Yuval B. 2008a. Bringing back the fruit into fruit fly-bacteria interactions. *Mol. Ecol.* **17:**1375–1386.

Behar A., Yuval B., Jurkevitch E. 2008b. Gut bacterial communities in the Mediterranean fruit fly (*Ceratitis capitata*) and their impact on host longevity. *J. Insect Physiol.* **54:**1377-1383.

Behar A., Ben-Yosef M., Lauzon C.R., Yuval B., Jurkevich E. 2009. Structure and function of the bacterial community associated with the Mediterranean fruit fly. In: Bourtzis K, Miller T (eds). Insect Symbiosis, Vol. 3.CRC Press: Boca Raton, pp 251–271.

Belcari A., Girolami V., Rivosecchi L., Zaitzev V.F. 1995. Diptera Tephritoidea. In: Minelli A., Ruffo S. end La Posta S. (eds.), Checklist delle specie della fauna italiana, 72. Calderini, Bologna.

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.

Blando S. & Mineo G. 2006. Generazioni primaverili della mosca delle olive sull'Olivastro (*Olea europaea* var. *Sylvestris* brot.) In sicilia. *Boll. Zool. Agra. Bachic. Ser* ii.**38:**173-178.

Blümel S. & Masalmeh M.A. 1989. Primary investigation on the microbial endoflora of the European cherry fruit fly, *Rhagoletis cerasi* L. (Trypetidae) pp. 339-344. Edited by R. Cavalloro. Proceedings of the CEC/IOBC International Symposium. Rome April 1987.

Breton C., Tersac M., Berville´ A. 2006. Genetic diversity and gene flow between the wild olive (oleaster, *Olea europaea* L.) and the olive: several Plio-Pleistocene refuge zones in the Mediterranean basin suggested by simple sequence repeats analysis. *J. Biogeogr.* **33:**1916–1928.

Buchner P. 1965. Endosymbiosis of Animals with Plant Microorganisms. New York: Interscience Publishers, Inc.

Buckley T.R., Simon C., Flook P.K., Misof P.K. 2000. Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. *Insect. Mol. Biol.* **9:**565–580.

Cannone J.J., Subramanian S., Schnare M.N. *et al.* (14 co-authors), 2002. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics.* **3:**2.

Capuzzo C., G. Firrao, L. Mazzon, A. Squartini, 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.

Carrión Y., Ntinou M., Badal E. 2010. *Olea europaea* L. in the North Mediterranean Basin during the Pleniglacial and the Early–Middle Holocene. *Quaternary Sci. Rev.* **29:** 952–968.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17:**540–552.

Charleston M.A. 1998. Jungles: A new solution to the host/parasite phylogeny reconciliation problem. *Math. Biosci.* **149:**191–223.

Charleston M.A. & Page R.D.M. 2002. TREEMAP (v2.0) Applications for Apple Macintosh. University of Oxford, Oxford.

Chen X.A., Li S., Aksoy S. 1999. Concordant evolution of a symbiont with its host insect species: molecular phylogeny of genus *Glossina* and its bacteriome-associated endosymbiont, *Wigglesworthia glossinidia*. *J. Mol. Evol.* **48:**49–58.

Chen R.T., Tsai C.F., Tzeng W.N. 2009. 16S and 28S rDNA sequences in phylogenetic analyses of freshwater prawns (macrobrachium bate, 1868) from Taiwan. *J. Crustacean Biol.* **29:**400-412.

Clark M.A., Moran N.A., Baumann P., Wernegreen J.J. 2000. Cospeciation between bacterial endosymbionts (*Buchnera*) and a recent radiation of aphids (*Uroleucon*) and pitfalls of testing for phylogenetic congruence. *Evolution* **54:**517– 525.

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

Cole J.R., Wang Q., Cardenas E., Fish J., Chai B., Farris R.J., Kulam-Syed-Mohideen A.S., McGarrell D.M., Marsh T., Garrity G.M., Tiedje J.M. 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* 37 (Database issue): D141-D145; doi: 10.1093/nar/gkn879.

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

Daser U. & Brandl R. 1992. Microbial gut floras of eight species of tephritids. *Biol. J. Linn. Soc.* **45:**155-165.

de Lorenzo V., Herrero M., Jakubzik U., Timmis K.N. 1990. Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *J. Bacteriol.* **172:**6568-72.

