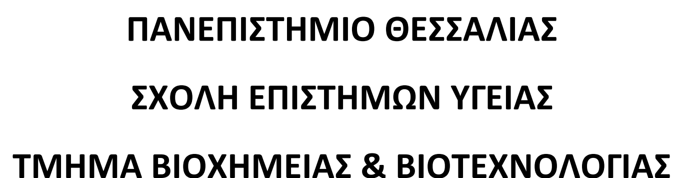

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ

ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ & ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

ΚΩΝΣΤΑΝΤΙΝΑ ΤΣΟΥΜΑΝΗ

ΜΟΡΙΑΚΗ ΑΝΑΛΥΣΗ ΤΗΣ ΔΟΜΗΣ ΚΑΙ ΤΗΣ ΟΡΓΑΝΩΣΗΣ ΤΟΥ ΓΟΝΙΔΙΩΜΑΤΟΣ ΤΟΥ ΣΗΜΑΝΤΙΚΟΤΕΡΟΥ ΠΑΡΑΣΙΤΟΥ ΤΗΣ ΕΛΙΑΣ, ΤΟΥ ΕΝΤΟΜΟΥ *BACTROCERA OLEAE*

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΛΑΡΙΣΑ 2012

UNIVERSITY OF THESSALY

SCHOOL OF HEALTH SCIENCES

DEPARTMENT OF BIOCHEMISTRY & BIOTECHNOLOGY

MOLECULAR ANALYSIS OF GENOME ORGANIZATION AND STRUCTURE OF THE OLIVE FRUIT FLY, *BACTROCERA OLEAE*

This dissertation is submitted for the degree of Doctor of Philosophy

by

KONSTANTINA TSOUMANI

LARISSA 2012

ABSTACT

The olive fruit fly, *Bactrocera oleae*, causes extensive damages in olive production. Apart from its great economic importance, little information is available at the genetic and molecular level. Genomic analysis of this insect is important not only with regard its basic research, but also due to further implications in biological control methods. Effective application of such methods, like the Sterile Insect Technique (S.I.T.), relies on the availability of fundamental genetic and molecular data. Consequently, the analysis of *B. oleae*'s genome structure may contribute towards the development of basic knowledge and tools for this insect.

In the context of the present analysis, the *B. oleae*'s genome size was determined through a quantitative Real-time PCR approach. Genome size was estimated at about 3.22 x 10^8bp, which places it at the lower end of the dipteran genome size range.

The isolation and characterization of 195 selected cDNA clones was also achieved. The sequence determination of these ESTs allowed further comparisons with known protein sequences deposited in the database and consequently the corresponding genes were characterized. In addition, based on *D. melanogaster* matches, the ESTs were functionally assigned and classified with Gene Ontology terms, allowing a primary estimation of the insect's transcriptome organization according to the ESTs' distribution to each functional category.

Moreover, the exact cytogenetic localization of 35 EST clones was determined by *in situ* hybridization to the insect's polytene chromosomes, enriching the available physical map with more than double entry points.

Additionally, it the synteny between *B. oleae* and *D. melanogaster* was also explored, firstly by demonstrating the colinearity of the chromosomal segments between the two species and secondly by the examination of the respective chromosomal positions among the mapped ESTs and their Drosophila orthologs. As a consequence, a synteny map of the two species was produced, enabling also the identification of some syntenic blocks.

Taking advantage of the availability of these new sequences, the overall codon frequencies of the active transcriptome were calculated, in order to examine the presence of a conserved profile. The comparative sequence analysis of the *B. oleae* ESTs with the homologous *D. melanogaster* genes led to the development of 17 nuclear EPIC-PCR markers (Exon Priming Intron Crossing-PCR). These markers were used for the amplification of the corresponding intron sequences in 11 Tephritids, which were subsequently used for phylogenetic comparisons.

Additionally, with the use of a segment of a putative LTR retrotransposon (named *Achilles*) as a probe, a *B. oleae* genomic library was screened and several phage clones were analyzed in order to isolate an intact copy of the retrotransposon. Molecular and *in silico* analysis of the isolated clones led to the clarification of the *B.* oleae's LTR retrotransposon *Achilles* organization, according to the respective structure of the homologous element *MAX* of *D. melanogaster*, which belongs to the BEL-Pao family. More specifically, the *Achilles* sequence was determined at about 7,487 bp, consisting of the 5' LTR end, the 5' non coding sequence and the open reading frame (ORF) which encodes the polyprotein gag-pol. In a quantitative Real-time PCR approach the element's copies were calculated at about 42 copies per genome. Moreover the activity of the element was examined, demonstrating the presence of at least one active copy of *Achilles* in the genome.

During the analysis of *Achilles* containing phage clones, a conserved centromeric satellite sequence of about ~ 300 bp was also discovered. This tandem repeat, named BoR300, was species specific, as determined by Southern analysis on several related species. This finding can render it a species diagnostic marker, particularly of the morphologically indistinguishable larval and pupal stages of several closely related species of the

Tephritidae family. Finally, BOR300 was estimated to occupy ~ 0.3 % of the insect's genome, whereas the examination of its distribution on polytene and mitotic *B. oleae* choromosomes enabled the first correlation between these two types of chromosomes, concerning 2 out of the six autosomes.