

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ  
Πρόγραμμα Μεταπτυχιακών Σπουδών  
του Τμήματος Βιοχημείας και Βιοτεχνολογίας  
«ΕΦΑΡΜΟΓΕΣ ΜΟΡΙΑΚΗΣ ΒΙΟΛΟΓΙΑΣ - ΜΟΡΙΑΚΗ ΓΕΝΕΤΙΚΗ, ΔΙΑΓΝΩΣΤΙΚΟΙ ΔΕΙΚΤΕΣ»

Τασσόγλου Σπύρος

**Ανάλυση πεπτιδόματος των αναπαραγωγικών αδένων του θηλυκού δάκου της ελιάς**

Λάρισα, 2016

**Ανάλυση πεπτιδόματος των αναπαραγωγικών αδένων του θηλυκού δάκου της ελιάς**  
**Peptidomic analysis of female accessory glands of *Bactrocera oleae***

Τριμελής εξεταστική επιτροπή:

1. Μαθιόπουλος Κωνσταντίνος, Καθηγητής Μοριακής Βιολογίας\*
2. Αμούτζιας Γρηγόριος, Επίκουρος Καθηγητής Βιοπληροφορικής στη Γενωμική\*
3. Κοντού Μαρία, Επίκουρος Καθηγήτρια Πρωτεϊνικής Χημείας\*

---

\* Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας

## Ευχαριστίες

Ολοκληρώνοντας τη μεταπτυχιακή μου εργασία, θέλω να ευχαριστήσω κάποιους ανθρώπους που αυτό το διάστημα με οδήγησαν, με βοήθησαν, μου έδειξαν μία διαφορετική οπτική, με έβαλαν στη θέση μου, με ανέχτηκαν, με όρισαν, ο καθένας με κάποιο τρόπο.

Δεν περιγράφεται η χαρά και η τύχη μου που ο Αλέξανδρος βρέθηκε στο δρόμο μου, φιλικά, επιστημονικά, συντροφικά. Πέρα από το ρόλο που έπαιξε στο να βρεθώ σε αυτό το εργαστήριο και το επιστημονικό πεδίο, με κερνάει διαρκώς μεράκι, σχολαστικότητα, πείσμα και πολλή πολλή δύναμη σε νοητικές συσκευασίες στις οποίες είμαι πολύ δεκτικός. Του αφιερώνω την εργασία αυτή, γιατί ειδικά το τελευταίο εξάμηνο θά 'θελα να τον έχω δει πολύ παραπάνω. Ευχαριστώ τον κ. Ματθιόπουλο που μου έδειξε τεράστια εμπιστοσύνη συμπεριλαμβάνοντάς με στην ερευνητική του ομάδα και αναθέτοντάς μου αυτό το θέμα. Είναι ένας αληθινός δάσκαλος που κατορθώνει και ισορροπεί τόσο καλά στην διδασκαλία του τη μεθοδική, αναλυτική σκέψη και τη φαντασία, δύο χαρακτηριστικά που και τα δύο μαζί δεν τα βρίσκεις συχνά σε αφθονία. Χάρη στον κ. Ματθιόπουλο βρέθηκα στο εργαστήριο βιοπληροφορικής της κ. Χατζηγεωργίου, την οποία επίσης ευχαριστώ πολύ που με δέχτηκε. Από το εργαστήριο μοριακής βιολογίας είμαι ευγνώμων που είχα τη Μαριλένα να καθοδηγεί τη σκέψη μου και τα πειράματά μου, να λύνει τις απορίες μου και να μου συμπεριφέρεται πάντα με απεριόριστη υπομονή. Ευχαριστώ ακόμα την Κωνσταντίνα, ειδικά επειδή με έσπρωξε την πιο κατάλληλη στιγμή να.. μιλήσω, τη Γιώτα που ήταν πάντα παρούσα και διαθέσιμη για – επιστημονικές και μη – ανησυχίες και ιδέες και την Έφη για την πολύ καλή της διάθεση και τα βοηθητικά της σχόλια. Ακόμα ευχαριστώ τα παιδιά στο εργαστήριο βιοπληροφορικής, τον Κωνσταντίνο, τη Μαρία, τη Δήμητρα, το Γιάννη και το Γιώργο που με αγκάλιασαν από την πρώτη στιγμή και είχαν πάντα διάθεση να κάνουν ουσιαστικές προτάσεις και παρατηρήσεις, τόσο προγραμματιστικά, όσο και βιολογικά. Ευχαριστώ επίσης την Ingrid Curtil για την πολύ επικοινωνιακή μας επικοινωνία που έχει καθορίσει σε μεγάλο βαθμό το περιεχόμενο και την ερευνητική προσέγγιση αυτής της διπλωματικής. Τέλος, ευχαριστώ θερμά τον κ. Αμούτζια και την κ. Κοντού που δέχθηκαν να συμμετέχουν στην εξεταστική μου επιτροπή, ενώ μάλιστα τους το ζήτησα με ιδιαίτερη καθυστέρηση.

Οφείλω να κάνω τουλάχιστον άλλο ένα σχόλιο: Αυτή η εργασία, όπως και οι περισσότερες υποθέτω, θα ήταν πολύ πιο ελλιπής αν δεν υπήρχε ο ιστότοπος sci-hub.io, που παρέχει τη δυνατότητα να «ανοίξουν» δημοσιεύσεις που είναι κλειδωμένες από τους εκδοτικούς τους οίκους και απαιτείται αντίτιμο για το κατέβασμά τους. Η Αλεξάνδρα Ελμπακιάν προσφέρει με το Sci-Hub ελεύθερη πρόσβαση στη γνώση στους πάντες και έχει την απεριόριστη εκτίμησή μου για αυτό. Οφείλουμε όλοι να κάνουμε ό,τι περνάει από το χέρι μας ώστε η γνώση να διακινείται ελεύθερα, και να ενισχύεται, μέχρι τη στιγμή που η ελεύθερη διάχυσή της θα είναι πια κάτι το αυτονόητο.

Χρυστάω πολλά σε όλους τους παραπάνω, μα κι άλλα σε άλλους τόσους. Γράφοντας βρέθηκα να προβάλλω κομμάτια των άλλων στον εαυτό μου και να αναρωτιέμαι αν ακούσια το μυαλό μου επέλεξε να ερεθιστεί από τα καλύτερά τους. Σίγουρα βρέθηκα να συλλέγω συμπεριφορές και τρόπους σκέψης και να ελπίζω πως, έστω λίγο, κατάφερα να αβαντάρω τα δικά μου καλύτερα κομμάτια έναντι των αδυναμιών μου, ώστε αυτά να ανταποδίδω στην καθημερινότητα, την επιστημονική αλληλεπίδραση και τη διαπροσωπική επαφή.

Μακάρι.

Περιεχόμενα	
Περίληψη .....	5
1. Εισαγωγή .....	6
1.1 Τεχνική Στείρου Εντόμου (Sterile Insect Technique – SIT).....	7
1.1.1 Η σύλληψη της SIT .....	7
1.1.2 Ανάπτυξη και πρώτη εφαρμογή .....	7
1.1.3 Επιτυχημένα SIT προγράμματα .....	7
1.1.4 Οι αδυναμίες της SIT και η μετεξέλιξή της .....	7
1.1.5 RNA παρεμβολή (RNA interference – RNAi).....	8
1.1.6 Γονίδια αυτοκατευθυνόμενων ενδονουκλεασών – Homing endonuclease genes (HEGs).....	8
1.1.7 Απελευθέρωση Εντόμων που φέρουν ένα Επικρατές Θνησιγόνο γονίδιο – Release of Insects carrying a Dominant Lethal gene (RIDL) .....	8
1.2 Αναπαραγωγή.....	11
1.2.1 Πρόλογος.....	11
1.2.2 Η μοριακή φυσιολογία της σύζευξης.....	11
1.2.3 Το Sex Peptide .....	11
1.2.4 Άλλες SFPs.....	12
1.2.5 Μετασυσζευκτική απόκριση και άλλοι ιστοί.....	12
1.2.6 Πλαστικότητα και διαειδική ποικιλομορφία της αναπαραγωγής των εντόμων .....	12
1.3 Πρωτεομική και πεπτιδομική .....	13
1.3.1 Φασματοσκοπία μάζας (Picó, 2015) .....	13
1.3.2 Είδη πηγών ιονισμού (Picó, 2015) .....	13
1.3.3 Πεπτιδομική .....	15
1.4 Στόχοι: .....	16
2. Υλικά και μέθοδοι.....	17
2.1 Εκτροφή εντόμων και σύζευξη .....	18
2.2 Απομόνωση πεπτιδίων – λυοφιλίωση .....	18
2.3 HPLC αντίστροφης φάσης – MALDI-TOF/TOF φασματοσκοπία μάζας .....	18
2.4 Αναζήτηση σε βάσεις δεδομένων με το Mascot .....	18
2.5 Χειροκίνητη επαλήθευση .....	18
2.6 Έλεγχος ποιότητας .....	19
2.7 Ανάλυση Gene Ontology – REVIGO.....	20
2.8 Σύγκριση με την τρανσκριπτομική ανάλυση.....	20
2.9 Ανάλυση TBLASTN και BLASTN με μετάγραφα που κωδικοποιούν για ACPs (Accessory Gland Proteins) άλλων οργανισμών .....	20
3. Αποτελέσματα και συζήτηση .....	21
3.1 Έλεγχος ποιότητας και Χειροκίνητη επαλήθευση (manual validation) .....	22
3.2 Ανάλυση όρων Gene Ontology .....	23
3.3 Σύγκριση πεπτιδικής ανάλυσης με τρανσκριπτομικά δεδομένα .....	23
3.4 Αναζήτηση για πιθανά ACP μόρια στους αναπαραγωγικούς αδένες του θηλυκού .....	28
3.5 Σύνοψη .....	29
4. Παράρτημα: Πίνακες .....	30
Βιβλιογραφία .....	43

## Περίληψη

Τα έντομα που παρασιτούν σε βάρους αγροκτηνοτροφικών οργανισμών υψηλής αξίας για τον άνθρωπο αντιμετωπίζονται παραδοσιακά με εντομοκτόνα. Η χρήση εντομοκτόνων όμως χαρακτηρίζεται από χαμηλή εξειδίκευση, ενώ οι πληθυσμοί στους οποίους εφαρμόζεται σταδιακά παρουσιάζουν ανοχή και προκύπτει διαρκώς η ανάγκη για ανάπτυξη νέων και ισχυρότερων δραστικών ουσιών, δημιουργώντας ένα φαύλο κύκλο. Τις τελευταίες δεκαετίες, χάρη στη ραγδαία ανάπτυξη μοριακών τεχνικών και διαγνωστικών μεθόδων και την αυξανόμενη γνώση πάνω στη γενετική και τη μοριακή φυσιολογία των εντόμων, έχουν προκύψει αξιόλογες μοριακές εναλλακτικές προτάσεις καταπολέμησης. Η RIDL, που αποτελεί μία από αυτές, συνίσταται στην απελευθέρωση στον αγρό εντόμων που φέρουν ένα επικρατές θνησιγόνο γονίδιο ώστε με τη σύζευξή τους με τα έντομα αγρίου τύπου να μην προκύψει απόγονος και να ελεγχθούν έτσι τα επίπεδα πληθυσμού. Το αναπαραγωγικό σύστημα αποτελεί ιδανικό πεδίο αναζήτησης γονιδίων που είναι απαραίτητα για την επιτυχία της αναπαραγωγής, αλλά και υποκινητών, μεταγραφικών παραγόντων και λοιπών ρυθμιστικών στοιχείων που θα μπορούσαν να χρησιμοποιηθούν για την εφαρμογή της RIDL. Σκοπός της παρούσας μεταπτυχιακής εργασίας ήταν η ανάλυση του πεπτιδόματος των αναπαραγωγικών ιστών του θηλυκού δάκου *Bactrocera oleae*, που είναι είδος μύγας της οικογένειας των Tephritidae και αποτελεί το σημαντικότερο εχθρό της ελαιοπαραγωγής. Η ανάλυση έγινε συγκριτικά μεταξύ 25 ασύζευκτων ατόμων και 25 ατόμων που έχουν συζευχθεί, ώστε να ταυτοποιηθούν βιοδραστικά πεπτίδια με ρόλο στην αναπαραγωγή. Πραγματοποιήθηκε απομόνωση των μικρών πεπτιδίων, κλασμάτωσή τους με HPLC αντίστροφης φάσης, ιονισμός και διαδοχική φασματοσκοπία μάζας με MALDI-TOF/TOF, βιοπληροφορική ανάλυση των φασμάτων και σύγκριση των ευρημάτων με προηγούμενες αναλύσεις του μεταγραφόματος.

## Abstract

Insects that have a parasitic relationship with organisms of agricultural importance have been traditionally treated with insecticides that are usually non-specific. More importantly though, the insect population the insecticide is directed to gradually develops resistance, constantly arising the need to develop new, more toxic substitutes, locking the issue in an infinite loop. Fortunately the last few decades, thanks to the booming development of molecular techniques and diagnostic methods and also to newly apprehended knowledge on insect genetics and molecular physiology, molecular counterproposals to insecticide use have been conceived and are under development for a number of parasite or vector species. RIDL is based on the Release of Insects carrying a Dominant Lethal gene at the crop and on the expectation that their mating with wild type insects will result in zero offspring, thus reducing total population numbers. The reproductive system is an ideal field in the quest to find genes necessary to reproductive success, as well as related promoters, transcription factors and other regulatory molecules that could be of use in a RIDL application. All that being said, the purpose of this thesis was to carry out analysis of the peptidome of the female reproductive tissue of *Bactrocera oleae*, which is a Tephritid fly species that bears destructive effects to the olive crops by egg-laying exclusively on the olive sap. A comparison of the peptide content between pools of 25 mated/unmated individuals was made, in an attempt to identify bioactive peptides with a role in reproduction. The small peptides were isolated, fractionized via reverse phase HPLC, then ionized and analyzed in fragments with tandem mass spectrometry (MALDI-TOF/TOF). Bioinformatics analysis of the resulting spectra and comparison of the peptidomics findings with prior transcriptomics analyses followed.

# 1. Εισαγωγή

## 1.1 Τεχνική Στείρου Εντόμου (Sterile Insect Technique – SIT)

### 1.1.1 Η σύλληψη της SIT

Η Τεχνική Στείρου Εντόμου (SIT) βασίζεται στην υπόθεση ότι η έκθεση ενός φυσικού πληθυσμού εντόμων σε έναν αριθμό στείρων μα κατά τ' άλλα σεξουαλικά ενεργών αρσενικών θα προκαλέσει εντονότερες συνέπειες στη δυναμική επιβίωσης του πληθυσμού, από ό,τι η μείωση ίσου αριθμού ατόμων από τον πληθυσμό, με θανάτωση – χρήση εντομοκτόνων – ή αφαίρεσή τους – χρήση παγίδων (Knippling, 1959).

### 1.1.2 Ανάπτυξη και πρώτη εφαρμογή

Αναπτύχθηκε τη δεκαετία του 1950 για την καταπολέμηση, της *Callitroga hominivorax*, ενός είδους μύγας που αποθέτει τα αυγά του σε ανοικτές πληγές και επιθηλιακές κυτταρικές σπινθήρες θηλαστικών, του ανθρώπου συμπεριλαμβανομένου. Οι Edward Knippling και Raymond Bushland, βασιζόμενοι στην παρατήρηση του Muller, το 1927, ότι έκθεση της *Drosophila melanogaster* σε ακτίνες X προκαλεί χρωμοσωμικές μεταλλάξεις (Muller, 1927), το 1951 πειραματίστηκαν ακτινοβολώντας βομβύκια *Callitroga hominivorax* διαφόρων σταδίων με ακτίνες X και το 1953 με ακτίνες Γ, που είναι πιο οικονομικές. Παρατήρησαν ότι ένα φυσιολογικό θηλυκό που είχε συζευχθεί με ένα στείρο αρσενικό, δεν έκανε επανασύζευξη με φυσιολογικά αρσενικά και απέθετε μόνο μη-γόνιμα αυγά. Ακόμα, ότι η αναλογία μη γόνιμων και εκκολαπτόμενων αυγών σε κλουβιά με στερωμένες και αγρίου τύπου μύγες ήταν σχετική με τους σχετικούς αριθμούς των στείρων και φυσιολογικών αρσενικών, αντίστοιχα. Υπέθεσαν ότι το φαινόμενο που παρατηρούσαν ήταν η έκφραση της Επικρατούς Θνησιγόνου Επίδρασης (Dominant Lethal Effect), μίας αδρής έννοιας που περιγράφει τη συσσώρευση μεταλλάξεων στο σπερματοζώαριο που δεν είναι ικανές να εμποδίσουν τη σύντηξη του με το ωάριο, μα προκαλούν πρόωρες αναπτυξιακές βλάβες στο ζυγωτό και οδηγούν στον εκφυλισμό του (Bushland & Hopkins, 1953).

Το 1954 προχώρησαν σε μαζική εκτροφή και απελευθέρωση στείρων αρσενικών *Callitroga hominivorax* στο νησί Κουρασάο, ανοικτά της ακτής της Βενεζουέλας, και πέτυχαν την πλήρη εξάλειψή του, οπότε η SIT σταδιακά άρχισε να κερδίζει έδαφος στο πεδίο του πληθυσμιακού ελέγχου εντόμων που αποτελούν κίνδυνο για τη δημόσια υγεία, για αγροκτηνοτροφικά είδη οικονομικής σημασίας. Ιδιαίτερο ρόλο στη θετική απόκριση της μεθόδου αποτέλεσε και το γεγονός ότι αποτελεί μία εναλλακτική πρόταση στη χρήση εντομοκτόνων, φιλική προς το περιβάλλον, που δεν αφήνει κατάλοιπα στο πεδίο και χαρακτηρίζεται από ειδοειδικότητα.

Ήδη από το 1959, έγιναν δύο κρίσιμες παρατηρήσεις αναφορικά με τη φυσιολογία της SIT: (1) ο τρόπος στείρωσης δεν είναι απαραίτητο να είναι η ακτινοβολήση – εκείνη την περίοδο προτάθηκε η χρήση χημικών παραγόντων που προκαλούν στείρωση – και (2) για να υπάρχει επιτυχία της μεθόδου πρέπει η αναπαραγωγική δεινότητα του στερωμένου αρσενικού να παραμένει κατά το δυνατόν ανεπηρέαστη (Knippling, 1959). Για να μπορεί δηλαδή να ανταγωνιστεί αναπαραγωγικά τα αρσενικά του φυσικού πληθυσμού, πρέπει να είναι ικανό να παρουσιάσει συζευκτική συμπεριφορά (αργότερα αναφέρεται ως courtship behaviour) που να πετυχαίνει στο να προκαλεί το θηλυκό σε συζευκτικό παιχνίδι και να είναι ικανό να ολοκληρώνει τη σύζευξη όπως ένα αρσενικό αγρίου τύπου.

### 1.1.3 Επιτυχημένα SIT προγράμματα

Ακολούθησαν πολλά επιτυχημένα προγράμματα μαζικών και διαδοχικών απελευθερώσεων στείρων

αρσενικών εντόμων, στη Φλόριντα και άλλες νοτιοανατολικές πολιτείες των Η.Π.Α. στις αρχές της δεκαετίας του 1960 ξανά για την καταπολέμηση του *Callitroga hominivorax*. Τα επόμενα 25 χρόνια, οι κτηνοτρόφοι στις Η.Π.Α. και το Μεξικό πίεσαν τις κυβερνήσεις ώστε να οργανωθεί από κοινού ένα διακρατικό πρόγραμμα εφαρμογής SIT για τη συγκεκριμένη μύγα. Το πρόγραμμα αυτό, σε συνεργασία με Ευρωπαϊκά και Αφρικανικά προγράμματα κατά τη δεκαετία του 1990 πήρε SIT μέτρα στη Λιβύη, όπου παρατηρήθηκε μεγάλη αύξηση του πληθυσμού *Callitroga hominivorax* και υπήρχε φόβος εξάπλωσής του στην υποσαχάρια Αφρική και τη μεσογειακή λεκάνη (A programme for the eradication of the New World Screwworm from North Africa, 1990).