Dillon R.J. & Dillon V.M. 2004. The gut bacteria of insects: Nonpathogenic Interactions. *Annu. Rev. Entomol.* **49:**71-92.

Donovan S.E., Purdy K.J., Kane M.D., Eggleton P. 2004. Comparison of Euryarchaea strains in the guts and food-soil of the soil-feeding termite *Cubitermes fungifaber* across different soil types. *Appl. Environ. Microbiol.* **70:**3884–3892.

Drew R.A.I. & Lloyd A.C. 1987. Relationship of fruit-flies (Diptera, Tephritidae) and their bacteria to host plants. *Ann. Entomol. Soc. Am.* **80:**629-636.

Drew R.A.I. & Lloyd A.C. 1991. Bacteria in the life cycle of tephritid fruit flies, p. 441-465. In: Barbosa, P., Krischik, V.A., Jones, C.G., eds., Microbial mediation of plant-herbivore interactions. John Wiley & Sons, New York. 625 p.

Dupanloup I. Schneider S. Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* **11:**2571–2581.

Duso C. 1991. Sulla comparsa in Italia di un Tefritide neartico del noce: *Rhagoletis completa* Cresson (Diptera: Tephritidae). *Bollettino di Zoologia Agraria e di Bachicoltura* **23**:203-209.

Duso C. & Dal Lago G. 2006. Life cycle, phenology and economic importance of the walnut husk fly *Rhagoletis completa* Cresson in northern Italy. *Ann. Soc. Entomol. Fr.* **42:**245-254.

Eberl L., Schukze R., Ammendola A., Geisenberger O., Erhart R., Sternberg C., Molin S., Amann R. 1997. Use of green fluorescent protein as a marker for ecological studies of activated sludge communities. *FEMS Microbiol. Lett.* **149:**77-83.

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., Laval G., Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinformatics Online* **1:**47-50.

Favia G. 2008. Bacteria of the Genus *Asaia*: A potential paratransgenic weapon against malaria. *Adv. Exp. Med. Biol.* **627:**49-59.

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.

Freidberg A. & Norrbom A.L. 1999. A generic reclassification and phylogeny of the tribe Myopitini (Tephritidae). Pp. 581-637. In Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior. (ed. by M. Aluja and A. L. Norrbom). CRC Press, Boca Raton.

Fu Y.X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147:**915–925.

Fukatsu T. & Hosokawa T. 2002. Capsule-transmitted gut symbiotic bacterium of the Japanese common plataspid stinkbug, *Megacopta punctatissima*. *Appl. Environ. Microbiol.* **68:**389–396.

Gil R., Latorre A., Moya A. 2004. Bacterial endosymbionts of insects: insights from comparative genomics. *Environ. Microbiol.* **6:**1109–1122.

Girolami V. 1973. Reperti morfo-istologici sulle batterio simbiosi 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, p. 74-76. *In* R. Cavalloro (ed.) Fruit flies of economic importance. Balkema, Rotterdam.

Girolami V. 1986. Mediterranean fruit fly associated bacteria: transmission and larval survival p. 135-146. *In* M. Mangel (ed.) Pest Control: Operations and Systems Analysis in Fruit Fly Management, NATO ASI Series, Vol. G 11. Springer-Verlag, Berlin.

Girolami V. & 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.

Goldman N., Anderson J.P., Rodrigo A.G. 2000. Likelihood-based tests of topologies in phylogenetics. *Syst. Biol.* **49:**652–670.

Gruwell M.E., Morse G.E., Normak B.B. 2007. Phylogenetic congruence of armored scale insects (Hemiptera: Diaspididae) and their primary endosymbionts from the Phylum Bacteriodetes. *Mol. Phylogenet. Evol.* **44:**267-280.

Guindon S. & Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52:**696-704.

Han H.Y. & McPheron B.A. 1997. Molecular phylogenetic study of Tephritidae (Insecta: Diptera) using partial sequences of the mitochondrial 16S ribosomal DNA. *Mol. Phylogenet. Evol.* **7:**17−32.