Παράλληλα η εφαρμογή SIT επεκτάθηκε και σε άλλα είδη εντόμων οικονομικής σημασίας ή εντόμων-φορέων παθογόνων. Προγράμματα μαζικής εκτροφής και διαδοχικών απελευθερώσεων στείρων αρσενικών στο πεδίο πραγματοποιήθηκαν ενάντια στη μεξικάνικη φρουτόμυγα *Anastrepha ludens* στο Texas και τη Νότια Καλιφόρνια (1964-σήμερα) (United States, Mexico, and Guatemala Fruit Fly Emergence and Release Facilities Review, 2008), την αυστραλιανή φρουτόμυγα *Bactrocera tryoni* στην Αυστραλία (1962-1990) (Fisher, 1996), τη μεσογειακή μύγα *Ceratitidis capitata* στο Μεξικό και τη Γουατεμάλα (1978-1982) (Villaseñor, 2000), την Καλιφόρνια και τη Φλόριντα (1994-σήμερα, VIENNA STRAIN) (Caceres, 2004), την Ιορδανία, το Ισραήλ και την Παλαιστίνη (1996-σήμερα, VIENNA STRAIN) (Cayol, 2004), τη μύγα του πεπονιού *Bactrocera cucurbitae* στην Ιαπωνία (1972-1993) (Yosiaki, Kakinohana, Yamagishi, & Kohama, 2003), διάφορα είδη μύγας Τσετσέ (είδη του γένους *Glossina*) στη Μπουρκίνα Φάσο (Politzar, 1984), τη Νιγηρία (Takken, 1986) και την Τανζανία (Msangi, 2000) (1970-1990) και στην Ουγκάντα (1990) (Oloo, 2000), το *Anopheles quadrimaculatus* (1959-1960) (Dame, Woodward, Ford, & Weidhaas, 1964) και το *Culex quinquefasciatus* (1969) (Patterson, 1970) στη Φλόριντα, το *Culex quinquefasciatus* στο Δελχί (1973) (Yasuno, 1978), το *Anopheles albimanus* στο Ελ Σαλβαδόρ (1972, 1977-1979) (Lofgren, et al., 1974) (Dame D. A., 1981), το *Aedes albopictus* στη γαλλική αποικία Ρεϋνιόν (2012) (Oliva, Jacquet, Gilles, Lempriere, & P.O. Maquart, 2012) και σε άλλα είδη.

Όμως, αξίζει να αναφερθεί ότι οι πιο επιτυχημένες περιπτώσεις εφαρμογής της SIT, στις οποίες αναφέρεται πλήρης εξάλειψη του πληθυσμού-στόχου, έλαβαν χώρα σε νησιά, σε περιοχές δηλαδή που περικλείονται από φυσικά εμπόδια που αποτρέπουν τη διαφυγή εντόμων από την εφαρμογή και τη μετανάστευση νέων εντόμων προς εμπλουτισμό του τοπικού πληθυσμού.

### 1.1.4 Οι αδυναμίες της SIT και η μετεξέλιξη της

Παρά την ευρεία εφαρμογή της μεθόδου και το μεγάλο αριθμό επιτυχημένων προγραμμάτων εξάλειψης ή ελέγχου πληθυσμών, η SIT σήμερα αποτελεί μία παρωχημένη τεχνική με αρκετά μειονεκτήματα:

- Η εκτροφή τεράστιου αριθμού εντόμων, η ακτινοβολήση, οι απελευθερώσεις και ο έλεγχος σε κάθε στάδιο της πορείας του προγράμματος από εξειδικευμένο προσωπικό αυξάνουν σημαντικά το κόστος.
- Στα περισσότερα είδη δεν υπάρχει η δυνατότητα να γίνει φυλοδιαχωρισμός από αρκετά νωρίς – ιδανικά από το στάδιο του εμβρύου (αυγού ή πρώτα προνομφικά στάδια) – ώστε να εκτραφούν μόνο αρσενικά και κάποια θηλυκά που θα χρησιμεύσουν μόνο για τη διατήρηση του εργαστηριακού πληθυσμού. Η



απελευθέρωση και των δύο φύλων που γινόταν αρχικά είχε πολλαπλά προβλήματα. Τα θηλυκά, παρά το γεγονός ότι ήταν στείρα, ωοαπέθεταν κανονικά, δημιουργώντας δευτερογενείς πηγές μόλυνσης στα φρούτα. Αφετέρου, ανταγωνίζονταν τις 'άγριες' θηλυκές για τη σύζευξη με τα στείρα εργαστηριακά αρσενικά που απελευθερώνονταν (Whitten, 1969), (Fletcher, 1989), (Vreysen, 2007).

- Η χρήση ακτινοβολίας αφενός επηρεάζει τη συζευκτική ικανότητα των αρσενικών, με αποτέλεσμα να αδυνατούν σε αρκετές περιπτώσεις να ανταγωνιστούν αναπαραγωγικά τα αρσενικά του φυσικού πληθυσμού, αφετέρου δε δεν εξασφαλίζει τη στειρότητα 100%. Η πρόκληση τόσο πολλών τυχαίων μεταλλάξεων το πιο πιθανό είναι ότι θα προκαλέσει στειρότητα, αλλά υπάρχει πάντα η πιθανότητα να προσδώσει, αρχικά στον εργαστηριακό πληθυσμό και επακόλουθα στο φυσικό, ανεπιθύμητα χαρακτηριστικά.

Από τη σύλληψη της SIT μέχρι σήμερα, οι επαναστάσεις στη γενετική μηχανική διαδέχονταν η μία την άλλη και απέφεραν απαραίτητα εργαλεία για την ακριβέστερη κατανόηση και τροποποίηση του γονιδιώματος. Οι ανακαλύψεις των πλασμιδίων (Lederberg, 1952), των DNA λιγασών (Weiss & Richardson, 1967) και των ενζύμων περιορισμού (Kelly & Smith, 1970) επέτρεψαν τον ανασυνδυασμό του γενετικού υλικού και τη μεταφορά του μεταξύ οργανισμών, ενώ η αλληλούχηση κατά Sanger (Sanger, 1977) και η αλυσιδωτή αντίδραση πολυμεράσης (PCR) (Bartlett & Stirling, 2003) αύξησαν δραματικά τον όγκο της διαθέσιμης βιολογικής πληροφορίας και τις δυνατότητες ταυτοποίησης και χαρακτηρισμού περιοχών του γονιδιώματος.

Τα επιτεύγματα αυτά οδήγησαν στον εμπλουτισμό του βασικού πρωτοκόλλου της SIT με γενετικές μεθόδους (Wilke, et al., 2009). Αντί της στέρωσης με ακτινοβολία ή χρήση χημικών παραγόντων, δοκιμάστηκε η χρήση μεταθετών στοιχείων, η πρόκληση μεταθέσεων και άλλων χρωμοσωμικών ανακατατάξεων, η επαγωγή της ασυμβατότητας σπερματοζωαρίου και ωαρίου (κυτταροπλασματική ασυμβατότητα) με διαμόλυνση με ενδοκυτταρικά παράσιτα (π.χ. *Wolbachia*). Τελικά προέκυψαν κάποιες ολοκληρωμένες προσεγγίσεις καταπολέμησης που συνοψίζονται παρακάτω, ενώ αποτελεί ενεργό και εξελισσόμενο πεδίο έρευνας η βελτίωσή τους και την προσαρμογή τους στις ιδιαιτερότητες κάθε οργανισμού-στόχου.

#### 1.1.5 RNA παρεμβολή (RNA interference – RNAi)

Η RNA παρεμβολή (Εικόνα 1.1.1) έχει χρησιμοποιηθεί σε μία παραλλαγή της SIT, κατά την οποία παρέχοντας στην τροφή των εντόμων τεχνητά δίκλινα τμήματα RNA (dsRNA), επιτυγχάνεται σίγηση ενός επιθυμητού γονιδίου. Δοκιμάστηκε σε κουνούπια του είδους *Aedes aegypti*, όπου προστέθηκαν στην τροφή των προνυμφών dsRNAs που στοχεύουν γονίδια που μεταγράφονται ιστοειδικά στους όρχεις, καθώς και ένα γονίδιο θηλυκού φυλοκαθορισμού (Whyard, et al., 2015). Το αποτέλεσμα ήταν να προκύπτουν αρσενικά κουνούπια με μειωμένη γονιμότητα και να μεταβληθεί η αναλογία αρσενικών-θηλυκών προς τα αρσενικά. Σε άλλο πείραμα με *Aedes aegypti*, στόχος των dsRNA ήταν μία επαναλαμβανόμενη αλληλουχία του RNA γονιδιώματος του ιού του δάγκειου πυρετού τύπου 2 (DEV2), ώστε να παρεμποδιστεί ο πολλαπλασιασμός του ιού στους ιστούς του εντόμου, καθιστώντας το ανθεκτικό στο συγκεκριμένο ιό, άρα και μη-φορέα (Franz, et al., 2006).

Η RNAi προσέγγιση δεν έχει τις αρνητικές συνέπειες της ακτινοβολήσης, ενώ ελαττώνεται η ανάγκη για φυλοδιαχωρισμό κατά την εκτροφή. Παρόλα αυτά, η σύνθεση dsRNA είναι πολύ ακριβή για να πραγματοποιηθεί στην απαιτούμενη κλίμακα.

#### 1.1.6 Γονίδια αυτοκατευθυνόμενων ενδονουκλεασών – Homing endonuclease genes (HEGs)

Τα HEGs χαρακτηρίζονται και ως εγωιστικά γονίδια, καθώς εκμεταλλεύονται τους μηχανισμούς επιδιόρθωσης για να αντιγραφούν και έτσι είναι ικανά να εξαπλωθούν σε έναν πληθυσμό. Σε έναν ετεροζυγώτη ως προς αυτά, είναι ικανά να κόψουν και να ενσωματώσουν τον εαυτό τους στο ανέπαφο αντίγραφο, μετατρέποντας το άτομο σε ομοζυγώτη. Με αυτόν το μηχανισμό ο αριθμός των HEGs στη γονιδιακή δεξαμενή ενός πληθυσμού αυξάνεται κατακόρυφα. Η εισαγωγή των HEGs (Εικόνα 1.1.1) μπορεί να γίνει στοχευμένα και να διακόψει τη φυσιολογική αλληλουχία γονιδίων σχετικών με την ανταγωνιστικότητα των φορέων, ή την αναπαραγωγή και το φυλοκαθορισμό, οδηγώντας σε θηλυκά που είναι ανθεκτικά σε προσβολή από παθογόνα, ή σε πληθυσμιακή μείωση. Τα HEGs έχουν εφαρμοστεί με επιτυχία σε εργαστηριακό επίπεδο στα κουνούπια *Aedes aegypti* (Traver, Anderson, & Adelman, 2009) και *Anopheles gambiae* (Windbichler, et al., 2011). Μία ιδιαιτερότητα των HEGs είναι το γεγονός ότι βασίζονται στην αναπαραγωγή καθ' αυτή για την εγκαθίδρυσή τους στον πληθυσμό. Αυτό σημαίνει κατ'αρχάς ότι δεν υπάρχει ανάγκη μαζικής εκτροφής εντόμων' αρκεί ένας ελάχιστος κρίσιμος αριθμός ώστε να εξασφαλιστεί ότι ο πληθυσμός θα εκτεθεί στα HEGs και στη συνέχεια θα ακολουθήσει εξάπλωση. Από την άλλη πλευρά, σημαίνει επίσης ότι είναι αδύνατον να στοχευτεί ευθέως η γονιμότητα και πρέπει να γίνει σοφότερη επιλογή του γονιδίου-στόχου (πχ. γονίδια φυλοκαθορισμού), και μάλιστα υπό το πρίσμα ότι το νεοεισαχθέν χαρακτηριστικό θα διαμείνει και θα αυξάνεται στη γονιδιακή δεξαμενή του φυσικού πληθυσμού μέχρι την εξάλειψή του.

#### 1.1.7 Απελευθέρωση Εντόμων που φέρουν ένα επικρατές Θνησιγόνο γονίδιο – Release of Insects carrying a Dominant Lethal gene (RIDL)

Η RIDL (Εικόνα 1.1.1) είναι η πιο σύγχρονη μετεξέλιξη της SIT. Αντί για την πρόκληση τυχαίων μεταλλάξεων, στη RIDL τα αρσενικά έντομα που απελευθερώνονται είναι διαγονιδιακά και εκθέτουν τον άγριο πληθυσμό σε θηλεοδραστικά διαγονίδια. Αυτά μπορεί να είναι θνησιγόνα σε ένα πρώιμο αναπτυξιακό στάδιο (πχ. βομβύκιο), ή στην ενήλικη ζωή του εντόμου. Εναλλακτικά, είναι δυνατόν να επιδρούν με έμμεσο τρόπο στην ικανότητα του εντόμου να συζευχθεί, να βρει τροφή, ή να ωοαποθέσει. Ταυτόχρονα, η ευεξία και η αναπαραγωγική συμπεριφορά και ικανότητα των αρσενικών εντόμων που φέρουν το διαγονίδιο επηρεάζεται στο ελάχιστο, αφού η μεταγραφή του καθοδηγείται από θηλεοειδικούς ή σταδιο-ειδικούς υποκινητές. Η κατασκευή του διαγονιδίου είναι αρκετά σύνθετη, καθώς βρίσκεται και υπό τον έλεγχο ενός καταστολέα τετρακυκλίνης, ώστε να επιτρέπεται η εκτροφή του εργαστηριακά, παρέχοντας στην τροφή το αντιβιοτικό τετρακυκλίνη.

Έχουν πραγματοποιηθεί επιτυχημένα πειράματα με εργαστηριακούς πληθυσμούς φυσιολογικών κουνουπιών και διαδοχικές εκθέσεις σε αρσενικά που έφεραν το γονίδιο flightless υπό τον έλεγχο του καταστολέα τετρακυκλίνης και ενός θηλεοειδικού υποκινητή. Ο φαινότυπος flightless χαρακτηρίζεται από ελαττωμένη έκφραση ενός απαραίτητου γονιδίου στους μύς των περυγών. Αποτέλεσμα ήταν τα θηλυκά που γεννιόνταν να είναι ανίκανα να πετάξουν προς αναζήτηση τροφής και ταιριού για σύζευξη. Η RIDL έχει δοκιμαστεί και στο πεδίο, στη Βραζιλία και στα νησιά Cayman (Harris, et al., 2011),

ενάντια σε *Aedes aegypti*, όπου παρατηρήθηκε μεν μειωμένη σύζευξη των διαγονιδιακών αρσενικών με τα θηλυκά αγρίου τύπου, μπόρεσε όμως να αντισταθμιστεί απελευθερώνοντας μεγάλους αριθμούς. Η εφαρμογή της RIDL απελευθερώνοντας συνδυαστικά βομβύκια και ενήλικα *Aedes aegypti* (Winskill, et al., 2014) έχει προταθεί ως μοντέλο ικανό να επιφέρει μακροπρόθεσμη καταστολή σε φυσικούς πληθυσμούς.

Η μέθοδος RIDL έχει το σημαντικό πλεονέκτημα ότι εκμεταλλεύεται τη γενετική μηχανική, ελαχιστοποιώντας τον κίνδυνο γενετικής μόλυνσης. Αντίθετα με τα HEGs που στηρίζονται εν πολλοίς στην ιδιότητά τους να μετατρέπουν έναν ετεροζυγώτη σε ομοζυγώτη επιδρώντας εξαιρετικά δραστικά στη γονοτυπική συχνότητα του υπό παρέμβαση γενετικού τόπου στον πληθυσμό, η RIDL στηρίζεται στις διαδοχικές απελευθερώσεις μεγάλων αριθμών διαγονιδιακών εντόμων στον πληθυσμό και στην αναπαραγωγική τους ανταγωνιστικότητα. Παρόλα αυτά το διαγονίδιο, από τη στιγμή που οδηγεί σε ένα θνησιγόνο φαινότυπο, ποτέ δε θα αποτελέσει στην πραγματικότητα μέρος της πληθυσμιακής γονιδιακής δεξαμενής.

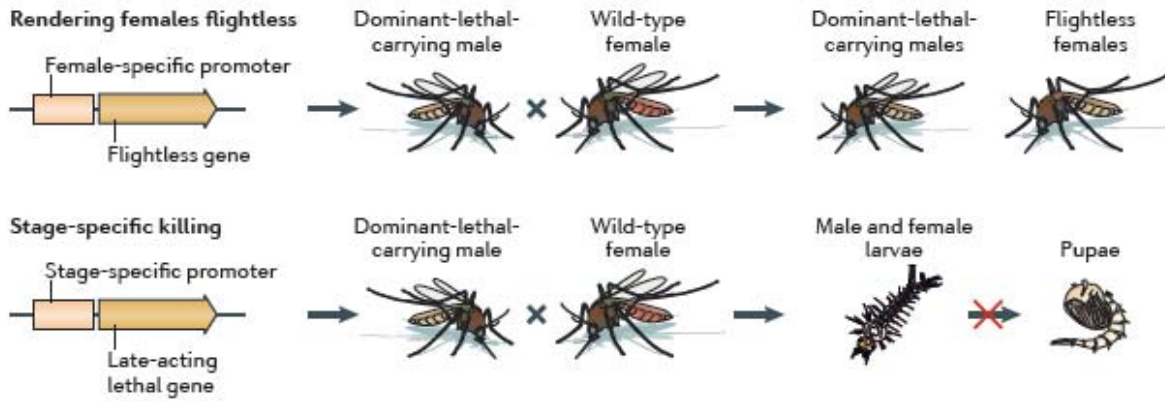
Οι αριθμοί των εντόμων που πρέπει να εκτραφούν αποτελούν, όμως, μειονέκτημα. Στα νησιά Cayman απελευθερώθηκαν 465 διαγονιδιακά αρσενικά ανά εκτάριο, ανά εβδομάδα για 4 εβδομάδες, ενώ μέθοδοι που εισάγουν επιθετικά ένα χαρακτηριστικό σε έναν πληθυσμό, όπως τα HEGs ή η απελευθέρωση εντόμων που φέρουν έναν ενδοσυμβιώτη (πχ. *Wolbachia*, που προκαλεί κυτταροπλασματική ιστοσυμβατότητα), απαιτούν πολύ μικρότερους αριθμούς. Το πλήθος των εντόμων που είναι ικανό να εξαλείψει ή να ελέγξει τα επίπεδα ενός φυσικού πληθυσμού αποτελεί καθοριστικό παράγοντα για τη φικτότητα της εφαρμογής μίας τέτοιας μεθόδου σε ευρεία κλίμακα, καθώς μεταφράζεται

οικονομικά σε εγκαταστάσεις, εξοπλισμό και αναλώσιμα, εξειδικευμένο προσωπικό που θα πραγματοποιήσει την εκτροφή, τις απελευθερώσεις και την επίβλεψη της διαδικασίας κατά τη διάρκεια και μετά τις απελευθερώσεις.

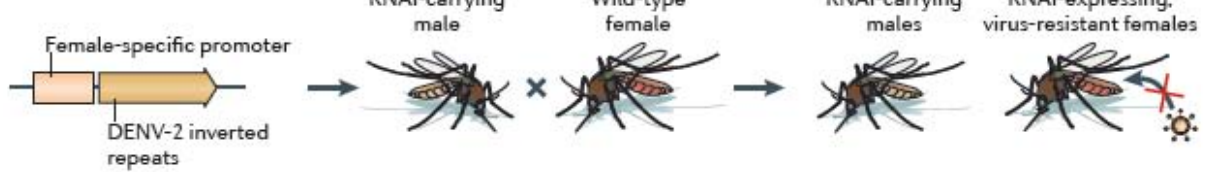
Συνοψίζοντας, η εμπειρία 65 ετών της έρευνας πάνω στην καταπολέμηση εντόμων αγροκτηνοτροφικού ή υγειονομικού ενδιαφέροντος δείχνει ότι καμία μέθοδος δεν είναι πανάκεια, ενώ τα καλύτερα αποτελέσματα μπορούν να αναμένονται από τη συνδυαστική χρησιμοποίηση συμβατικών τεχνικών (παγίδες – εντομοκτόνα), βιολογικών (εισαγωγή ειδών-εχθρών των εντόμων-στόχων στο μικροπεριβάλλον) και γενετικών προσεγγίσεων (SIT) (Global Strategic Framework For Integrated Vector Management, 2004). Επίσης, σε κάθε παρεμβατική προσέγγιση στο φυσικό πεδίο πρέπει να προηγείται, αλλά και να ακολουθεί πολύ αυστηρή παρακολούθηση της δυναμικής του οικοσυστήματος και των πιθανών μεταβολών στην οικολογία.

Η στροφή των μεθόδων καταπολέμησης προς βιοτεχνολογικές και γενετικές προσεγγίσεις αυτόματα καθιστά επιτακτική την ανάγκη για πληρέστερο μοριακό χαρακτηρισμό διαδικασιών των εντόμων-στόχων. Τα επιμέρους συστήματα, με επίκεντρο το αναπαραγωγικό, και η μεταξύ τους συντονισμένη αλληλεπίδραση μπορούν να ιδωθούν με μια εργαλειακή ματιά ως ακριβή σημεία ελέγχου (checkpoints) της δημιουργίας απογόνων, επόμενα και της πληθυσμιακής δυναμικής. Ο τελικός λοιπόν στόχος στη μελέτη της αναπαραγωγής για την καταπολέμηση ειδών εντόμων είναι η παρεμπόδιση ή ο αποπρογραμματισμός κάποιου μονοπατιού που εμφανίζεται εξειδικευμένα σε ένα στάδιο της αναπαραγωγής χωρίς τα παρελκόμενα της μη-ειδικής πρόκλησης γενετικών βλαβών.

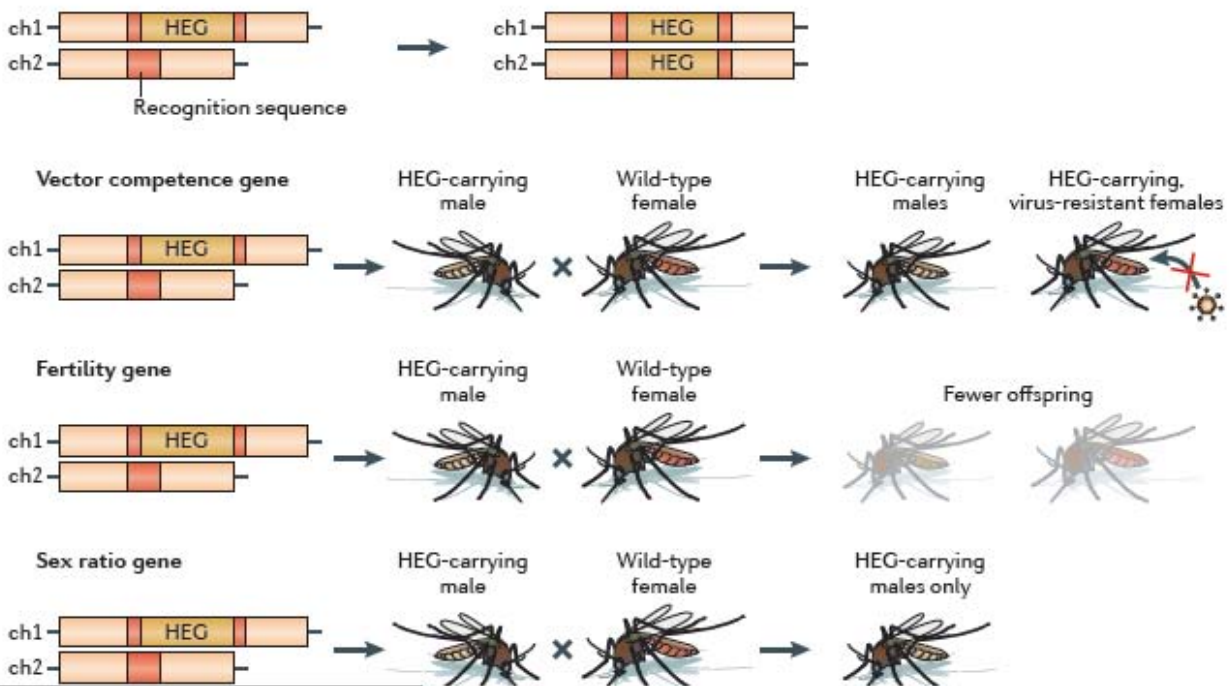
**a RIDL**



**b RNAi**



**c HEGs**



Εικόνα 1.1.1 Καταπολέμηση εντόμων με τις μεθόδους RIDL, RNAi και HEGs (McGraw EA, 2013)

## 1.2 Αναπαραγωγή

### 1.2.1 Πρόλογος

Υπό το πρίσμα της Εξέλιξης, όπως απαιτεί να παρατηρούμε τη Βιολογία ο Ντομπζάνσκι (Dobzhansky, 1973), ίσως ο Χρόνος είναι το πεδίο όπου επιτρέπεται στα μόρια να αλληλεπιδρούν και να συνθέτουν όλο και πιο περίπλοκες δομές οι οποίες με τη σειρά τους αλληλεπιδρούν η μία με την άλλη, συνιστώντας συνεχώς ένα υπερσύνολο που ισορροπεί ανάμεσα στο παρόν και το μέλλον, τη στατικότητα και τη δυναμικό στοιχείο, το προσαρμοσμένο και το προσαρμοζόμενο. Η ισορροπία αυτή πιθανότατα ξεκίνησε από μία ανόργανη σούπα τυχαίων και απλοϊκών αντιδράσεων (Miller, 1953), με το κρισιμότερο ίσως σημείο της μετάβασης από τη μη-ζωή στη ζωή να είναι η δημιουργία της λιπιδικής διπλοστιβάδας (Tessera, 2009) (Monnard, 2015). Αυτή μείωσε το διαθέσιμο χώρο, οπότε αύξησε τον, πρωτόγονο ακόμα, εξελικτικό ρυθμό. Απομονώνοντας, μάλλον με τυχαίο τρόπο, μόρια μέσα σε μεμβράνες, στα πρώιμα κύτταρα, το παιχνίδι της αλληλεπίδρασης σα να απέκτησε δομικές μονάδες, μα ποτέ δεν έγινε πιο σύνθετο: Πετυχημένο είναι κάθε σύστημα που οργανώνεται με τρόπο αρκούντως λειτουργικό ώστε να διατηρηθεί στο ενδιαίτημά του και να αυτοαντιγραφεί – Κάθε δυνατή αλληλεπίδραση είναι αποδεκτή – Οι χαμένοι χάνονται.

Στους σημερινούς ζωντανούς οργανισμούς διακρίνονται τα συστήματα που είναι προς το παρόν οι νικητές αυτής της ατέλειωτης διαδικασίας. Και ο τρόπος που «βρίσκει» κάθε είδος για να εξασφαλίσει τη συνέχισή του, η αναπαραγωγή του, δεν αποτελεί απλά ένα χαρακτηριστικό του Έμβριου, ούτε μονάχα μία περίοδο στη ζωή του· είναι μια έννοια συνυφασμένη με την ανάπτυξη του από τη στιγμή που υπήρξε ζυγωτό μέχρι τη στιγμή που θα συζευχθεί. Φαίνεται λοιπόν, πάντα υπό το πρίσμα της Εξέλιξης, ότι η αναπαραγωγή είναι η μηχανή της ζωής. Έτσι καταδεικνύεται ότι η κατανόηση της βιολογίας της αναπαραγωγής μέσα από τη βασική έρευνα αποτελεί αναγκαιότητα συνολικά για την ανθρωπότητα και πηγή έμπνευσης για τον εκάστοτε ερευνητή.

Στο πεδίο του ελέγχου εντόμων που αποτελούν παράσιτα της αγροκτηνοτροφίας (Integrated Pest Management, IPM) ή φορείς ασθενιών (Integrated Vector Management, IVM), η μελέτη της αναπαραγωγής είναι απαραίτητη για την ταυτοποίηση των μορίων κλειδιών και των αλληλεπιδράσεων μεταξύ τους. Όσο καλύτερα μπορούμε να χαρακτηρίσουμε τα επιμέρους στάδια της αναπαραγωγής, τόσο πιο έξυπνες, αποδοτικές, ακριβείς και οικονομικές προσεγγίσεις ελέγχου είμαστε ικανοί να σχεδιάσουμε και να υλοποιήσουμε.

### 1.2.2 Η μοριακή φυσιολογία της σύζευξης

Η σύζευξη είναι ένα σύνθετο σύνολο διεργασιών που ακόμα δεν είναι απόλυτα χαρακτηρισμένες και κατανοητές. Για την επιτυχία της είναι απαραίτητος ο κατάλληλος χωροχρονικός εντοπισμός και η αλληλεπίδραση μεγάλου αριθμού μορίων πατρικής και μητρικής προέλευσης. Στα έντομα, τα στάδια της αναπαραγωγής έχουν μελετηθεί αρκετά στην κοινή μύγα των φρούτων *Drosophila melanogaster*, που αποτελεί οργανισμό-μοντέλο, ενώ έντομα που επηρεάζουν σημαντικά την ανθρώπινη υγεία και δραστηριότητα, όντας φορείς παθογόνων ή παράσιτα αγροκτηνοτροφικής σημασίας, αποτελούν τους πιο άψιμους οργανισμούς υπό μελέτη.

Κατά τη σύζευξη, τα σπερματοζωάρια που βρίσκονται σε μία υγρή φάση, που ονομάζεται σπερματικό υγρό (seminal fluid), μεταφέρονται στο θηλυκό. Στο σπερματικό υγρό συμπεριλαμβάνονται επίσης μεμβρανικά κυστίδια (Corrigan, 2014), καθώς και μόρια που αποκαλούνται πρωτεΐνες του σπερματικού υγρού (Seminal Fluid Proteins, SFPs), που σχετίζονται με την επιτυχή αποθήκευση του σπέρματος (Qazi & Wolfner, 2003) στα θηλυκά εξειδικευμένα όργανα, τις σπερματοθήκες (spermatheca)

και το σπερματικό δοχείο (seminal receptacle), τη διατήρησή του στην κατάλληλη φυσικοχημική κατάσταση (Wong, 2008), την απελευθέρωσή του (Avila, Mattei, & Wolfner, 2010), την ανοσολογική απόκριση του θηλυκού (Findlay, Yi, Maccoss, & Swanson, 2008) και τη λεγόμενη μετασυζευκτική απόκριση (post-mating response), όπως ονομάζεται συλλογικά μία σειρά ιδιαίτερων συμπεριφορικών και φυσιολογικών μεταβολών που προκαλούνται στο θηλυκό έντομο από τη σύζευξη. Στη *Drosophila melanogaster*, χαρακτηρίζεται από αύξηση των ρυθμών ωοπαραγωγής και ωοαπόθεσης και μείωση της δεκτικότητας για επανασύζευξη για τουλάχιστον 5 ημέρες (Mapping, 1962).

Ο αναπαραγωγικός σωλήνας του θηλυκού εντόμου είναι ένα σύνθετο όργανο, του οποίου οι υποπεριοχές έχουν ξεχωριστές μορφολογικές και λειτουργικές ταυτότητες μετασυζευκτικά. Μέχρι το 2015, η μελέτη τους γινόταν με τομές ιστών, γεγονός που απέτρεπε την παρατήρηση των αλληλεπιδράσεων και του συντονισμού μεταξύ των επί μέρους οργάνων του αναπαραγωγικού σωλήνα *in situ* μεταξύ υπό-σύζευξη ή συζευγμένων μυγών. Με microCT scans σε *Drosophila melanogaster* αγρίου τύπου και σε knock-outs, οι (Mattei, Riccio, Avila, & Wolfner, 2015) κατόρθωσαν να οπτικοποιήσουν τις συντονισμένες αναδιαμορφώσεις των αναπαραγωγικών οργάνων από τη σύζευξη μέχρι την ωοαπόθεση και να αποκτήσουν μορφολογικές πληροφορίες που συμπληρώνουν τα μοριακά δεδομένα. Η σύζευξη πραγματοποιείται με μεγάλη ακρίβεια, με τους κολπικούς οδόντες (vaginal teeth) να συναρμολογούν με τα άγκιστρα (claspers) του αρσενικού. Ένα πρώιμο φυσιολογικό αποτέλεσμα της είναι η πρόκληση ενός συζευκτικού τραύματος στον κόλπο του θηλυκού από το εισδύον όργανο (intromittent organ) του αρσενικού, γεγονός που δικαιολογεί την παρουσία μορίων του σπερματικού υγρού στην αιμολήμφα και στόχων τους εκτός του αναπαραγωγικού σωλήνα. Η ωοπαραγωγή πραγματοποιείται κατά κύματα, ή παρτίδες, καθώς εκκρίματα τόσο του αρσενικού (Heifetz Y. L., 2000) όσο και του θηλυκού (Sun, 2013) ελέγχουν τον αριθμό των ωριμων ωοκυττάρων στις ωοθήκες. Οι SFPs επιδρούν στη μωλική συστολή, προκαλώντας κινήσεις των σαλπίνγων με πιθανό ρυθμιστικό ρόλο στη διέλευση του ωαρίου προς τη μήτρα, καθώς και αύξηση του όγκου της μήτρας, που θα διευκολύνει την κατάλληλη τοποθέτηση του ωαρίου για να γονιμοποιηθεί.

### 1.2.3 To Sex Peptide

To Sex Peptide (SP, ή Acp70A) είναι ένα πεπτίδιο μήκους 55 αμινοξέων που έχει ταυτοποιηθεί ως ο σημαντικότερος μοριακός διακόπτης της μετασυζευκτικής απόκρισης. Παράγεται από τους αναπαραγωγικούς αδένες του αρσενικού και αποτελεί συστατικό του σπερματικού υγρού (seminal fluid) που μεταφέρεται στο θηλυκό άτομο κατά τη σύζευξη. Έχει αποδειχθεί με πειράματα σίγησης ότι η μειωμένη έκφρασή του στα αρσενικά οδηγεί σε μειωμένη έκταση μετασυζευκτικά γεγονότα στο θηλυκό (Charpman, et al., 2003) και σε αδυναμία απελευθέρωσης του αποθηκευμένου σπέρματος (Avila, Mattei, & Wolfner, Sex peptide is required for the efficient release of stored sperm in mated *Drosophila* females, 2010), ενώ η ένεση καθαρισμένου ή συνθετικού SP σε θηλυκά παρθένα άτομα (Chen, et al., 1988) είναι αρκετή από μόνη της για να προκαλέσει μία βραχυπρόθεσμη μετασυζευκτική απόκριση μίας με δύο ημερών. Επίσης, το SP φαίνεται να επάγει μηχανισμούς που επιταχύνουν τη διέλευση της τροφής στο γαστρεντερικό σωλήνα (Arger-McGlaughon, 2013) και να μειώνει το χρόνο του απογευματινού ύπνου (siesta sleep). Το siesta sleep είναι μια συμπεριφορά που εμφανίζει σεξουαλικό διμορφισμό – παρατηρείται κυρίως στα αρσενικά έντομα και λιγότερο σε



παρθένα, σεξουαλικά μη-ώριμα θηλυκά – ενώ η σύζευξη και συγκεκριμένα το SP αναστέλλουν αυτή τη συμπεριφορά, διατηρώντας το θηλυκό σε εγρήγορση και προσφέροντας ένα πλεονέκτημα απέναντι σε περιβαλλοντικές απειλές και θηρευτές (Isaac, 2010). Τέλος, το SP είναι υπεύθυνο για μεταβολές στις διατροφικές συνθήκες του θηλυκού, συγκεκριμένα, για την αύξηση της κατανάλωσης τροφής και την προτίμηση σε τροφές που περιέχουν ζύμες, που αποτελούν πηγή πρωτεϊνών και θεωρείται ότι διευκολύνουν την ωοπαράγωγη (Carvalho, Karahi, Anderson, & Benzer, 2006).

Υπό φυσιολογικές συνθήκες, το SP μεταφέρεται στο θηλυκό και σε ελεύθερη μορφή, αλλά και συζευγμένο με τις ουρές των σπερματοζωαρίων, από όπου αποδεσμεύεται σταδιακά όσο εκείνα βρίσκονται αποθηκευμένα, οδηγώντας σε μία παράταση της μετασυσζευκτικής απόκρισης (Peng, 2005). Το ελεύθερο SP είναι ανιχνεύσιμο και στην αιμοδέμφο του θηλυκού, για ένα μικρό διάστημα, μέχρι μιας ημέρας (Pirpel, 2008).

Το SP για να δράσει είναι απαραίτητο να προσδεθεί στο Sex Peptide Receptor (SPR), έναν υποδοχέα GPCR που εκφράζεται στους νευρικούς ιστούς και στον αναπαραγωγικό σωλήνα του θηλυκού. Η παρεμβολή στην έκφραση του SPR οδηγεί σε παρόμοιο φαινότυπο με την απουσία του SP (Avila, Mattei, & Wolfner, Sex peptide receptor is required for the release of stored sperm by mated *Drosophila melanogaster* females, 2015).

#### 1.2.4 Άλλες SFPs

Πέρα από το SP έχουν ταυτοποιηθεί και άλλα μόρια SFPs (ή ACPs, Accessory Gland Proteins) που συμβάλλουν ποικιλοτρόπως στη μετασυσζευκτική απόκριση. Η οβουλίνη (Acp26Aa) αυξάνει το ρυθμό ωορρηξίας (Heifetz, Lung, Frongillo, & Wolfner, 2000), η Acp36DE την αποθήκευση του σπέρματος (Neubaum & Wolfner, 1999), η CG33943 είναι απαραίτητη για την πλήρη ενεργοποίηση της ωοαπόθεσης, ενώ τέσσερις ακόμα ACPs (CG1652, CG1656, CG17575, CG9997) απαιτούνται για τη συνεχόμενη ωοαπόθεση, την απελευθέρωση του σπέρματος από τα αποθηκευτικά όργανα και τη μειωμένη δεκτικότητα στη επανασύζευξη (Ram & Wolfner, 2007). Οι λειτουργίες της πλειοψηφίας των ACPs παραμένουν άγνωστες και αποτελούν αντικείμενο ενεργούς μελέτης.

#### 1.2.5 Μετασυσζευκτική απόκριση και άλλοι ιστοί

Παρόλο που η σύζευξη και τα κυριότερα μετασυσζευκτικά γεγονότα λαμβάνουν χώρα στους αναπαραγωγικούς αδένες του θηλυκού, πειράματα μικροσυστοιχιών έχουν αποδείξει ότι και άλλοι ιστοί του υπόκεινται σε μεγάλη έκτασης μεταγραφικές μεταβολές λόγω της σύζευξης.

Φαίνεται ότι οι ACPs κυκλοφορούν στην αιμοδέμφο και επηρεάζουν τη γονιδιακή έκφραση σε ιστούς στο κεφάλι ήδη από τις δύο ώρες μετά τη σύζευξη (Dalton, 2010). Καταστέλλονται μόρια που σχετίζονται με το στρες (Heat-shock proteins) και την αποτοξικοποίηση (τρανσφεράσες της γλουταθειόνης), ενώ επάγονται γονίδια που σχετίζονται με το μεταβολισμό της ντοπαμίνης, που σχετίζεται με τη δεκτικότητα του θηλυκού στη σύζευξη, μεταγραφικοί παράγοντες σχετικοί με την ανάπτυξη

(diminutive), και θετικοί ρυθμιστές της διάθεσης για θρέψη (short neuropeptide F precursor).

Στις 24 ώρες μετά τη σύζευξη, οι περισσότερες διεργασίες που υπεραντιπροσωπεύονται αφορούν την ενεργοποίηση του μεταβολισμού και της πρόσληψης θρεπτικών συστατικών. Η ενεργοποίηση του μεταβολισμού είναι επίσης αισθητή και στον ιστό fat body. Πρόκειται για έναν ιστό ανάλογο του λιπώδους ιστού των θηλαστικών, ο οποίος αφενός αποτελεί έναν δυναμικό αποθηκευτικό χώρο για γλυκογόνο, λιπίδια και πρωτεΐνες που ανταποκρίνεται στις εκάστοτε ενεργειακές ανάγκες (Liu, et al., 2009), αφετέρου, μπορεί να επηρεάσει παρακρινώς άλλους ιστούς, απελευθερώνοντας ουσίες στην αιμοδέμφο (Lazareva AA, 2007).

Στον εγκέφαλο επηρεάζονται, μεταξύ άλλων, κάποια μόρια που κωδικοποιούν για τασεο-εξαρτώμενα κανάλια ιόντων (pHCl, shaker, cacophony, η β-υπομονάδα της πρωτεΐνης Ca<sup>2+</sup>-channel, το ομόλογο 1 της tipE), για καθοδήγηση του νευράξονα (smooth, φωσφατάση της τυροσίνης 99a), και για μεταφορά νευροδιαβιβαστών (complexin) γεγονός που υποδεικνύει νευροφυσιολογικές μετασυσζευκτικές μεταβολές (Dalton, 2010).