Han H.Y. & McPheron B.A. 1999. Nucleotide sequence data as a tool to test phylogenetic relationships among higher groups of Tephritidae: a case study using mitochondrial ribosomal DNA. Pp. 115−132. In Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior, (ed. by M. Aluja and A. L. Norrbom). CRC Press, Boca Raton.

Han H.Y., Ro K.E., McPheron B.A. 2006. Molecular Phylogeny of the Subfamily Tephritinae (Diptera: Tephritidae) based on mitochondrial 16S rDNA sequences. *Mol. Cells* **22:**78-88.

Han H.Y. & Ro K.E. 2009. Molecular phylogeny of the family Tephritidae (Insecta: Diptera): new insight from combined analysis of the mitochondrial 12S, 16S, and COII genes. *Mol. Cells* **27:**55-66.

Hosokawa T., Kikuchi Y., Nikoh N., Shimada M., Fukatsu T. 2006. Strict host-symbiont cospeciation and reductive genome evolution in insect gut bacteria. *Plos Biol.* **4:**e337.

Howard D.J., Bush G.L., Breznak J.A. 1985. The evolutionary significance of bacteria associated with *Rhagoletis*. *Evolution* **39:**405-417.

Huber T., Faulkner G., Hugenholtz P. 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20:**2317–2319.

Huelsenbeck J.P. & Rannala B. 2004. Frequentist properties of Bayesian Posterior Probabilities of phylogenetic trees under simple and complex substitution models. *Syst. Biol.* **53:**904-913.

Hunt D.E., Klepac-Ceraj V., Acinas S.G., Gautier C., Bertilsson S., Polz M.F. 2006. Evaluation of 23S rRNA PCR primers for use in phylogenetic studies of bacterial diversity. *Appl. Environ. Microb.* **72:**2221–2225.

Husseneder C. & Grace J.K. 2005. Genetically engineered termite gut bacteria (*Enterobacter cloacae*) deliver and spread foreign genes in termite colonies. *Appl. Microbiol. Biotechnol.* **68:**360-367.

Jesus J., Harris D.J., Brehm A. 2007. Relationships of Afroablepharus Greer, 1974 skinks from the Gulf of Guinea islands based on mitochondrial and nuclear DNA: Patterns of colonization and comments on taxonomy. *Mol. Phylogenet. Evol.* **45:**904-914.

Jobb G., Haeseler A.Von, Strimmer K., 2004. TREEFINDER: A powerful graphical analysis environment for molecular phylogenetics. *BMC Evol. Biol.* **4:**18-26

Jow H., Gowri-Shankar V., Guillard B. 2005. PHASE: a software package for phylogenetics and sequence evolution. http://www.bioinf.man.ac.uk/resources/phase.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences *J. Mol. Evol.* **16:**111-120.

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

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.

Lauzon C.R., Sjogren R.E., Wright S.E. Prokopy R.J. 1998. Attraction of *Rhagoletis pomonella* (Diptera: Tephritidae) flies to odor of bacteria: apparent confinement to specialized members of enterobacteriaceae. *Environ. Entomol.* **27:**853–857.

Lauzon C.R., 2003. Symbiotic relationships of tephritids. In: Insect symbiosis v.1. Ed. by Kostas Bourtzis and Thomas A. Miller. 115–129.

Lauzon C.R., McCombs S.D. Potter S.E. 2009. Establishment and vertical passage of *Enterobacter* (Pantoea) *agglomerans* and *Klebsiella pneumoniae* through all life stages of the Mediterranean fruit fly. *Ann. Entomol. Soc. Am.* **102:**85-95.

Lecointre W.P. & Deleporte P. 2005. Total evidence requires exclusion of phylogenetically misleading data. *Zool. Scr.* **34:**101-117.

Legendre P. & Anderson M.J. 1998. Program DistPCoA. Département de Sciences Biologiques, Université de Montréal, 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. *Mol. Phylogenet. Evol.* **45:**997-1013.

Lloyd A.C., Drew R.A.I., Teakle D.S., Hayward A.C. 1986. Bacteria associated with some Dacus species (Diptera: Tephritidae) and their host fruit in Queensland. *Aust. J. Biol. Sci.* **39:**361-368.

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.