#### 1.2.6 Πλαστικότητα και διαειδική ποικιλομορφία της αναπαραγωγής των εντόμων

Η *Drosophila melanogaster* αποτελεί τον οργανισμό μοντέλο και όλες οι μελέτες που έχουν πραγματοποιηθεί σε άλλα έντομα γίνονται με αυτήν ως σημείο αναφοράς και κυρίως επικεντρώνονται στην ταυτοποίηση ομόλογων SFPs. Μελέτες έχουν πραγματοποιηθεί στο σπερματικό υγρό των κουνουπιού *Anopheles gambiae* (Dottorini, 2007), της μέλισσας *Apis mellifera* (Baer, 2009), της μεσογειακής μύγας *Ceratitis capitata* (Scolari, 2012) και του γρύλου *Teleogryllus oceanicus* (Simmons, 2013). Τόσο η φαινοτυπική πλαστικότητα σε όλες τις μορφοφυσιολογικές πτυχές της σύζευξης (Hodin, 2009), όσο και τα μοριακά δεδομένα που μαρτυρούν ότι τα μόρια που σχετίζονται με την αναπαραγωγή εξελίσσονται ταχύτατα (Swanson, 2001) (Begun, 2000), υποδεικνύουν πως κάθε έντομο υπό μελέτη αποτελεί ένα νέο σύστημα, στο οποίο οι προσδοκίες να βρεθούν πανομοιότυποι μηχανισμοί επίτευξης της αναπαραγωγικής διαδικασίας πρέπει να παραμένουν χαμηλές.

Παρόλα αυτά, ομόλογα του SP και άλλων μορίων που έχουν ταυτοποιηθεί ως ACPs στη *Drosophila melanogaster*, έχουν βρεθεί σε άλλα έντομα (Bono, Matzkin, Hoang, & Brandsmeier, 2015), υποδεικνύοντας πως κύριοι ρυθμιστές κάποιων βημάτων της μετασυσζευκτικής απόκρισης είναι συντηρημένοι σε κάποιο βαθμό. Το μικρό μέγεθος του Sex Peptide, καθώς και το γεγονός ότι στη διακυτταρική επικοινωνία συχνά εμπλέκονται με ρόλο εξωκυτταρικού προσδέτη μικρού μοριακού βάρους πρωτεΐνες, πεπτιδικές ορμόνες και νευροδιαβιβαστές (Russell, Hertz, & McMillan, 2011), (Lodish, Berk, & Zipursky, 2000), υπογραμμίζουν τη μελέτη του προτεόματος και δη του μικρού προτεόματος, ή πεπτιδόματος, ως ερευνητικό πεδίο απαραίτητο για τη διαλεύκανση των μοριακών παικτών της αναπαραγωγής.

### 1.3 Πρωτεομική και πεπτιδομική

#### 1.3.1 Φασματοσκοπία μάζας (Picó, 2015)

Η φασματοσκοπία μάζας αποτελεί ένα ισχυρό και ευέλικτο αναλυτικό εργαλείο για την ταυτοποίηση μιας άγνωστης ουσίας, το μοριακό της βάρος, τη σύστασή και τη χημική της δομή. Μπορεί να χωριστεί στη στοιχειακή, ή ατομική, φασματοσκοπία, η οποία καθιστά ικανή την αναγνώριση και ποσοτικοποίηση σχεδόν όλων των στοιχείων του περιοδικού πίνακα, και στη μοριακή φασματοσκοπία μάζας, που παρέχει πληροφορίες για τη δομή οργανικών και ανόργανων μορίων και την ποιοτική ή ποσοτική σύσταση σύνθετων μιγμάτων.

Στο φασματογράφο, τα μόρια του αναλύτη μετατρέπονται σε ιόντα εξαιτίας της ενέργειας που τους ασκείται. Τα ιόντα διαχωρίζονται με βάση το λόγο μάζας/φορτίου ( $m/z$ ) και κατευθύνονται σειριακά ή ταυτόχρονα προς έναν μετατροπέα που μετατρέπει τον αριθμό των ιόντων (ion abundance) σε ηλεκτρικό σήμα. Η συνάρτηση του αριθμού των ιόντων προς το λόγο  $m/z$  συνιστά ένα φάσμα μάζας. Στην περίπτωση που τα ιόντα που παράγονται από την πηγή ιονισμού έχουν μονό φορτίο, το φάσμα είναι η αποτύπωση του αριθμού ιόντων προς τη μάζα.

Οι μάζες στη φασματοσκοπία εκφράζονται σε ατομικές μονάδες μάζας (amu) ή Ντάλτον (Da), όπου  $1\text{amu} = 1\text{Da} = 1/12$  της μάζας του ουδέτερου ατόμου του  $^{12}\text{C}$ .

Ο λόγος  $m/z$  είναι μία αδιάστατη τιμή (δεν έχει μονάδες μέτρησης) που ισοδυναμεί με τη μάζα του εκάστοτε ιόντος προς το φορτίο του. Έτσι, αν, π.χ., ένα μονά φορτισμένο ιόν έχει μάζα 71.08Da, έχει τιμή  $m/z = 71.08$ , ενώ αν το ίδιο ιόν είναι διπλά φορτισμένο, έχει  $m/z = 71.08/2 = 35.54$ .

Παρόλο που υπάρχουν πολλών ειδών φασματογράφοι που εξυπηρετούν καλύτερα διαφορετικές ερευνητικές ανάγκες, τα μέρη τους σε μεγάλο βαθμό παραμένουν τα ίδια. Πρόκειται για:

- το σύστημα εισόδου, μέσω του οποίου τα δείγματα, στερεά, υγρά, ή αέρια, ανάλογα με το είδος της πηγής ιονισμού, εισέρχονται στην περιοχή κενού του φασματογράφου,
- την πηγή ιονισμού, όπου τα συστατικά του δείγματος μετατρέπονται με βομβαρδισμό ηλεκτρονίων, φωτονίων, ιόντων, ή μορίων σε αέρια ιόντα,
- τον αναλυτή μάζας, στον οποίο τα αέρια ιόντα διαχωρίζονται βάσει του λόγου  $m/z$  τους
- το μετατροπέα ιόντων που τα συλλέγει και μετατρέπει σε ηλεκτρικό σήμα
- ένα υπολογιστικό σύστημα που ελέγχει τις παραμέτρους της φασματογράφησης, αναλύει τα δεδομένα και παράγει τα φάσματα

Ο ιονισμός, η ανάλυση μάζας και η συλλογή των ιόντων γίνονται πάντα υπό συνθήκες χαμηλής πίεσης ώστε να ελαχιστοποιείται η πιθανότητα να συγκρουστούν μεταξύ τους τα προκύπτοντα ιόντα.

#### 1.3.2 Είδη πηγών ιονισμού (Picó, 2015)

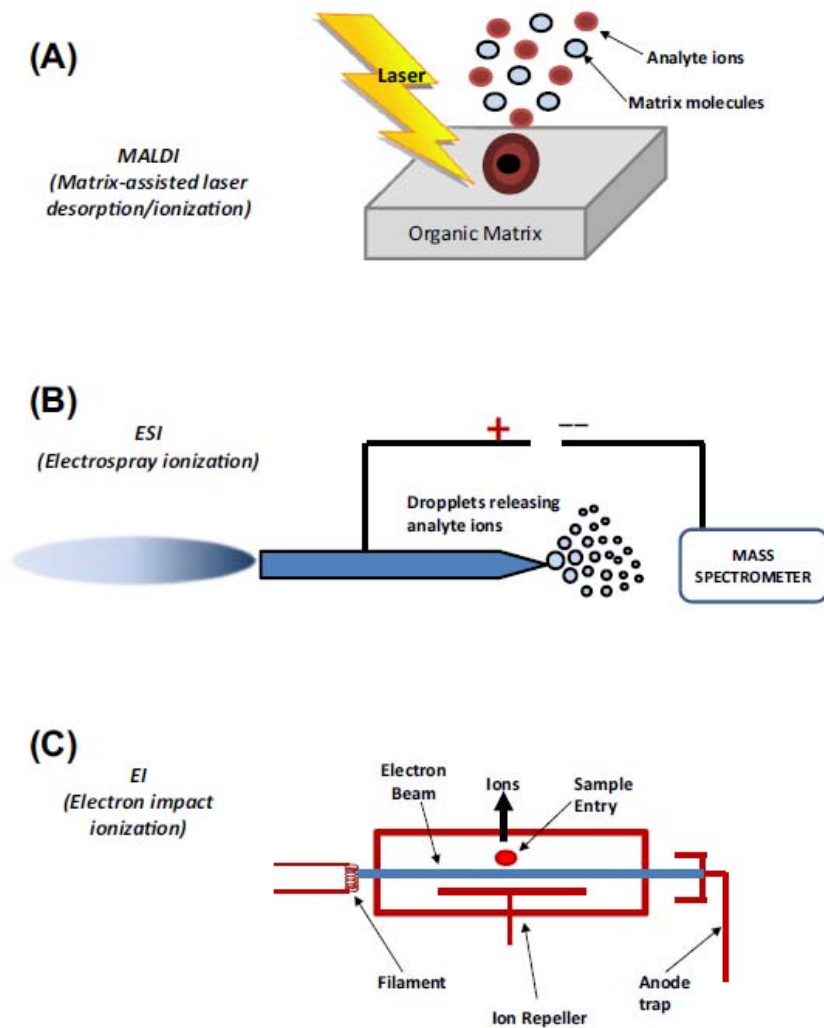
Η φασματογράφηση ξεκινά με τη δημιουργία αέριων ιόντων του αναλύτη κατά τη διαδικασία του ιονισμού. Οι μέθοδοι ιονισμού χωρίζονται σε δύο κατηγορίες, τις πηγές αέριας φάσης (gas-phase sources) και τις πηγές εκρόφησης (desorption sources). Στις πηγές αέριας φάσης ο αναλύτης εξατμίζεται και στη συνέχεια ιονίζεται, ενώ στις πηγές εκρόφησης το δείγμα, σε στερεή ή υγρή κατάσταση, μετατρέπεται απ' ευθείας σε αέρια ιόντα.

Η πιο ευρέως χρησιμοποιούμενη μέθοδος ιονισμού, που ανήκει στην πρώτη κατηγορία, αποκαλείται πρόσκρουση ηλεκτρονίων (electron impact, EI). Συνιστά το βομβαρδισμό των μορίων με ηλεκτρόνια υψηλής ενέργειας και την παραγωγή ανιόντων, κατιόντων και ουδέτερων ριζών, από τα οποία τα κατιόντα κατευθύνονται στον αναλυτή μάζας. Κατά την EI, η ακτίνα ηλεκτρονίων είναι τόσο ισχυρή που προκαλεί πολλά θραύσματα, χρήσιμα για τον προσδιορισμό των μοριακών ριζών των υπό ανάλυση μορίων.

Οι πιο δημοφιλείς μέθοδοι (Εικόνα 1.3.1) από τις πηγές εκρόφησης είναι ο ιονισμός με ηλεκτροψεκασμό (Electrospray Ionization, ESI) και ο ιονισμός εκρόφησης laser με τη βοήθεια υλικού μήτρας (Matrix Assisted Laser Desorption Ionization, MALDI). Στον ESI τα μόρια του αναλύτη ψεκάζονται υπό την επίδραση υψηλής τάσης και, φορτισμένα, φτάνουν στον αναλυτή. Ο ιονισμός MALDI είναι μία μέθοδος "μαλακού ιονισμού" κατά την οποία χρησιμοποιούνται ασυνεχείς, κοντοί παλμοί laser αέριου αζώτου, συνήθως στα 237nm, για να ιονιστεί ο αναλύτης. Το υλικό μήτρας σχηματίζει κρυστάλλους με τον αναλύτη και είναι απαραίτητο για τον ιονισμό του, ενώ κατά τη διαδικασία ιονίζεται και το ίδιο. Ένα ακόμα σημαντικό χαρακτηριστικό αυτού του είδους ιονισμού είναι το γεγονός ότι παράγει σχεδόν αποκλειστικά μονά φορτισμένα ιόντα, των οποίων η μετα-ανάλυση είναι πιο εύκολη.

Ιδανικό θα ήταν ένας αναλυτής μάζας να έχει τη δυνατότητα να ξεχωρίζει απειροελάχιστες διαφορές στη μάζα και ταυτόχρονα να επιτρέπει τη διέλευση ικανού αριθμού ιόντων, ώστε να παράγονται μετρήσιμα ρεύματα ιόντων. Επειδή οι δύο αυτές ιδιότητες δεν είναι απολύτως συμβατές μεταξύ τους, έχουν αναπτυχθεί πολλοί διαφορετικοί τύποι αναλυτών.

Οι αναλυτές μαγνητικού τομέα (sector analyzers) διαχωρίζουν με βάση την εκτροπή των ιόντων από ένα μαγνητικό πεδίο. Η τροχιά κάθε ιόντος εξαρτάται από την τιμή  $m/z$  του και το μαγνητικό πεδίο μεταβάλλεται αργά, ώστε να φέρει ιόντα διαφορετικών  $m/z$  στον ανιχνευτή. Υπάρχουν και μαγνητικοί αναλυτές διπλής εστίασης, στους οποίους ένας ηλεκτρικός τομέας που προηγείται του μαγνητικού ασκεί ένα ηλεκτροστατικό πεδίο και εστιάζει ιόντα που χαρακτηρίζονται από πολύ μικρό εύρος κινητικής ενέργειας στην οπή που οδηγεί στο μαγνητικό τομέα, επιτυγχάνοντας έτσι πολύ υψηλή διακριτική ικανότητα.

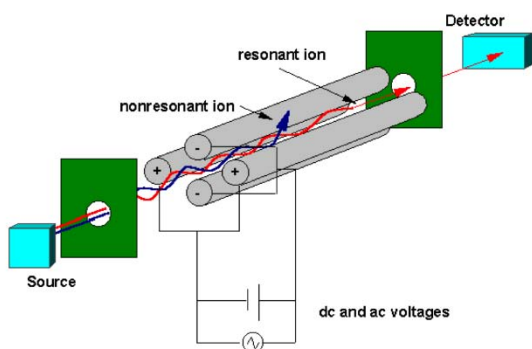


Εικόνα 1.3.1 Οι πιο συνήθεις πηγές ιονισμού: MALDI, ESI και EI (Picó, 2015)

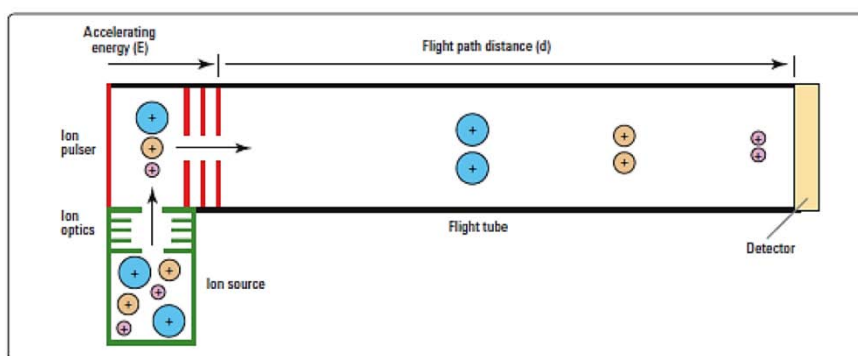
Οι τετραπολικί αναλυτές (quadrupole mass analyzers) αποτελούνται από τέσσερις κυλινδρικές ράβδους, συμμετρικά τοποθετημένες περιμετρικά της δέσμης ιόντων (Εικόνα 1.3.2). Στις δύο εφαρμόζεται εναλλασσόμενο ρεύμα και στις άλλες δύο συνεχές και ρυθμίζοντας την τάση επιτρέπεται κάθε φορά σε ιόντα διαφορετικού  $m/z$  να εξέλθουν προς το μετατροπέα ιόντων. Έτσι επιτρέπεται σε πολύ μικρό χρόνο να σαρωθεί ένα εύρος  $m/z$ . Χαρακτηρίζονται όμως από χαμηλή διακριτική ικανότητα. Στους αναλυτές χρόνου πτήσης (time-of-flight/TOF analyzers) ασκείται στα ιόντα μία σταθερή κινητική ενέργεια και στη συνέχεια αφήνονται να διασχίσουν ένα σωλήνα κενού που καταλήγει στον ανιχνευτή (Εικόνα

1.3.3). Αφού  $K=1/2mv^2$ , η ταχύτητα κίνησής τους είναι αντιστρόφως ανάλογη της μάζας τους· τα ελαφρύτερα ιόντα ανιχνεύονται πρώτα και τα βαρύτερα στη συνέχεια. Οι αναλυτές χρόνου πτήσης είναι απλές συσκευές, που επιτρέπουν ταχύτερη ανάλυση και δεν περιορίζονται σε ένα εύρος μαζών. Είναι όμως σχετικά χαμηλής διακριτικής ικανότητας.

Υπάρχουν ακόμα οι παγιδευτές ιόντων (ion trap) και οι αναλυτές ιοντικού κυκλοτρονικού συντονισμού με μετασχηματισμό Fourier (Fourier transform ion cyclotron resonance analyzers), που έχουν ιδιαίτερα υψηλό κόστος και περίπλοκη κατασκευή, αλλά είναι εξαιρετικά ακριβή.



Εικόνα 1.3.2 Τετραπολικοί αναλυτές



Εικόνα 1.3.3 Αναλυτές Χρόνου Πτήσης

Στη μοριακή φασματοσκοπία το σύνηθες είναι να πραγματοποιείται διαδοχική φασματογράφιση (tandem MS ή MS/MS), δηλαδή να υπάρχουν δύο αναλυτές μάζας συνδεδεμένοι σε σειρά. Η πηγή ιονισμού παράγει μοριακά ιόντα, αλλά και ιόντα θραύσης και φέρονται στον πρώτο αναλυτή μάζας. Από αυτά επιλέγεται ένα συγκεκριμένο ιόν, το πρόδρομο (precursor ion), το οποίο στον πυρήνα αλληλεπίδρασης θα διασπαστεί, είτε αυθόρμητα, είτε αντιδρώντας με κάποιο αέριο, είτε μετά από ακτινοβολία σε θραύσματα, τα παράγωγα ή θυγατρικά ιόντα (product/daughter ions). Τα θυγατρικά ιόντα αναλύονται από το δεύτερο αναλυτή μάζας και ανιχνεύονται από το μετατροπέα. Η διαδοχική φασματοσκοπία παράγει τεράστιες ποσότητες πληροφορίας, που επιτρέπουν τη δομική επεξήγηση ενός μορίου, καθώς και την ανάλυση μιγμάτων. Κατά κανόνα, απαιτείται χρωματογραφία ή διαχωρισμός με ηλεκτροφόρηση για να παρουσιάζεται μία ένωση κάθε φορά στο φασματογράφο.

### 1.3.3 Πεπτιδομική

Παρότι το πρώτο νευροπεπτίδιο που ανακαλύφθηκε ήταν η ουσία P το 1931, χρειάστηκε σχεδόν 70 χρόνια για να εμφανιστεί στη βιβλιογραφία ο όρος πεπτιδομική, που περιγράφει τη μελέτη του συνόλου των μικρών ενεργών πεπτιδίων που παράγονται και δρουν σε κάποιο κύτταρο, ιστό ή οργανισμό μία χρονική στιγμή. Μέσα σε αυτή τη χρονική διάρκεια έγιναν σημαντικές ανακαλύψεις στους χώρους της φυσικής, της χημείας και της μηχανικής που κατέστησαν δυνατή την κατασκευή φασματογράφων μάζας με τα κατάλληλα χαρακτηριστικά ώστε να μπορούν να αναλύσουν σύνθετα βιολογικά μίγματα. Στο ίδιο χρονικό πλαίσιο, η ποσότητα της πληροφορίας που διαχειριζόταν η επιστημονική κοινότητα εκτινάχθηκε στα ύψη, χάρη στο Πρόγραμμα Αλληλούχησης του Ανθρώπινου Γονιδιώματος και συναφή προγράμματα και το γεγονός αυτό απέδωσε

ολοκληρωμένες και περιεκτικές βάσεις δεδομένων που μπορούσαν να αξιοποιηθούν και για πρωτεομικές ή πεπτιδομικές μελέτες.

Από το 2000 μέχρι σήμερα, η ανάλυση του μικρού πρωτεόματος, του πεπτιδόματος, έχει σημειώσει αρκετές επιτυχίες στους τομείς της νευροενδοκρινολογίας και της αναζήτησης βιοδεικτών και φαρμακευτικών δραστικών ουσιών σχετικών με την ογκολογία, τις νευροεκφυλιστικές νόσους, την οστεοπόρωση, καρδιοαγγειακά νοσήματα, τη γονιμότητα, το μεταβολισμό των λιπιδίων οξέων και την ανοσολογία. Ωστόσο, ακόμα και σήμερα αναπτύσσεται στη σκιά της πρωτεομικής για μια σειρά λόγων:

1. Σε αντίθεση με τα πειράματα πεπτιδομικής όπου πραγματοποιείται πέψη του συνόλου των πρωτεϊνών σε μικρότερα πεπτιδία, συνήθως με τρυψίνη, πριν τη φασματογράφησή τους, στην πεπτιδομική αυτό δε συμβαίνει, καθώς έχουν ήδη απομονωθεί πεπτιδία μικρού μεγέθους. Αυτό έχει ως αποτέλεσμα ο χώρος της αναζήτησης να αυξάνεται πολλαπλάσια, καθώς δεν αναμένεται να εμφανιστούν τα άκρα των θραυσμάτων συγκεκριμένα αμινοξέα – τα σημεία κοπής.
2. Οι βαριές μετα-μεταφραστικές τροποποιήσεις που συχνά παρατηρούνται στα μικροπεπτιδία είναι παράγοντες που επηρεάζουν το φάσμα που προκύπτει και κάνουν τη μετα-ανάλυση δύσκολη.
3. Οι αναζητήσεις για πεπτιδία που συμβαδίζουν με τα δεδομένα του κάθε φάσματος γίνονται με τη χρήση πρωτεϊνικών ή τρανσκριπτομικών βάσεων δεδομένων, οι οποίες είναι φτωχές σε καθαρά πεπτιδικά δεδομένα. Αν κάποιο πεπτιδίο είναι προϊόν της μετάφρασης κάποιου μικρού αναγνωστικού πλαισίου (sORF), ή τμήμα πρωτεΐνης που δεν έχει ακόμα ταυτοποιηθεί, είναι απίθανο να βρεθεί.



4. Το σύστημα απόδοσης score των προγραμμάτων αναζήτησης είναι σχεδιασμένο για πειράματα πρωτεομικής: η απόδειξη ότι μία πρωτεΐνη υπάρχει σε ένα πρωτεϊνικό δείγμα είναι το να βρεθούν πολλαπλά φάσματα που να αντιστοιχούν σε πεπτίδια-θραύσματά της. Στην

πεπτιδομική λοιπόν παρατηρείται συχνά ομαδοποίηση των φασμάτων κάτω από μία πρωτεΐνη και μάλιστα με υψηλά scores, ενώ το καθένα στην πραγματικότητα μπορεί να έχει πιο ποιοτικές αντιστοιχίσεις με διαφορετικά πεπτίδια.

#### 1.4 Στόχοι:

Ο σκοπός της παρούσας διπλωματικής εργασίας ήταν να γίνει μια πρώτη ταυτοποίηση των μικρών πεπτιδίων και τμημάτων πρωτεϊνών που είναι παρόντα στους αναπαραγωγικούς αδένες του θηλυκού (Female Accessory Glands) κατά το ενήλικο στάδιο του δάκου της ελιάς *Bactrocera oleae*, στο οποίο έχει επέλθει αναπαραγωγική ωριμότητα. Η μελέτη των μορίων αυτών έγινε με το σκεπτικό ότι όσο πιο καλά χαρτογραφηθούν οι αναπαραγωγικές διαδικασίες που πραγματοποιούνται στο δάκο και τα εμπλεκόμενα σε αυτές μόρια, τόσο πιο κοντά βρισκόμαστε στο σχεδιασμό μίας αποτελεσματικής και οικονομικής μοριακής προσέγγισης καταπολέμησης του παρασίτου αυτού.

Έμφαση δόθηκε στη συζευκτική διαδικασία και για το λόγο αυτό πραγματοποιήθηκε ανάλυση απομονωμένων πεπτιδίων από ασύζευκτο και συζευγμένο αναπαραγωγικό ιστό. Τα μόρια που αναμενόταν να βρεθούν ανήκουν στις ακόλουθες ευρείες κατηγορίες. Η πρώτη αφορά μόρια που μεταφέρονται από το αρσενικό στο θηλυκό μέσω του σπερματικού υγρού (Seminal Fluid Proteins) κατά τη σύζευξη και προκαλούν τις απαραίτητες βιοχημικές και φυσιολογικές μετασυζευκτικές μεταβολές για να πραγματοποιηθεί επιτυχώς η γονιμοποίηση και τελικά η ωοαπόθεση. Η δεύτερη αποτελείται από μόρια

του ίδιου του θηλυκού που είτε είναι απαραίτητα για να είναι προετοιμασμένο και ώριμο το θηλυκό να πραγματοποιήσει με επιτυχία τη σύζευξη, είτε επάγονται λόγω της σύζευξης από τη δράση των SFPs και εμπλέκονται σε κάποια μετασυζευκτική μεταβολή.

Το γεγονός ότι η πεπτιδομική προσέγγιση που ακολουθήθηκε (MALDI/TOF-TOF) δεν έχει ποσοτικό χαρακτήρα αντιμετωπίστηκε εν μέρει αξιοποιώντας προηγούμενες αναλύσεις τρανσκριπτόματος που είχαν πραγματοποιηθεί στα FAGs, τα MAGs (Male Accessory Glands), τους όρχις (testes) και σε γαντρεντερικό ιστό αρσενικών και θηλυκών ενήλικων δάκων (GUT) που αποτέλεσε τον ιστό ελέγχου. Έτσι, επιχειρήθηκε μία σύνδεση προϋπάρχοντων RNA-seq δεδομένων, που προσφέρουν μία εικόνα της μεταγραφής των ανάλογων γονιδίων προσυζευκτικά και μετασυζευκτικά, με τα νέα δεδομένα της πεπτιδομικής.

Επίσης, χρησιμοποιήθηκαν ταυτοποιημένες πρωτεΐνες των αναπαραγωγικών αδένων του αρσενικού (Acps) των εντόμων *Drosophila melanogaster*, *Anopheles gambiae* και *Ceratitis capitata* σε μία προσπάθεια να ταυτοποιηθούν ομόλογές τους στα πεπτίδια που απομονώθηκαν.

## 2. Υλικά και μέθοδοι

## 2.1 Εκτροφή εντόμων και σύζευξη

Το εργαστηριακό στέλεχος δάκου που χρησιμοποιήθηκε προέρχεται από το Εθνικό Κέντρο Έρευνας Φυσικών Επιστημών "Δημόκριτος" και εκτρέφεται στο εντομοτροφείο μας από το 2000. Η εκτροφή (Sagri, et al., 2014) πραγματοποιείται σε ξύλινα κλουβιά 30×30×30εκ, θερμοκρασία 25°C, φωτοπερίοδο 12 ωρών και θρεπτικό μέσο που περιγράφεται στον Πίνακα 2.1 (Zygouridis, 2014). Για τη σύζευξη χρησιμοποιήθηκαν 7 ημερών (αναπαραγωγικά ώριμα) αρσενικά και θηλυκά έντομα. Για να εξασφαλιστεί ότι δεν είχαν επανασυζευχθεί, διαχωρίστηκαν στα δύο φύλα αμέσως μόλις εξήλθαν από τα βομβύκιά τους και μεγάλωσαν σε ξεχωριστά κλουβιά μέχρι τη στιγμή της αναπαραγωγής. Τότε τοποθετήθηκαν σε κοινό κλουβί και κάθε ζευγάρι που ξεπέρασε τη μισή ώρα σύζευξης θεωρήθηκε επιτυχημένο.

Τροφή Προνομφών (0.5kg)	Τροφή Ενηλίκων (0.424kg)		
Νιπαγίνη	1g	Υδρολυμένη Μαγιά	8g
Σορβικό Κάλιο	0.25g	Σακχαρόζη	32g
Ελαιόλαδο	10ml	Κρόκος Αυγού	2.4g
Tween 80	3.75ml		
Νερό (βρύσης)	275ml		
Μαγιά	37.5g		
Σόγια	15g		
Σακχαρόζη	10g		
HCl 10M	3ml+15ml H <sub>2</sub> O		
Κυτταρίνη	150g		

Τα θηλυκά διατηρήθηκαν απομονωμένα μετά τη σύζευξη για 12 ώρες και στη συνέχεια απομονώθηκαν σε στερεοσκόπιο οι αναπαραγωγικοί τους ιστοί.

## 2.2 Απομόνωση πεπτιδίων – λυοφιλίωση

Ακολούθησε καθαρισμός και απομόνωση των πεπτιδίων με διάλυμα απομόνωσης μεθανόλης/νερού/οξικού οξέος 90:9:1:

1. Πραγματοποιήθηκε ελαφριά φυγοκέντρηση (spin down), ώστε όλοι οι ιστοί να κατακαθίσουν στον πάτο του σωληναρίου και να μην μείνουν υπολείμματα τους στα τοιχώματα.
2. Ομογενοποιήθηκαν με pestle και μηχανικό αναδευτήρα και ακολούθησε θραύση των κυτταρικών μεμβρανών με υπερήχους (sonication).
3. Τα δείγματα φυγοκεντρήθηκαν για 10min, στα 7500g και συλλέχθηκε το υπερκείμενο. Το βήμα [3] επαναλήφθηκε άλλες δύο φορές.
4. Το υπερκείμενο συλλέχθηκε σε tips πιπέτας τύπου Ziptip C18 Millipore P10, που περιέχουν στήλη ικανή να συγκρατεί πεπτιδία και μικρού μοριακού βάρους πρωτεΐνες.
5. Ακολούθησαν εκλούσεις μέσα στο tip-στήλη, με διάλυμα έκλουσης 350μl Acetonitrile 100% και 150μl TFA 0.1%.
6. Η στήλη καθαρίστηκε σε νέο σωληνάριο με TFA 0.1% και πραγματοποιήθηκε λυοφιλίωση.

Τα Ziptips C18 είναι κατάλληλα για απομόνωση μικρών πεπτιδίων και πρωτεϊνών <50kDa. Το TFA (τριφλουορο-οξικό οξύ) είναι ένας παράγοντας ιοντικής δέσμωσης με πολύ ισχυρότερη δυνατότητα πρόσδεσης από το οξικό οξύ, που χρησιμοποιείται ευρέως στην υγρή χρωματογραφία. Το Acetonitrile αποτελεί έναν ικανοποιητικό διαλύτη για μεγάλο εύρος ιοντικών και μη πολικών μορίων.

## 2.3 HPLC αντίστροφης φάσης – MALDI-TOF/TOF φασματοσκοπία μάζας

Τα λυοφιλωμένα πεπτιδία αναδιαλύθηκαν σε 0.1% TFA, ακολούθησε vortex και sonication για 5min και φυγοκέντρηση στις 13000rpm για 10min, με σκοπό την καθίζηση τυχόν άχρηστων υπολειμμάτων. Τα υπερκείμενα μεταφέρθηκαν σε νέο σωληνάριο, από όπου ενέθηκαν στη στήλη της HPLC αντίστροφης φάσης για διαχωρισμό. Τα κλάσματα που προέκυψαν μεταφέρθηκαν σε σημεία (spots) της εξειδικευμένης πλάκας MALDI από μία εξειδικευμένη ρομποτική συσκευή (ProteinEer FC II; Bruker Daltonics) και ακολούθησε η ανάλυσή τους σε φάσματα από το φασματογράφο μάζας MALDI-TOF/TOF (Ultraflexxtreme; Bruker Daltonics). Τόσο για τα ασύζευκτα, όσο και για τα συζευγμένα έντομα, προέκυψε από ένα αρχείο (CompoundList.xml), όπου καταγράφονται βασικές τιμές κάθε φάσματος υπό έναν κωδικό και ένα αρχείο με τις κορυφές κάθε φάσματος (analysis.baf).

## 2.4 Αναζήτηση σε βάσεις δεδομένων με το Mascot

Τα φάσματα που προέκυψαν από τη διαδοχική φασματοσκοπία μάζας υπέστησαν την απαραίτητη επεξεργασία από τα προγράμματα BioTools και FlexAnalysis (της Bruker Daltonics) και ακολούθησε αναζήτηση σε βάσεις δεδομένων με το λογισμικό Mascot [έκδοση 2.4.1.] για να συσχετιστούν με υπάρχουσες πρωτεΐνες. Οι παράμετροι που χρησιμοποιήθηκαν από το Mascot, οι βάσεις δεδομένων και τα αρχεία που προέκυψαν παρουσιάζονται στον Πίνακα 2.2.

Digestion enzyme	None
Precursor mass tolerance	100ppm
Fragment mass tolerance	0.8Da
Modifications allowed	Carbamidomethyl (C) Oxidation (M)
Databases	-Bactrocera oleae transcriptome library -Swissprot [Drosophila (fruit flies)] -Swissprot [All entries] -Ceratitis capitata transcriptome library
Resulting files	-F07115_BM_against_Bactrocera_oleae.xml -F070154_BM_against_Drosophila.xml -F071150_BM_against_All.xml -F071152_BM_against_Ceratitis_capitata.xml

## 2.5 Χειροκίνητη επαλήθευση

Η χειροκίνητη επαλήθευση (manual validation) των αποτελεσμάτων μίας αναζήτησης σε βάσεις δεδομένων είναι ένα απαραίτητο βήμα μέχρι στιγμής για την αξιόπιστη εξαγωγή συμπερασμάτων. Πρόκειται για μια διαδικασία που είναι σε κάποιο βαθμό υποκειμενική, καθώς διάφορες εργαστηριακές ομάδες δίνουν βαρύτητα σε διαφορετικά κριτήρια για να δεχθούν ή να απορρίψουν μία αντιστοιχία.

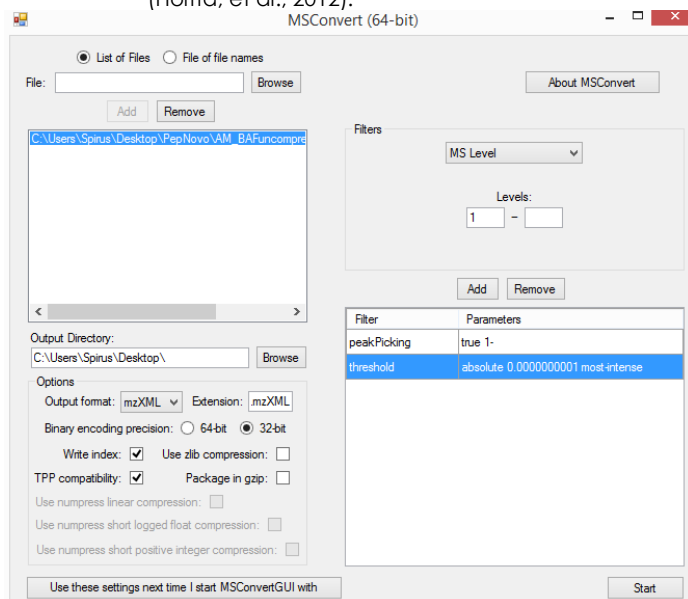
- Με τη χρήση του CompassXport της Bruker Daltonics έγινε μετατροπή των αρχείων analysis.baf στη μορφή .mzXML με την εξής εντολή, στο Command Prompt των Windows:  
>CompassXport.exe -a  
C:\Users\Spirus\Downloads\BM\BM\_BAF.d -o  
CompassXportOfAnalysisBaf -mode 0
- Στη συνέχεια τα αρχεία CompassXportOfAnalysisBaf.mzXML που προέκυψαν υπέστησαν επεξεργασία, ώστε να αφαιρεθούν τα φάσματα που περιείχαν μηδενικές τιμές και οι κορυφές οι οποίες λόγω

παρουσίας ισωτόπων του  $^{12}\text{C}$  είχαν αναλυθεί σε περισσότερες από μία, να αποκατασταθούν (deisotoping/deconvolution). Αυτή η επεξεργασία έγινε με το λογισμικό MSConvert της σουίτας ProteoWizard με τις παραμέτρους που φαίνονται στην Εικόνα 2.5.1 και τα αρχεία που προέκυψαν είναι τα *AM\_BAF32bitUncompressed.mzXML* και *BM\_BAF32bitUncompressed.mzXML*.

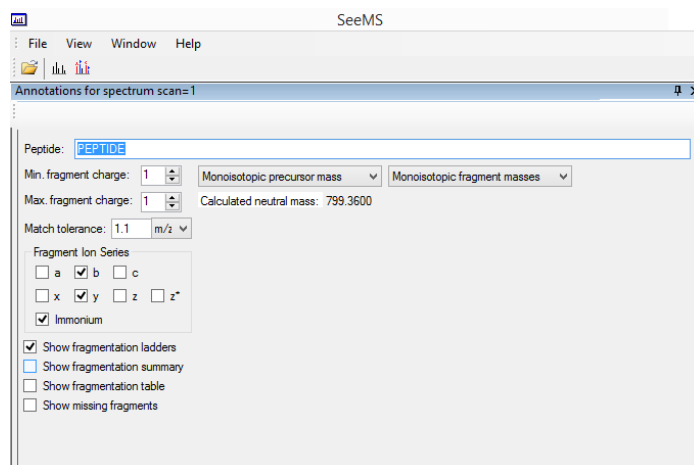
- Επειδή δεν υπήρχε πρόσβαση στο απαραίτητο λογισμικό της Mascot για την πραγματοποίηση της χειροκίνητης επαλήθευσης, αναπτύχθηκαν και χρησιμοποιήθηκαν μικρά προγράμματα στη γλώσσα προγραμματισμού Perl, ώστε να συσχετιστούν οι κωδικοί των φασμάτων των *CompoundList.xml* με τους αριθμούς σάρωσης από το *A/BM\_BAF32bitUncompressed.mzXML* και τις Αντιστοιχίες-Φάσματος-Πεπτιδίου (Peptide-Spectrum-Matches, PSMs) από τα .xml αρχεία του Mascot. Με τον τρόπο αυτό δημιουργήθηκε ένα αρχείο-οδηγός που υποδεικνύει σε ποιό φάσμα πρέπει να επαληθευτεί ποιό πεπτίδιο.
- Με το πρόγραμμα SeeMS της σουίτας ProteoWizard άνοιξαν τα τελικά αρχεία .mzXML και έγινε δυνατή η οπτικοποίηση των PSMs, με τη χρήση των παραμέτρων που φαίνονται στην Εικόνα 2.5.2. Αποθηκεύτηκαν εικόνες από κάθε αντιστοιχία φάσματος-πεπτιδίου, για πιθανή αξιολόγηση στο μέλλον. Οι κανόνες (De Novo PSM Rules, n.d.) που ακολουθήθηκαν για την επαλήθευση είναι οι εξής:

1. Το φάσμα πρέπει να έχει καλό λόγο σήματος/θρούβου και οι πιο προεξέχουσες κορυφές να έχουν αντιστοιχηθεί.
2. Θραύσματα b ή y που περιέχουν τα αμινοξέα R, K, Q, N τείνουν να χάνουν αμμωνία, οπότε να εμφανίζουν μία συμπληρωματική κορυφή -17amu, συνήθως μικρότερης έντασης, αριστερά της κύριας.
3. Ομοίως, θραύσματα b ή y που περιέχουν S, T, E μπορεί να χάσουν νερό και να εμφανίσουν μία συμπληρωματική κορυφή -18amu αριστερά της κύριας.
4. Η ένταση των b ιόντων είναι δυνατόν να χάνεται όταν ακολουθεί P, G, H, K ή R.
5. Αν στο N-τελικό άκρο ενός θραύσματος βρίσκεται P ή H, ενδέχεται να προκύψουν συμπληρωματικές εσωτερικές θραύσεις. Παραδείγματος χάριν, αν σε ένα πεπτίδιο VKGDCPRIVKL υπάρχει το θραύσμα PRIVKL, ενδέχεται να υπάρχουν και κορυφές που να αντιστοιχούν στα PRIVK, PRIV και PRI. Συνήθως το y θραύσμα που έχει P ή H στο N-τελικό του άκρο αντιστοιχεί στην πιο υψηλή κορυφή του φάσματος.
6. Παρουσία P, και λιγότερο, H, K, R ενδέχεται το y ιόν να ανταλλάσσει ένταση με το b, και αντίστροφα.
7. Όταν μία θραύση γίνεται πριν/μετά από R, η κορυφή της απώλειας της αμμωνίας ίσως είναι υψηλότερη από την κύρια κορυφή του θραύσματος.
8. Παρουσία D ενδέχεται το σήμα των θραυσμάτων να σβήνει.
9. Κάποιες φορές εμφανίζονται στην αρχή του φάσματος ιόντα immonium, που είναι ενδεικτικά της παρουσίας κάποιου μεμονωμένου αμινοξέος στο πεπτίδιο. Η απουσία τους όμως δεν υποδηλώνει ότι το εν λόγω αμινοξύ δεν υπάρχει στην αλυσίδα.
10. Όταν στο πεπτίδιο υπάρχει R και D ή E,

συνήθως η υψηλότερη κορυφή είναι το θραύσμα b ή y του D ή E που περιέχει R (Höhlitz, et al., 2012).