Magri D., Vendramin G.G., Comps B., Dupanloup I., Geburek T., Gömöry D., Latałowa M., Litt T., Paule L., Roure J.M., Tantau I., Knaap W.O., Petit R.J., Beaulieu J.L. 2006. A new scenario for the quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytol.* **171:**199–221.

Marchini D., Marri L., Rosetto M., Manetti A.G.O., Dallai R.1997. Presence of antibacterial peptides on the laid egg chorion of the medfly *Ceratitis capitata*. *Biochem. Biophys. Res. Commun.* **240:**657–663.

Marchini D., Rosetto M., Dallai R., Marri L. 2002. Bacteria associated with the oesophageal bulb of the medfly *Ceratitis capitata* (Diptera: Tephritidae). *Curr. Microbiol.* **44:**120-124.

Mazzon L., Piscedda A., Simonato M., Martinez-Sañudo 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. Micr.* **58:**1277–1287.

Mazzon L., Martinez-Sañudo 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.

Merz B. 1991. *Rhagoletis completa* Cresson und *Rhagoletis indifferens* Curran zwei wirtschaftlich bedeutende nordamerikanische Früchtfliegen, neu für Europa (Diptera: Tephritidae). *Mitt. Schweiz. Entomol. Ges.* **64:**55-57.

Merz B. 1994. Diptera: Tephritidae. Insecta Helvetica Fauna 10, 198 p.

Merz B. 1999. Phylogeny of the Palaearctic and Afrotropical genera of the Tephritis group (Tephritinae: Tephritini). Pp. 629−669. In: Fruit Flies (Tephritidae): Phylogeny and Evolution of Behavior (ed. by M. Aluja and A. L. Norrbom), CRC Press, Boca Raton.

Miller J.H. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

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

Moran N.A., 2001. The Coevolution of Bacterial Endosymbionts and Phloem-Feeding Insects. *Ann. Mo. Bot. Gard.* **88:**35-44.

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

Moran N.A., Russell J.A., Koga R., Fukatsu T. 2005. Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Appl. Environ. Microbiol.* **71:**3302–3310.

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.

Nardi F., Carapelli A., Dallai R., Frati F. 2003. The mitochondrial genome of the olive fly *Bactrocera oleae*: two haplotypes from distant geographic locations. *Insect Mol. Biol.* **12:**605–611.

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., Vontas J.G., Dallai R., Roderick G.K., Frati F. 2006. Geographical distribution and evolutionary history of organophosphate resistant Ace alleles in the olive fly (*Bactrocera oleae*). *Insect Biochem. Mol. Biol.* **36:**593–602.

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.

Naylor G.J.P. & Brown W.M. 1998. Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. *Syst. Biol.* **47:**61–76.

Norrbom A.L., Carroll L.E., Thompson F.C., White I.M., Freidberg A. 1999. Status of knowledge. Pp. 9-48. In: Thompson, F.C. (ed.), Fruit fly expert identification system and systematic information database. Backhuys Publishers, Leiden, Myia pp. 9-48.

Ochando M.D. & Reyes A. 2000. Genetic population structure in olive fly *Bactrocera oleae* (Gmelin): gene flow and patterns of geographic differentiation. *J. Appl. Entomol.* **124:**177–183.

Ochando M.D., Reyes A., Callejas C., Segura D., Fernandez P. 2003. Molecular genetic methodologies applied to the study of fly pests. *Trends Entomol.* **3:**73–85.

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

Page R.D.M., Cruickshank R.H., Dickens M., Furness R.W., Kennedy M., Palma R.L., Smith V.S. 2004. Phylogeny of "Philoceanus complex" seabird lice (Phthiraptera: Ischnocera) inferred from mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* **30:**633-652.

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.

Panchal M. 2007. The automation of nested clade phylogeographic analysis. *Bioinformatics* **23:**509–510.

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.

Pavari A. (1916). Studio preliminare sulla coltura delle specie forestali esotiche in Italia. *Ann. R. Ist. Sup. For. Naz.* I, 159-379.

Peloquin J.J., L. Kuzina, C. Lauzon and T. Miller. 2000. Transformation of internal extracellular bacteria isolated from *Rhagoletis completa* Cresson gut with enhanced green fluorescent protein. *Curr. Microbiol.* **40:**367–371.