Εικόνα 2.5.1 Παράμετροι του προγράμματος MSConvert



Εικόνα 2.5.2 Παράμετροι του προγράμματος SeeMS

Η χειροκίνητη επαλήθευση και όλη η μετέπειτα ανάλυση πραγματοποιήθηκε στις Αναζητήσεις Mascot στις βάσεις δεδομένων (i) του μεταγραφώματος του *Bacitracina oleae* και (ii) του συνόλου των καταχωρημένων πρωτεϊνών του γένους *Drosophila* της Swissprot.

Με το ελεύθερο λογισμικό jVenn (Bardou, Mariette, Escudié, Djemiel, & Klopp, 2014), δημιουργήθηκαν σχήματα Venn που περιγράφουν τον αριθμό των PSMs που θεωρήθηκαν έγκυρα, καθώς και την αλληλεπικάλυψη σε πεπτίδια μεταξύ των δύο συνθηκών.

## 2.6 Έλεγχος ποιότητας

Η ποιότητα των υπό ανάλυση δειγμάτων ελέγχθηκε με δύο τρόπους.

Αρχικά χρησιμοποιήθηκε το στατιστικό SQS (Spectrum Quality Score), από το λογισμικό PeptideNovo+ ως ένας δείκτης της ποιότητας του κάθε φάσματος (Nesvizhskii, et al., 2006), (PeptideNovo+ Readme File), (Frank & Pevzner, 2005). Το SQS είναι ένας αλγόριθμος μηχανικής μάθησης που εκπαιδεύτηκε χρησιμοποιώντας φάσματα καλής και κακής ποιότητας και λαμβάνει υπ' όψιν βασικά χαρακτηριστικά στο φάσμα – τον αριθμό των υψηλών

κορυφών, την κατανομή τους σε όλο το εύρος του φάσματος, την παρουσία συμπληρωματικών b και γ ιόντων, το κατά πόσο η απόσταση μίας υψηλής κορυφής από την επόμενη ισούται με το μοριακό βάρος κάποιου αμινοξέος κοκ. Παίρνει τιμές από 0 μέχρι 1 και αντιπροσωπεύει την πιθανότητα το εκάστοτε φάσμα να αποτελεί ένα ικανά ποιοτικό φάσμα για να αναλυθεί σε πεπτίδιο. Το SQS υποβοηθήσε τη χειροκίνητη επαλήθευση, αλλά χρησιμοποιήθηκε και συγκριτικά μεταξύ του ασύζευκτου και του συζευγμένου δείγματος: Κατασκευάστηκαν, με τη στατιστική γλώσσα προγραμματισμού R, γραφήματα boxplot (θηκογράμματα) που παρέχουν μία σαφή εικόνα για τη μέση τιμή, αλλά και τη διασπορά της ποιότητας των φασμάτων στις δύο συνθήκες.

Μετά τη χειροκίνητη επαλήθευση, κατασκευάστηκαν σε R γραφήματα των, αποδεκτών και μη, PSMs ως προς την τιμή του Mascot Score, τα οποία επίσης συνδέονται με την ποιότητα των φασμάτων, αναδεικνύουν δε και τη σημασία της χειροκίνητης επαλήθευσης.

### 2.7 Ανάλυση Gene Ontology – REVIGO

Έγινε συλλογή των GO terms των έγκυρων πεπτιδίων από τη UNIPROT: για τα αποτελέσματα της αναζήτησης με το γένος *Drosophila*, χρησιμοποιήθηκαν τα UNIPROT ID των αποτελεσμάτων απ'ευθείας, ενώ για την αναζήτηση με το μεταγράφημα του *Bactrocera oleae*, συλλέχθηκαν τα UNIPROT ID των BLAST hits των σχετικών μεταγράφων.

Χρησιμοποιήθηκε το πρόγραμμα REVIGO (Šupek, Bošnjak, Škunca, & Šmuc, 2011) για την ανάλυση και καλύτερη παρουσίασή τους. Το REVIGO επιτρέπει την εισαγωγή p-value, ή άλλων τιμών που να αποδίδουν στατιστική σημασία, ανά GO term, επόμενα επιλέχθηκε η παράμετρος "some other quantify where higher is better" και τα GO terms αναλύθηκαν με το ανάλογο για το καθένα Mascot Score, που εκφράζει τη στατιστική δύναμη της εμφάνισής τους.

### 2.8 Σύγκριση με την τρανσκριπτομική ανάλυση

Οι λίστες που περιγράφουν τη διαφορική έκφραση

όλων των μεταγράφων του *B. oleae* στους ιστούς MAGs (Male Accessory Glands), testes, FAGs (Female Accessory Glands), προσυζευκτικά και μετασυζευκτικά, καθώς και στον ιστό ελέγχου GUT, φιλτραρίστηκαν ώστε να περιέχουν μόνο τα μεταγγραφα που αντιστοιχούν σε πεπτίδια. Σχεδιάστηκαν ραβδογράμματα της έκφρασης κάθε μεταγράφου, ως προς τον ιστό ελέγχου ( $p < 0.005$ ). Πραγματοποιήθηκε αναζήτηση για ORFs στα μεταγγραφα αυτά και ανάλυση BLASTP για τις υποθετικές πεπτιδικές αλληλουχίες των ORF.

### 2.9 Ανάλυση TBLASTN και BLASTN με μετάγραφο που κωδικοποιούν για ACPs (Accessory Gland Proteins) άλλων οργανισμών

Συλλέχθηκαν οι αλληλουχίες από τα μεταγγραφα από ACPs που συγκεντρώνονται στις δημοσιεύσεις των (Charman, 2004), (Davies, 2006) και (Baldini, Gabrieli, Rogers, & Catteruccia, 2012) στα είδη *Drosophila melanogaster*, *Ceratitis capitata* και *Anopheles gambiae* και δημιουργήθηκε η βάση δεδομένων `Acp` με το πρόγραμμα makeblastdb και την εντολή:

```
>makeblastdb -in acp.fa -dbtype nucl
```

Έγινε tBLASTn στη βάση δεδομένων `Acp` με τα έγκυρα πεπτίδια, τροποποιώντας τις παραμέτρους ειδικά για εύρεση μικρών αλληλουχιών. Εντολή:

```
>tblastn -query valid.fa -db Acp -out results.csv -outfmt '6 qseqid sseqid pident length qlen slen evalue' -evalue 20000 -seg no -comp_based_stats 0 -word_size 2 -ungapped -num_threads 4
```

Έγινε BLASTn με query όσα μεταγγραφα του δάκου αντιστοιχήθηκαν σε πεπτίδια και βάση δεδομένων την Acp, κρατώντας τις τυπικές επιλογές. Εντολή:

```
>blastn -query valid_transcripts.fa -db Acp -out resultsBlast.csv -outfmt '6 qseqid sseqid pident length qlen slen evalue' -num_threads 4
```

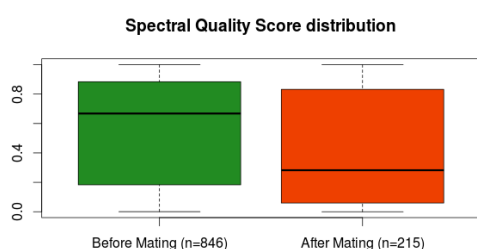
Έγινε tBLASTx με ίδια χαρακτηριστικά με το BLASTn, στο site του BLAST για καλύτερη οπτικοποίηση του αποτελέσματος.

### 3. Αποτελέσματα και συζήτηση

### 3.1 Έλεγχος ποιότητας και Χειροκίνητη επαλήθευση (manual validation)

Στον Πίνακα 1 του Παραρτήματος συγκεντρώνονται ανά συνθήκη – συζευγμένα/ασύζευκτα – και ανά βάση δεδομένων – τρανσκριπτομική βιβλιοθήκη/πρωτεϊνική βιβλιοθήκη – όλα τα πεπτίδια που θεωρήθηκαν έγκυρα κάνοντας επαλήθευση με το λογισμικό SeeMS της σουίτας ProteoWizard. Οι κανόνες που ακολουθήθηκαν περιγράφονται στο κεφάλαιο Υλικά και Μέθοδοι.

Αξίζει να σημειωθεί ότι είναι άγνωστο το αν τα απομονωμένα πεπτίδια αποτελούν τμήματα μεγαλύτερων πρωτεϊνών υπό αποικοδόμηση, ή είναι στην πραγματικότητα πρωτεολυμένα βιοενεργά πεπτίδια, ή προϊόντα μετάφρασης μικρών αναγνωστικών πλαισίων (sORFs). Τα μετάγραφα και οι πρωτεΐνες που αντιστοιχίζονται πολλαπλά με φάσματα είναι πιο πιθανό να τοποθετούνται υπό την πρώτη εκδοχή, ενώ τα μετάγραφα/πρωτεΐνες όπου υπάρχει αντιστοιχία ένα προς ένα, ίσως υπόκεινται στη δεύτερη.

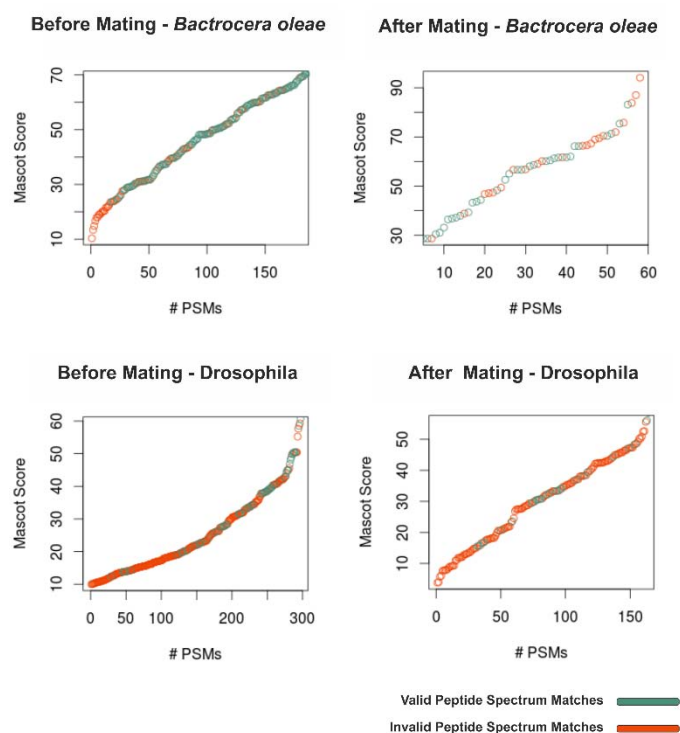


Εικόνα 2.1.1 Θηκόγραμμα κατανομής των τιμών SQS ανά δείγμα

Παρατηρείται επίσης ότι τα δύο δείγματα διαφέρουν αριθμητικά κατά πολύ ως προς τα επαληθευμένα πεπτίδια: προσυζευκτικά έχουν επαληθευτεί 160 και 57 PSMs στην τρανσκριπτομική και την πρωτεϊνική βάση αντίστοιχα, ενώ μετασυζευκτικά 38 και 30. Αυτό δεν είναι λογικό να αποτελεί κάποια βιολογική πληροφορία, αντίθετα, αφορά στην ποιότητα των δύο δειγμάτων. Όπως φαίνεται και στο θηκόγραμμα της Εικόνας 3.1.1:

1. Η κατανομή των SQS στα δύο δείγματα διαφέρει. Στο ασύζευκτο δείγμα η μέση τιμή είναι υψηλότερη και η διασπορά των τιμών είναι μικρότερη.
2. Το πλήθος των φασμάτων που προέκυψαν από τη φασματοσκοπία των μη συζευγμένων ιστών είναι τετραπλάσιο από αυτό των συζευγμένων.

Είναι αναμενόμενο λοιπόν το γεγονός ότι τα έγκυρα πεπτίδια στο δείγμα των μη συζευγμένων βοηθητικών αδένων του θηλυκού θα είναι πιο πολλά και αποδίδεται μάλλον σε κάποιο τεχνικό ζήτημα στη διαδικασία της απομόνωσής τους, παρά σε κάποιο βιολογικό αίτιο.



Εικόνα 3.1.2 Γραφήματα των PSMs ως προς το Mascot Score

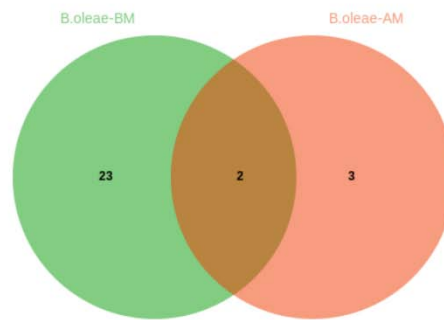
Στην Εικόνα 3.1.2 παρουσιάζονται σε γραφήματα όλα τα PSMs, ως προς το Mascot Score. Με μπλε χρώμα εμφανίζονται τα έγκυρα PSMs, ενώ με πορτοκαλί τα μη αποδεκτά. Τα γραφήματα αυτά αποτελούν ένα ακόμα στοιχείο που υποδηλώνει ότι η ποιότητα των απομονωμένων πεπτιδίων των μη συζευγμένων ιστών είναι υψηλότερη. Ακόμα, παρατηρείται πως είναι μεγαλύτερο το ποσοστό των έγκυρων πεπτιδίων της αναζήτησης Mascot με βάση δεδομένων το μεταγράψωμα του δάκου, παρά με το σύνολο των πρωτεϊνών του γένους *Drosophila*. Αυτό είναι φυσιολογικό, καθώς η βιβλιοθήκη του SWISSPROT απαρτίζεται από πρωτεΐνες που απέχουν εξελικτικά από τα απομονωμένα πεπτίδια, αντίθετα από τα μετάγραφα. Μάλιστα, όπως αναφέρεται και στο κεφάλαιο Εισαγωγή, το αναπαραγωγικό σύστημα αποτελεί ένα χώρο όπου η εξέλιξη μπορεί να λειτουργήσει ραγδαία και δραστικά, υπό την έννοια ότι μεταλλάξεις σε μόριά του είναι ικανές να γεννήσουν ασυμβατότητες και περιορισμούς μέσα στο είδος, που πιθανόν να σημάνουν την αρχή μίας ειδογένεσης: ίσως πρέπει να αναμένεται ότι οι μοριακές συγκρίσεις μεταξύ κοντινών ειδών θα αποφέρουν λίγο λιγότερους καρπούς στο αναπαραγωγικό σύστημα, από ό,τι σε άλλους ιστούς.

Συμπληρωματικά, σχεδιάστηκαν δύο διαγράμματα Venn (Εικόνα 3.1.3) που δείχνουν την κατανομή των κωδικών μεταγράφων/πρωτεϊνών στις δύο αναζητήσεις.





**Common UNIPROT IDs:**  
 OR35A\_DROME  
 THR\_DROVI  
 SPNE\_DROAN  
 MCM5\_DROME



**Common transcript IDs:**  
 c12447\_g1\_i1  
 c57633\_g2\_i2

Εικόνα 3.1.3 Διαγράμματα Venn που παρουσιάζουν τους κωδικούς που εμφανίστηκαν και στις δύο συνθήκες υπό μελέτη

### 3.2 Ανάλυση όρων Gene Ontology

Στην Εικόνα 3.2.1 συνοψίζονται οι κυριότεροι όροι GO του κλάδου Βιολογική Διεργασία (BP, Biological Process) που εκπροσωπούν τα πεπτιδία προσυζευκτικά και μετασυζευκτικά. Στο άνω μέρος των σχεδίων συγκεντρώνονται οι όροι με τις υψηλότερες τιμές Mascot Score. Παρατηρούνται όροι γενικής λειτουργίας που σχετίζονται με τον αναδιπλασιασμό του DNA, τη ρύθμιση της γονιδιακής έκφρασης, τη μετάφραση και τη ρύθμιση του μεταβολισμού, αλλά και όροι που σχετίζονται με την αναπαραγωγή και την εμβρυική ανάπτυξη (ωογένεση, ανοσοαπόκριση, επιμήκυνση βλαστικής ζώνης, κορυφαία στένωση που εμπλέκεται στη γαστριδίωση). Αποτελεί ευχάριστη επιβεβαίωση η εμφάνιση του οσφρητικού υποδοχέα OR35a, καθώς σε προηγούμενα πειράματα (Saghi, et al., 2014) είχε επιβεβαιωθεί η παρουσία άλλων οσφρητικών μορίων σε αναπαραγωγικούς ιστούς του *Bactrocera oleae*.

Στους Πίνακες 2 και 3 του Παραρτήματος είναι καταχωρημένοι οι όροι και των τριών κλάδων του Gene Ontology ανά πεπτιδίδια, προσυζευκτικά και μετασυζευκτικά, και έχουν σημειωθεί με bold όροι σχετικοί με την αναπαραγωγή και την εμβρυική ανάπτυξη.

### 3.3 Σύγκριση πεπτιδικής ανάλυσης με τρανσκριπτομικά δεδομένα

Τα πεπτιδομικά δεδομένα είναι ποιοτικά· δεν παρέχεται κάποια εικόνα για την ποσότητα του κάθε πεπτιδίου που ταυτοποιείται. Για το λόγο αυτό, πραγματοποιήθηκε αναζήτηση σε πίνακες από προηγούμενη τρανσκριπτομική ανάλυση που περιγράφουν τη διαφορική έκφραση των μεταγράφων στους ιστούς MAGs, testes, FAGs και με τον ιστό ελέγχου GUT για όσα μεταγράφα βρέθηκαν να αντιστοιχίζονται με πεπτιδία, ώστε να δανειστούμε ποσοτική πληροφορία, τουλάχιστον για τη μεταγραφική δραστηριότητα του καθενός. Στον Πίνακα 4 του Παραρτήματος και στις Εικόνες 3.3.1 και 3.3.2 παρουσιάζεται η διαφορική έκφραση των μεταγράφων αυτών στους αναπαραγωγικούς ιστούς ως προς τον ιστό ελέγχου και σημειώνονται – γκρι υπογράμμιση και αστερίσκος αντίστοιχα – τα  $\log_2FC$  που έχουν  $p$ -value < 0.005. Επίσης, χρησιμοποιήθηκε το πρόγραμμα ORF Finder ([http://www.bioinformatics.org/sms2/orf\\_find.html](http://www.bioinformatics.org/sms2/orf_find.html)) για να βρεθούν ανοικτά αναγνωστικά πλαίσια στα μεταγράφα που αντιστοιχήθηκαν με πεπτιδία από την

φασματοσκοπική ανάλυση. Εξετάστηκε κατά πόσο οι μεταφράσεις των ORFs περιέχουν τα πεπτιδία που βρέθηκαν φασματοσκοπικά και πραγματοποιήθηκε χαρακτηρισμός με το πρόγραμμα BLASTP για τα πεπτιδία στα οποία μεταφράζονται τα μεγαλύτερα ORFs από κάθε μετάγραφο. Στον Πίνακα 5 του Παραρτήματος παρουσιάζονται τα αποτελέσματα της ανάλυσης αυτής, πρώτα για τα μεταγράφα που εμφανίζουν μεταγραφική ενεργότητα, ακολούθως για το μετάγραφο που είναι μεταγραφικά ενεργό μόνο στον ιστό Testes και τέλος για μεταγράφα που δε φαίνεται από την τρανσκριπτομική ανάλυση να είναι μεταγραφικά ενεργά.