Peloquin J.J., Lauzon C.R., Potter S., Miller T.A. 2002. Transformed bacterial symbionts re-introduced to and detected in host gut. *Curr. Microbiol.* **45:**41-45.

Petri L. 1909. Ricerche sopra i batteri intestinali della mosca olearia. Memorie della Regia Stazione di Patologia Vegetale di Roma.

Posada D. & Crandall K.A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817-818.

Posada D., Crandall K.A., Templeton A.R. 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Mol. Ecol.* **9:**487–488.

Posada D. & Buckley T.R. 2004. Model selection and model averaging in phylogenetics: advantages of AIC and BAyesian approaches over likelyhood ratio tests. *Syst. Biol.* **53:**793-808.

Posada D., Crandall K.A., Templeton A.R. 2006. Nested Clade Analysis statistics. *Mol. Ecol. Notes* **6:**590-593.

Rani A., Sharma A., Rajagopal R., Adak T., Bhatnagar R.K. 2009. Bacterial diversity analysis of larvae and adult midgut microflora using culture-dependent and culture-independent methods in lab-reared and field-collected *Anopheles stephensi*-an Asian malarial vector. *BMC Microbiol.* **9**: 96

Ratner S.S. & Stoffolano J.G. 1984. Ultrastructural changes of the esophageal bulb of the adult female apple maggot, *Rhagoletis pomonella* (Walsh) (Diptera: Tephritidae). *Int. J. Insect. Morphol. Embryol.* **13:**191-208.

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

Ricklefs R.E., Fallon S.M., Bermingham E. 2004. Evolutionary relationships, cospeciation, and host switching in avian malaria parasites. *Syst. Biol.* **53:**111–119.

Rokas A., Williams B.L., King N., Carroll S.B. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature,* **425:**798-804.

Ronquist F. & Huelsenbeck J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19:**1572–1574.

Rossiter M.A., Howard D.J. Bush G.L. 1983. Symbiotic bacteria of *Rhagoletis pomonella*. In Fruit Flies of Economic Importance (R. Cavalloro, Ed.), pp. 77-82. Balkema, Rotterdam, the Netherlands

Roy B.A. 2001. Patterns of association between crucifers and their flower-mimic pathogens: host jumps are more common than coevolution or cospeciation. *Evolution* **55:**41–53.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145:**1219–1228.

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.

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.

Shimodaira H. & Hasegawa M. 1999. Multiple comparisons of loglikelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16:**1114–1116.

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 7–8th April 2008. Abstract book, Vol. 15 Ed. by Miranda-Chueca MA Universitat de les Illes Balears, 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.

Simon R., Priefer U., Pulher A. 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria. *Nat. Biotechnol.* **1:**784-791.

Skillman L.C, Sutherland I.W., Jones M.V., Goulsbra A. 1998. Green fluorescent protein as a novel species-specific marker in enteric dual-species biofilms. *Microbiol.-UK* **144:**2095-2101.

Spanos L., Koutroumbas G., Kotsyfakis M., Louis C. 2000. The mitochondrial genome of the Mediterranean fruit fly, *Ceratitis capitata*. *Insect Mol. Biol.* **9:**139–144.

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

Steinhaus E.A. 1960. The importance of environmental factors in the insects-microbe ecosystem. *Bacteriol. Rev.* **24:**365-373.

Suarez A., Güttler A., Strätz M., Staendner L.H., Timmis K.N., Guzmàn C.A. 1997. Green fluorescent protein-based reporter systems for genetic analysis of bacteria including monocopy applications. *Gene* **196:**69-74.

Swofford D.L., Olsen G.L., Waddell P.J., Hillis D.M. 1996. Phylogenetic inference, pp. 407-514. In D.M. Hillis, C. Moritz, and B.K. Mable (eds.), Molecular Systematics. Sinauer, Sunderland, MA.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:**585-595.

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.

Tanaka K., Shimizu T., Zakria M., Njoloma J., Saeki Y., Sakai M., Yamakawa T., Minamisawa K., Akao S. 2006. Incorporation of a DNA sequence encoding green fluorescent protein (GFP) into endophytic diazotroph from sugarcane and sweet potato and the colonizing ability of these bacteria in *Brassica oleracea*. *Microbes Environ.* **21:**122-128.

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.

Templeton A.R., Routman E., Phillips C. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. *Genetics* **140:**767-782.

Thaochan N., Drew R.A.I., Hughes J.M., Vijaysegaran S., Chinajariyawong A. 2010. Alimentary tract bacteria isolated and identified with API-20E and molecular cloning techniques from Australian tropical fruit flies, *Bactrocera cacuminata* and *B. tryoni*. *J. Insect Sci.***10:**131.

Tsiropoulos G.J. 1976. Bacteria associated with the Walnut Husk fly, *Rhagoletis completa*. *Environ. Entomol*. **5:**83-86.

Tzanakakis M.E. 2006. Insects and Mites Feeding on Olive: Distribution, Importance, Habits, Seasonal Development and Dormancy. Leiden: Brill Acad. Publ. 182 pp.

Vasanthakumar A., Delalibera I., Jr., Handelsman J., Klepzig K.D., Schloss P.D., Raffa K.F. 2006. Characterization of gut-associated bacteria in larvae and adults of the southern pine beetle, *Dendroctonus frontalis* Zimmermann. *Environ. Entomol.* **35:**1710-1717.

Ward D.M., Weller R., Bateson M.M. 1990. 16S rRNA reveal numerous uncultured microorganisms in a natural community. *Nature* **345:**63–65.

Weiblen G.D. & Bush G.L. 2002. Speciation in fig pollinators and parasites. *Mol. Ecol.* **11:**1573–1578.

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.

Wernegreen J.J. 2004. Endosymbiosis: Lessons in Conflict Resolution. *Plos Biol.* **2:**e68.

White I.M. 1988. Handbooks for the Identification of British Insects. Vol. X, Part 5 (a). Tephritid Flies, Diptera: Tephritidae. Royal Entomological Society, London. 1-134.

White I.M. 2006. Taxonomy of the Dacina (Diptera: Tephritidae) of Africa and the Middle East. *Afr. Entomol. Mem.* **2:**1–152.

Xia X. & Xie Z. 2001. DAMBE: Software package for data analysis in molecular biology and evolution. *J. Hered.* **92:**371–373.

Xiong B. & Kocher T.D. 1991. Comparison of mitochondrial DNA sequences of seven morphospecies of black flies (Diptera: Simuliidae). *Genome* **34:**306-311.

Yoder M. & Gillespie J. 2004-present. jRNA. Exploring insect phylogeny using RNA secondary structure. Web pages at http://hymenoptera.tamu.edu/rna.

Zwickl D.J. 2006. GARLI, Genetic Algorithm for Rapid Likelihood Inference, Version 0.951. Available from: http://www.zo.utexas.edu/faculty/antisense/garli/Garli.html.

Zwölfer H. 1983. Life systems and strategies of resource exploitation in tephritids. In: Cavalloro, R. (ed.), Fruit flies of economic importance. Proc. CEC/IOBC Int. Symp. Athens, 16-19 November 1982. A.A. Balkema, Rotterdam.

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.

Acknowledgments

Firstly I would like to thank my supervisor Prof. Vincenzo Girolami, and Dott. Luca Mazzon, who assisted and supported me for the three years of this experience.

I thank Prof. Andrea Squartini, microbiologist at the University of Padova, Prof.ssa Laura Marri microbiologist at the University of Siena and Mauro Simonato for their kindly collaborations and precious advices on data analysis.

A special thanks go to my research colleague, mate of office and friend Isabel Martínez-Sañudo.

I thank all my friends and colleagues of the department Alberto, Andrea, Caterina, Diego, Edoardo, Elisabetta, Ewelina, Fernanda, Haya, Linda, Lorenzo, Matteo, Mauro, Paola S., Paola T., Shakeel, Teshale. Thanks also to the entomology staff at the department (Professors Andrea Battisti, Carlo Duso, Luigi Masutti, Giuseppina Pellizzari; Dott. Massimo Faccoli, Gabriella Frigimelica, Nicola Mori and technicians Patrizia dall'Ara, Gabriella Fenici, Paolo Paolucci). Many thanks also to those professors and colleagues that provide me *Bactrocera oleae* specimens.

And finally I want to thank Daniel who always gave me happiness and encouragement and my family for their precious support.