#### 3.3.1 Μετάγραφή στους ιστούς MAGs - FAGs

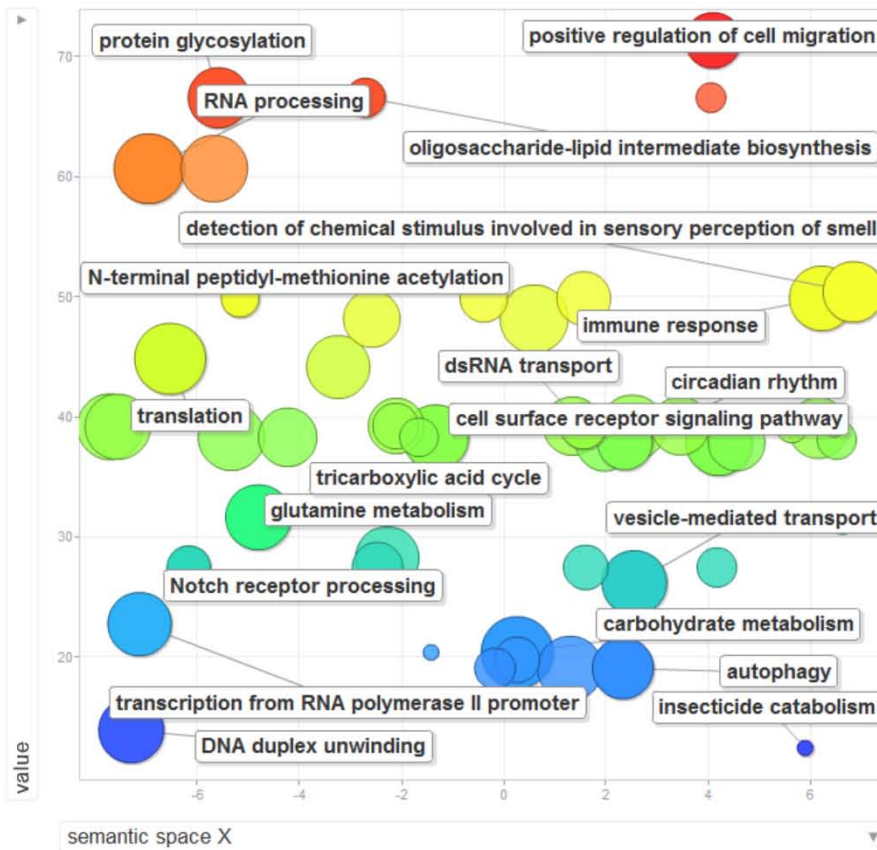
##### c12447\_g1\_i1 – Κερατοξίνη A

Βρέθηκε στο αναγνωστικό πλαίσιο 1 του μεταγράφου αυτού ένα ORF μήκους 192 νουκλεοτιδίων. Μεταφράζεται σε ένα πεπτιδίδια 64 αμινοξέων που με BLASTP φαίνεται να φέρει ομολογία με την Κερατοξίνη A της *Ceratitis capitata*. Οι κερατοξίνες A και B είναι αντιμικροβιακά μόρια μεγέθους 71 αμινοξέων και μοριακού βάρους 3kDa που έχουν απομονωθεί από τους βοηθητικούς αδένες της θηλυκής μεσογειακής μύγας. Τα περισσότερα αντιμικροβιακά πεπτιδία που έχουν βρεθεί στην αιμοδόμο των εντόμων επάγονται από μικροβιακές μολύνσεις. Αντιθέτως, η έκφραση των κερατοξινών είναι ιδιόσυστα ενεργή αποκλειστικά στους αναπαραγωγικούς αδένες του σεξουαλικά ώριμου θηλυκού, ενώ επάγεται ελάχιστα μετά τη σύζευξη (Marchini, et al., 1995). Η κερατοξίνη A είναι παρούσα και στην επιφάνεια όπου αποτίθενται τα αυγά της μεσογειακής μύγας, επομένως έχει προταθεί ότι παίζει ρόλο στην προστασία από μικροβιακές μολύνσεις κατά τα πρώιμα στάδια της ανάπτυξης της προνύμφης (Gomulski, et al., 2012). Η επαγωγή της έκφρασης αντιμικροβιακών μορίων ως απόκριση στη σύζευξη είναι αδύναμη στη μεσογειακή μύγα, συγκριτικά με τη *Drosophila melanogaster*, και ομοιάζει με τα πρότυπα που παρατηρούνται στο *Anopheles gambiae* και την *Apis mellifera*, που χαρακτηρίζονται από απουσία μετασυζευκτικής ανοσοαπόκρισης (Kocher, Richard, Taryp, & Grozinger, 2008), (Rogers, et al., 2008).

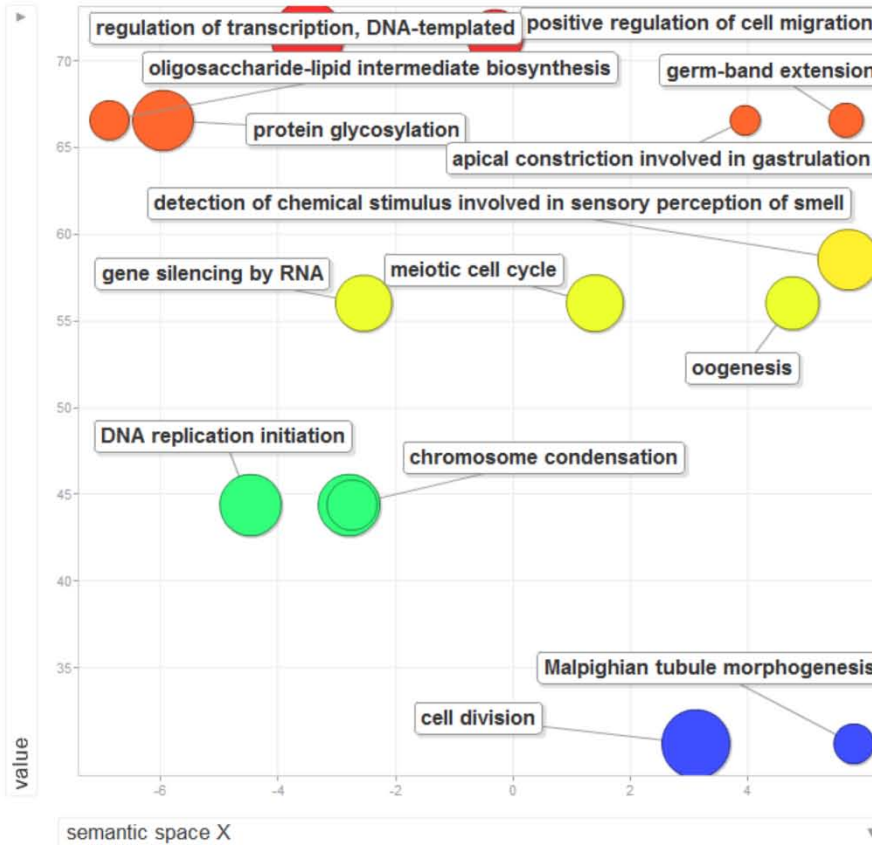
Το παρατηρούμενο πρότυπο έκφρασης στο δάκο της ελιάς δε φαίνεται να απαντάται σε κάποια από τις δύο παραπάνω κατηγορίες: η έκφραση παραμένει χαμηλή και σχετικά αμετάβλητη από τη σύζευξη, στο θηλυκό, ενώ υψίσταται και στο αρσενικό, στο οποίο παρατηρείται πτώση μετασυζευκτικά, επομένως δεν πρόκειται για θηλεο-ειδικό γονίδιο.



### Before Mating distribution



### After Mating distribution



Εικόνα 3.2.1 Παρουσίαση όρων Gene Ontology προσζευκτικά και μετασζευκτικά

#### **c97331\_g1\_i1 – Πρωτεΐνη εξωτερικής μεμβράνης A**

Βρέθηκε στο αναγνωστικό πλαίσιο 3 ένα ORF μήκους 252 νουκλεοτιδίων. Μεταφράζεται σε ένα πεπτιδίο 84 αμινοξέων που με BLASTP φαίνεται να φέρει ομολογία με την πρωτεΐνη ompA του θαλάσσιου α-πρωτεοβακτηρίου *Maritimibacter alkaliphilus*. Πρόκειται για ένα καλά συντηρημένο μέλος της μεγαλύτερης υπερ-οικογένειας προκαρυωτικών εξωμεμβρανικών πρωτεϊνών, με πιθανό ρόλο στη σταθερότητα της επαφής της εξωτερικής μεμβράνης με το στρώμα πεπτιδογλυκάνης και στη βακτηριακή κυτταρική προσκόλληση (Aykanat & Dincturk, 2007). Το γεγονός ότι το συγκεκριμένο βακτήριο απαντάται σε υδάτινο και συγκεκριμένα θαλάσσιο περιβάλλον (Lee, Choo, Γιωναννοή, & Cho, 2007), υποδεικνύει ότι δεν αποτελεί μέρος της μικροχλωρίδας του εντόμου, αλλά δεν αποκλείεται η πιθανότητα η συγκεκριμένη πρωτεΐνη να είναι στην πραγματικότητα μέλος του μετα-πρωτεόματος του δάκου.

Το ίδιο πρότυπο έκφρασης με το παραπάνω μετάγραφο – c12447 – παρατηρείται και σε αυτήν την περίπτωση. Το γεγονός ότι υπάρχει μεγαλύτερη έκφραση του μεταγράφου αυτού στο αρσενικό, σε συνδυασμό με την ομολογία του με μόρια σχετικά με την κυτταρική προσκόλληση, υποδεικνύουν πως αξίζει να μελετηθεί περαιτέρω, καθώς δε θα ήταν απίθανη η εμπλοκή του σε ουσιαστικές διεργασίες της αναπαραγωγής, όπως η ακροσωμική αντίδραση, η αποθήκευση του σπέρματος, ή η διατήρηση των φυσικοχημικών του ιδιοτήτων.

#### **c19149\_g1\_i1 – Ριβοσωμική πρωτεΐνη RS3A**

Στο αναγνωστικό πλαίσιο 1 του μεταγράφου βρέθηκε ένα ORF μήκους 807 νουκλεοτιδίων που μεταφράζεται σε ένα πεπτιδίο 269 αμινοξέων που φέρει τέλεια ομολογία με τη ριβοσωμική πρωτεΐνη RS3A της *Drosophila yakuba* και 40S ριβοσωμικές πρωτεΐνες πολλών ειδών. Η RS3A είναι συστατικό της 40S (μικρής) ριβοσωμικής υπομονάδας, με πληθώρα εναλλακτικών μεταγράφων που οφείλονται σε διαφορετικά σημεία εκκίνησης της μεταγραφής (Transcription Start Sites – TSS) και ψευδογονιδίων.

#### **c49413\_g1\_i1 – Β-αλυσίδα της καρβοξυλάσης του μεθυλ-κροτονοϋλ-CoA**

Βρέθηκε στο αναγνωστικό πλαίσιο 1 του μεταγράφου ένα ORF μήκους 1749 νουκλεοτιδίων. Μεταφράζεται σε ένα πεπτιδίο 583 αμινοξέων που είναι ομόλογο με τη Β-αλυσίδα της καρβοξυλάσης του μεθυλ-κροτονοϋλ-CoA του *Bactrocera oleae* και άλλων Τερηρίτιδαι. Το ένζυμο αυτό εμπλέκεται σε βασικά μονοπάτια βιοσύνθεσης και καταβολισμού αμινοξέων και λιπαρών οξέων. Σε κανένα από τα ORFs που βρέθηκαν δεν υπάρχει η αλληλουχία του μικρού πεπτιδίου που ταυτοποιήθηκε στην πεπτιδομική ανάλυση.

#### **c51247\_g1\_i1 – Περιοχή serendipity H-1**

Βρέθηκαν πολλά ORFs διαφόρων μεγεθών. Το μεγαλύτερο από αυτά (3099nt) βρίσκεται στο αναγνωστικό πλαίσιο 3 και κωδικοποιεί για ένα μεγάλο πεπτιδίο μήκους 1033 αμινοξέων, πλούσιο σε περιοχές πρόσδεσης ψευδαργύρου και DNA. Ταυτοποιείται απόλυτα με την ομολογία με το serendipity locus protein H-1 (sry H-1) του δάκου της ελιάς. Το sry H-1 ανήκει σε μία οικογένεια γονιδίων που είναι μεταγραφικά ενεργά κατά την πρώιμη ανάπτυξη, ενώ πιθανώς να παρέχονται και μητρικά. Το γεγονός ότι διακατέχονται από DNA-binding sites σε συνδυασμό με το μεταγραφικό τους πρότυπο υποδεικνύει ότι μπορεί να έχουν κάποιο ρόλο στον πρώιμο προγραμματισμό της έκφρασης του εμβρύου (Vincent, Keizlarová-Lepesant, Segalat, Yanicostas, & Lepesant, 1988). Σε κανένα από τα ORFs που βρέθηκαν δεν υπάρχει η αλληλουχία του μικρού πεπτιδίου που ταυτοποιήθηκε στην πεπτιδομική ανάλυση.

#### **c56147\_g1\_i1 – Υπομονάδα του συμπλέγματος επιμήκυνσης**

Βρέθηκαν πολλά ORFs διαφόρων μεγεθών στο μετάγραφο αυτό. Η αλληλουχία του μικρού πεπτιδίου που ταυτοποιήθηκε στην πεπτιδομική ανάλυση βρίσκεται σε ένα πολύ μικρό ORF το οποίο δε δίνει BLAST hits. Το μεγαλύτερο ORF που βρέθηκε (804 νουκλεοτιδία) βρίσκεται στο αναγνωστικό πλαίσιο 3 και κωδικοποιεί για ένα πεπτιδίο μήκους 268 αμινοξέων, ομόλογο του ELP6 που αποτελεί υπομονάδα του συμπλέγματος επιμήκυνσης της RNA πολυμεράσης II και εμπλέκεται στη μεταγραφή της μεσογειακής μύγας.

#### **c56988\_g2\_i1 – Πρωτεΐνη 1 σχετική με την ουβικιτίνη**

Το μεγαλύτερο ORF του μεταγράφου αυτού βρίσκεται στην αρχή του, στο αναγνωστικό πλαίσιο 1, με μήκος 1761 νουκλεοτιδία, και μεταφράζεται σε ένα πεπτιδίο 587 αμινοξέων που φέρει ομολογία με την πρωτεΐνη UBAP1 (ubiquitin associated protein 1) του *Bactrocera dorsalis*. Το πεπτιδίο που ταυτοποιήθηκε με το Mascot δεν εμφανίζεται σε κανένα από τα ORFs του ORF Finder.

#### **c57771\_g2\_i1 – Μη χαρακτηρισμένη πρωτεΐνη KIA0100**

Αυτό το μετάγραφο μήκους 7436 νουκλεοτιδίων έχει ένα μεγάλο ORF που μεταφράζεται σε ένα πεπτιδίο 2287 αμινοξέων με ομολογία με την άγνωστης λειτουργίας πρωτεΐνη KIA0100 του *Bactrocera oleae*. Στο C-τελικό άκρο του βρίσκεται η συντηρημένη περιοχή Art1 που περιέχει ένα σινιάλο εντοπισμού στο σύμπλεγμα Golgi. Το πεπτιδίο από τη φασματοσκοπική ανάλυση δεν εμφανίζεται σε κανένα από τα ORFs.

#### **c57872\_g1\_i1 – Γλυκοσυλτρανφεράση της πυροφωσφορικής δολιχόλης**

Το μετάγραφο c57872 στο αναγνωστικό πλαίσιο 3 έχει ένα μεγάλο ORF μήκους 1599nt που περιέχει το πεπτιδίο που ταυτοποιήθηκε από την πεπτιδομική ανάλυση. Κωδικοποιεί για ένα πεπτιδίο μήκους 533 αμινοξέων που φέρει τέλεια ομολογία με την α-1,3-γλυκοσυλτρανφεράση του δολιχολικού πυροφωσφορικού του *Bactrocera oleae*. Το ένζυμο αυτό καταλύει ένα ενδιάμεσο βήμα στη συγκρότηση του ολιγοσακχαρίτη Glc3Man9GlcNAc2 πάνω στην πυροφωσφορική δολιχόλη. Η Glc3Man9GlcNAc2-PP-Dol είναι ο δότης γλυκάνης κατά τη μεταμεταφραστική τροποποίηση της N-συνδεδεμένης γλυκοζυλίωσης (Burda & Markus, 1999).

#### **c58251\_g1\_i1 – Παράγοντας επιμήκυνσης 1-α**

Το πεπτιδίο που ταυτοποιήθηκε από την πεπτιδομική ανάλυση βρίσκεται στο C-τελικό άκρο του προϊόντος του μεγαλύτερου σε μέγεθος ORF, μήκους 464 αμινοξέων, και εμφανίζει ομολογία με τον παράγοντα επιμήκυνσης 1-α του *Bactrocera oleae* και πολλών άλλων εντόμων. Ο παράγοντας επιμήκυνσης προωθεί την GTP-εξαρτώμενη πρόσδεση του φορτωμένου tRNA στη θέση A του ριβοσώματος κατά τη μετάφραση, γεγονός που δικαιολογεί και τον εξαιρετικό βαθμό συντήρησης που παρατηρείται με το BLASTP.

Τα παραπάνω 8 μετάγραφα (c19149 – c58251) εμφανίζουν παρόμοιο μοτίβο έκφρασης. Βρίσκονται σε υψηλή μεταγραφική κατάσταση στο θηλυκό προσυζευκτικά, ενώ μετασυζευκτικά υπάρχει πτώση προς τα επίπεδα του ιστού ελέγχου και του αρσενικού. Όπως αναφέρεται στην εισαγωγή, θα ήταν αναμενόμενο να εμφανίζεται κορύφωση της έκφρασης μετασυζευκτικά για γονίδια που σχετίζονται με την ενεργοποίηση του μεταβολισμού, της μετάφρασης καθώς και για γονίδια σχετικά με τον εμβρυικό προγραμματισμό. Ενδεχομένως αυτή η πληροφορία χάνεται διότι ο ιστός έχει απομονωθεί 12 ώρες μετά τη σύζευξη και η

αύξηση στην έκφραση μπορεί να υπάρχει σε μετέπειτα χρονική στιγμή.

#### **c42884\_g1\_i1 – Μιτοχονδριακή ριβοσωμική πρωτεΐνη L33**

Το μεγαλύτερο ORF που παρατηρείται στο μετάγραφο αυτό έχει μήκος 65 αμινοξέων και εμφανίζει τέλεια ομολογία με την 39S μιτοχονδριακή ριβοσωμική πρωτεΐνη L33 του *Bactrocera cucurbitae* και άλλων ειδών. Η μελέτη των ριβοσωμικών υπομονάδων παρουσιάζει ιδιαίτερο ενδιαφέρον, γιατί αρκετές από αυτές κωδικοποιούνται από απλοανεπαρκή γονίδια. Η απλοανεπαρκεία είναι το φαινόμενο κατά το οποίο για να παρατηρηθεί ο φαινότυπος αγρίου τύπου είναι απαραίτητο να υφίστανται σε λειτουργική μορφή και τα δύο αντίγραφα – σε ένα διπλοειδή οργανισμό – κάποιοι γονιδίου. Στη *Drosophila melanogaster* έχει παρατηρηθεί πως αρκετά απλοανεπαρκή γονίδια που κωδικοποιούν για ριβοσωμικές πρωτεΐνες αποδίδουν το φαινότυπο *Minute*, που χαρακτηρίζεται από επιμήκυνση της εμβρυϊκής αναπτυξιακής διαδικασίας, μειωμένη γονιμότητα και βιωσιμότητα (Marygold, 2007). Μέχρι στιγμής δεν έχουν ταυτοποιηθεί μιτοχονδριακά ριβοσωμικά *Minute* γονίδια, όμως το φαινόμενο εμφανίζει ιδιαίτερο ενδιαφέρον καθώς πιθανότατα θα μπορούσε να αποτελέσει εργαλείο τεχνικών πληθυσμιακού ελέγχου εντόμων.

Τα δύο πεπτίδια που ταυτοποιήθηκαν από τη φασματοσκοπική ανάλυση δεν αποτελούν τμήματα των πρωτεϊνών που κωδικοποιούνται από τα ORFs.

#### **c51099\_g2\_i2 – Ικαραπίνη**

Το μεγαλύτερο ORF παρατηρείται στο αναγνωστικό πλαίσιο 3 και το φασματοσκοπικά ταυτοποιημένο πεπτίδιο NIDDPVYK βρίσκεται στο C-τελικό άκρο του. Το ORF κωδικοποιεί για ένα πεπτίδιο 210 αμινοξέων με ομολογία με μια μη-χαρακτηρισμένη πρωτεΐνη του *Bactrocera dorsalis* (XP\_011199953.1) και με τμήμα της icararin-like του *Bactrocera oleae*. Η ικαραπίνη είναι μία αλλεργιογόνος εκκρινόμενη πρωτεΐνη 223 αμινοξέων που βρίσκεται σε χαμηλές συγκεντρώσεις στο δηλητήριο της μέλισσας *Apis mellifera carnica* (Peiren, et al., 2006).

#### **c57492\_g2\_i1 – Άγνωστη πρωτεΐνη DDB\_G0288805**

Το φασματοσκοπικά ταυτοποιημένο πεπτίδιο δε βρίσκεται σε κανένα από τα ORF που εμφανίζονται στην αλληλουχία του μεταγράφου αυτού. Το μεγαλύτερο εξ αυτών βρίσκεται στο δεύτερο αναγνωστικό πλαίσιο και μεταφράζεται σε ένα πεπτίδιο 448 αμινοξέων που αντιστοιχεί στη μη χαρακτηρισμένη πρωτεΐνη DDB\_G0288805 του δάκου (ref|XP\_014092398.1|).

Τα μετάγραφα (c42884 – c57492) εμφανίζουν χαμηλότερη έκφραση στους αναπαραγωγικούς ιστούς από ό,τι στον ιστό ελέγχου και παρατηρείται περαιτέρω πτώση τους μετασυσζευκτικά στο θηλυκό.

#### **c51296\_g1\_i1 – Μυο-ειδική πρωτεΐνη 20**

Το φασματοσκοπικά ταυτοποιημένο πεπτίδιο AGSNKGATQAGQNIQASRKGK βρίσκεται στο C-τελικό άκρο του μεγαλύτερου ORF που εμφανίζεται στο αναγνωστικό πλαίσιο 3 του μεταγράφου αυτού και κωδικοποιεί για μία πρωτεΐνη 185 αμινοξέων, με τέλεια ομολογία με τη Muscle-specific protein 20 που σχετίζεται με τη συγκόλληση των μυοβλαστών και την ένωσή τους σε ενιαίες συγκιαιτικές ίνες που ονομάζονται μυοσωλήνες. Έχουν καταγραφεί γονίδια σχετικά με τη μυϊκή σύσπαση (*Mrp20*, *sls*, *Tm2*, *CaMKII*, *Mbs*, *bt*, *GsfS1*, *TrnC73F*, *Rya-r44F*, *CG1776*, *CG5023*, *zormin*) να υπεραντιπροσωπεύονται στο θηλυκό, και δη συζευγμένο,

ιστό και υπάρχει η υπόθεση ότι η διαφορική λειτουργία της μυϊκής σύσπασης είναι πιθανόν να υποδηλώνει ρόλο των ACPs στην ωορρηξία, την ωοαπόθεση και την αποθήκευση του σπέρματος (Baker, 2007).

Η *Mrp20* παρουσιάζει πτώση και στους δύο ιστούς μετά τη σύζευξη. Ενδεχομένως το φαινόμενο αυτό να υποδηλώνει την έναρξη των μορφολογικών αλλαγών και μυϊκών κινήσεων στον αναπαραγωγικό σωλήνα του θηλυκού που, όπως αναφέρθηκε στην Εισαγωγή, αποτελούν χαρακτηριστικό της αναπαραγωγικής φυσιολογίας.

#### **c57633\_g2\_i2 – Βακτηριακή φωσφατάση**

Το πεπτίδιο που ταυτοποιείται φασματοσκοπικά με το μετάγραφο αυτό δεν αποτελεί τμήμα των πεπτιδικών αλληλουχιών που προκύπτουν από τη μετάφραση των ORF. Το μεγαλύτερο ORF (48 αμινοξέα) βρίσκεται στο πλαίσιο ανάγνωσης 2 και εμφανίζει μικρή ομοιότητα με τη βακτηριακή φωσφατάση εφρίνης (WP\_004512139.1) του *Geobacter metallireducens*.

#### **c75091\_g1\_i1 – Υποθετική βακτηριακή πρωτεΐνη**

Το μοναδικό ORF που βρίσκεται στο μετάγραφο c75091 έχει μήκος 38 αμινοξέα και αντιστοιχεί σε μία υποθετική πρωτεΐνη του *Pseudoalteromonas flavipulchra* (WP\_010607725.1) με πολύ χαμηλά σκορ. Το πεπτίδιο που ταυτοποιείται φασματοσκοπικά με το μετάγραφο αυτό δεν αποτελεί τμήμα του μεταφρασμένου ORF.

Η χαμηλή ομοιότητα που παρουσιάζουν τα παραπάνω δύο μετάγραφα με βακτηριακές πρωτεΐνες δεν επιτρέπει να γίνουν υποθέσεις αναφορικά με το μετα-πρωτόμα του δάκου.

### **3.3.2 Μεταγραφή στον ιστό TESTES**

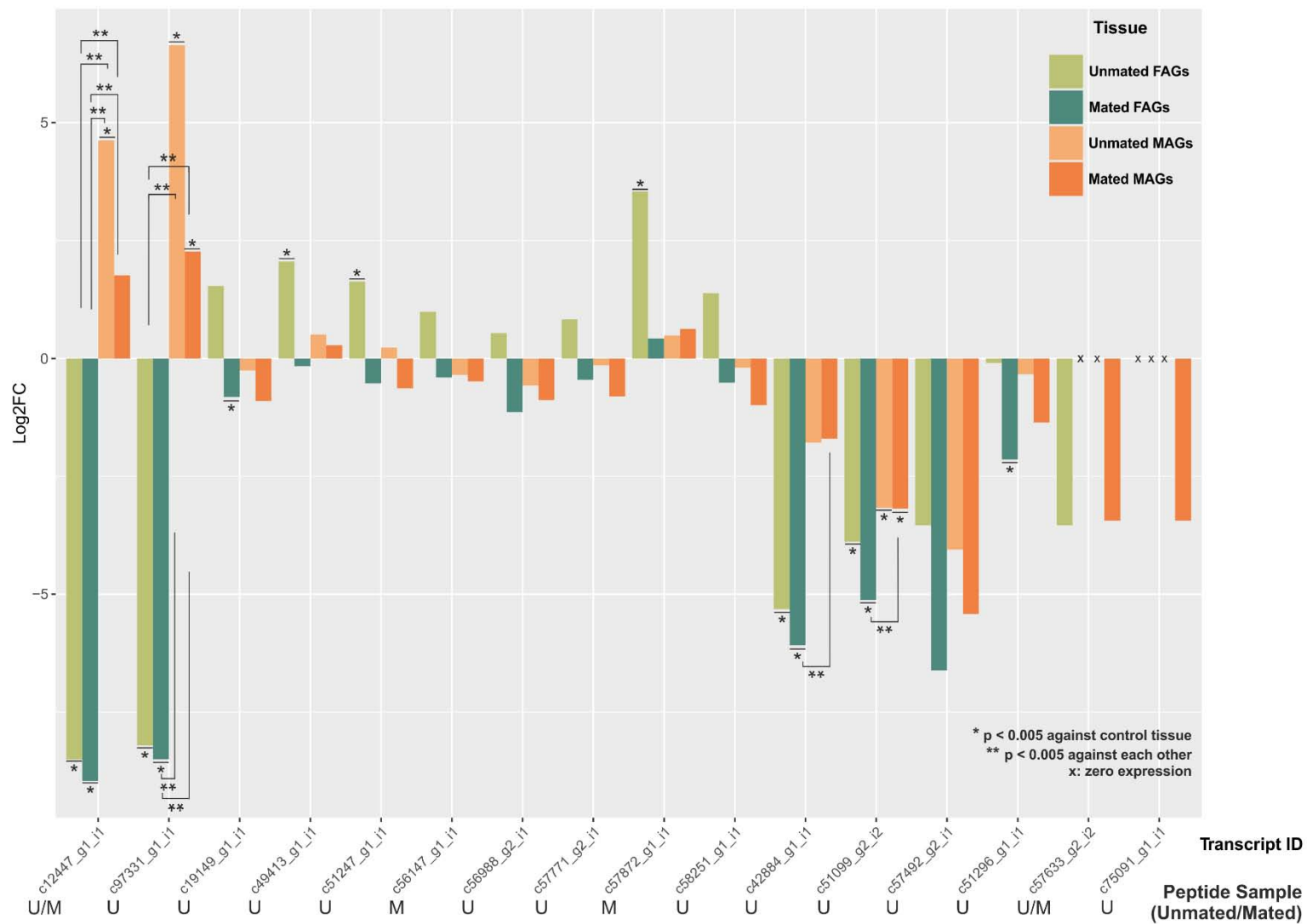
#### **c28791\_g1\_i1**

Το πεπτίδιο QLPAVQLPNLVQL που βρέθηκε φασματοσκοπικά δεν αποτελεί μέρος κανενός εκ των ανοικτών αναγνωστικών πλαισίων που προκύπτουν από την αλληλουχία του μεταγράφου αυτού. Τα δύο μεγαλύτερα ORFs βρίσκονται στο πλαίσιο ανάγνωσης 1. Το πρώτο φέρεται να κωδικοποιεί για μία αλληλουχία 97 αμινοξέων που έχει μικρή ομολογία με μία μη ταυτοποιημένη πρωτεΐνη του πρωτεοβακτηρίου *Burkholderia stagnalis* (WP\_060184636.1), ενώ το δεύτερο για μία αλληλουχία 95 αμινοξέων με μικρή ομολογία με ένα μέλος της οικογένειας μεταγραφικών παραγόντων *LysR* ενός *Rhizobium* (WP\_057471200.1).

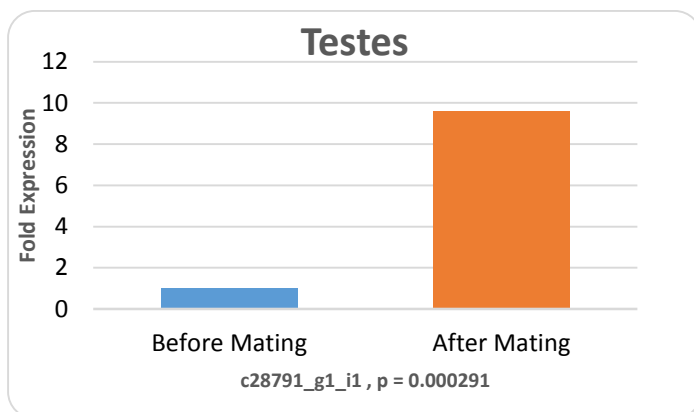
Παρά την αδυναμία εξαγωγής συμπερασμάτων σχετικών με την πιθανή λειτουργία του μεταγράφου, η έκφρασή του αυξάνεται στους όρχεις μετά τη σύζευξη (Εικόνα 3.3.2).

#### **Απουσία μεταγραφής σε αναπαραγωγικούς ιστούς**

Στα μετάγραφα όπου αντιστοιχήθηκαν από την αναζήτηση Mascot με πεπτίδια και δεν παρατηρείται μεταγραφή από την ανάλυση RNA-seq (c40343\_g1\_i1, c53521\_g2\_i1, c50896\_g1\_i1, c32983\_g1\_i1, c84026\_g1\_i1, c41479\_g1\_i2, c102929\_g1\_i1, c49224\_g1\_i1, c49805\_g2\_i1, c40025\_g1\_i1 και c55040\_g1\_i1), εμφανίζονται κάποια ORFs, όμως, αποδίδουν πολύ αδύναμες στοιχίσεις, συχνά με προκαρυωτικές αλληλουχίες με τον αλγόριθμο BLASTP και τα πεπτίδια του Mascot δεν είναι παρόντα σε κανένα ανοικτό αναγνωστικό πλαίσιο.



Εικόνα 3.3.1 Σχετική έκφραση των μεταγράφων που αντιστοιχούν σε πεπτιδία στους βοηθητικούς αδένες του αρσενικού (MAGs) και του θηλικού (FAGs) ως προς τον ιστό ελέγχου (GUT)



Εικόνα 3.3.2 Έκφραση του μεταγράφου c28791 στους όρχεις

### 3.4 Αναζήτηση για πιθανά ACP μόρια στους αναπαραγωγικούς αδένες του θηλυκού

Τα μετάγραφα που έχει ταυτοποιηθεί ότι κωδικοποιούν πρωτεΐνες των βοηθητικών αδένων του αρσενικού (ACPs) στη *Drosophila melanogaster*, την *Ceratitidis capitata* και το *Anopheles gambiae* (Πίνακας 6 Παραρτήματος) χρησιμοποιήθηκαν για την κατασκευή μίας βάσης δεδομένων. Ο λόγος που επιλέχθηκαν τα μετάγραφα και όχι οι πρωτεΐνες οι ίδιες είναι το γεγονός ότι σε αρκετές περιπτώσεις η αμινοξική αλληλουχία δεν είναι γνωστή, καθώς η μελέτη έχει γίνει σε επίπεδο μεταγραφής. Επιχειρήθηκε με χρήση των προγραμμάτων BLAST να μελετηθεί το κατά πόσο κάποια ACP βρίσκεται παρούσα στα δείγματα. Χρησιμοποιήθηκαν τόσο τα πεπτίδια που ταυτοποιήθηκαν προσυζευκτικά, όσο και μετασυζευκτικά, με το σκεπτικό ότι ενδεχομένως να υπάρχει κάποια πρωτεΐνη ή ένα πεπτίδιο που παράγεται από το θηλυκό και έχει παρόμοιο ρόλο και αλληλουχία με κάποια ACP που μεταφέρεται από το αρσενικό κατά τη σύζευξη. Η ανάλυση των πεπτιδίων με βάση δεδομένων ένα σύνολο μεταγράφων (TBLASTN) αποτελεί μία προβληματική μέθοδο που δίνει ασαφή αποτελέσματα χαμηλής στατιστικής σημασίας, καθώς:

- i. Τα κριτήρια του BLAST προσαρμόστηκαν για να βρίσκουν αντιστοιχία μεταξύ μικρών ακολουθιών και να μην την απορρίπτουν. Το επιτρεπτό όριο e-value ρυθμίστηκε στο 20000, ο πίνακας αμινοξικών αντικαταστάσεων ήταν ο PAM30, δηλαδή ο ενδεικνυόμενος για μεγάλες εξελικτικές αποστάσεις, και κάποια φίλτρα που αγνοούν περιοχές χαμηλής πολυπλοκότητας απενεργοποιήθηκαν.
- ii. Το μήκος των πεπτιδίων που

ταυτοποιήθηκαν κυμαίνεται από 6 μέχρι 43 αμινοξέα. Το μεγάλο μέγεθος των μεταγράφων που απαρτίζουν τη βάση δεδομένων αυξάνει σημαντικά το χώρο αναζήτησης (search space) και την πιθανότητα να παρατηρηθούν μικρές στοιχίσεις.

Επομένως, οι στοιχίσεις που παρατηρήθηκαν με την ανάλυση TBLASTN δεν αποτελούν ασφαλή ένδειξη για παρουσία ACPs στα πεπτιδομικά δείγματα, παρά τα σχετικά χαμηλά e-values και την πλήρη στοιχίση. Είναι πιο πολύ ένα αδύναμο σημάδι ότι τα σχετικά πεπτίδια πιθανόν να μοιράζονται κοινούς ρόλους με κάποια μόρια ACPs.

Αν κάποιο από τα πεπτίδια της πεπτιδομικής ανάλυσης είναι στην πραγματικότητα τμήμα μίας πρωτεΐνης ετερόλογης για κάποια ACP, θα ήταν αναμενόμενο η νουκλεοτιδική τους αλληλουχία να εμφανίζει σημαντική ομοιότητα. Για να αποσαφηνιστεί αυτό το ενδεχόμενο, πραγματοποιήθηκε αυτή τη φορά BLASTN, με τις παραμέτρους στις συνήθεις τιμές, μεταξύ των μεταγράφων που αντιστοιχήθηκαν με πεπτίδια και των μεταγράφων των ACPs. Από αυτήν την ανάλυση ξεχώρισαν μόνο οι τρεις ισομορφές του μεταγράφου c51099\_g2, που στοιχήθηκαν ακριβώς στον ίδιο τόπο με το ACP μετάγραφο της *Ceratitidis capitata* DQ406812 (σημανση με γκρι στον Πίνακα 6). Στην συνέχεια, πραγματοποιήθηκε και tBLASTx (Εικόνα 3.4.1) μεταξύ των παραπάνω αλληλουχιών για να εξακριβωθεί ότι το πεπτίδιο που αντιστοιχεί με το c51099\_g2, το `ARSPSDDAEVIDVPR`, βρίσκεται στο ίδιο αναγνωστικό πλαίσιο. Το DQ406812 έχει άγνωστη λειτουργία, ενώ στο c51099\_g2, όπως αναφέρθηκε στην παραπάνω ενότητα, αποδίδεται μία ομοιότητα με μία αλλεργιογόνο, άγνωστης λειτουργίας, εκκρινόμενη πρωτεΐνη Icararin-like του *Bactrocera oleae*.



gij89277017|gb|DQ406812.1| Ceratitis capitata clone 33a mRNA sequence

Sequence ID: Icl|Query\_201721 Length: 559 Number of Matches: 19

Range 1: 126 to 332 [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Positives	Gaps	Frame
140 bits(301)	2e-62	54/69(78%)	61/69(88%)	0/69(0%)	+3/-3
Query 636	RRHSNDDSDSSVGEPDFKDFRVIDTGVGYPFLOPHPLPFSFNLGFFDAFEDIFRRLRN	815	RR ++DSDE S GE DF+DF+GVIDTGVGYPFLOPHPLPFSFNLGFFDAFEDIF+RLRN		
Sbjct 332	RRGYDESDSDEGSAGETDFRDFKGVIDTGVGYPFLOPHPLPFSFNLGFFDAFEDIFQRLRN	153			
Query 816	QIWSSSDDD 842				
Sbjct 152	Q+WS + D QLWSQALPD 126				

Range 2: 3 to 122 [Graphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

Score	Expect	Identities	Positives	Gaps	Frame
80.8 bits(170)	2e-62	32/40(80%)	35/40(87%)	0/40(0%)	+3/-3
Query 837	DDLGAFDTSKGNNTTSTVKVVDGHHKIEVNETVYGDADNLFK	956	+ + + D SKGNNTTSTVKVVDGHHKIEVNETVYGD DNLFK		
Sbjct 122	ETISSLDPSKGNNTTSTVKVVDGHHKIEVNETVYGDADNLFK	3			

Range 3: 345 to 398 [Graphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

Score	Expect	Identities	Positives	Gaps	Frame
40.0 bits(81)	2e-62	16/18(89%)	17/18(94%)	0/18(0%)	+3/-3
Query 588	VPARSPSDDAEVIDVPRR 641				
Sbjct 398	+PA SPSDDAEVIDVPRR LPAGSPSDDAEVIDVPRR 345				

Εικόνα 3.4.1 TBLASTX μεταξύ του μεταγράφου c51099\_g2 και του Acp μεταγράφου της Ceratitis capitata DQ406812. Το πεπτιδίο ARSPSDDAEVIDVPR παρατηρείται στο C-τελικό άκρο.

### 3.5 Σύνοψη

Στην παρούσα εργασία επιχειρήθηκε να ταυτοποιηθούν βιοενεργά πεπτιδία και τμήματα πρωτεϊνών που έχουν ένα σημαντικό ρόλο στην αναπαραγωγή του δάκου της ελιάς, *Bactrocera oleae*. Έμφαση δόθηκε συγκεκριμένα στη συζευκτική διαδικασία και στο θηλυκό αναπαραγωγικό σύστημα, από το οποίο έγινε απομόνωση των μικρών πεπτιδίων προσυζευκτικά και μετασυζευκτικά και φασματοσκοπική τους ανάλυση. Ταυτοποιήθηκε ένας αριθμός τμημάτων πρωτεϊνών που σχετίζονται με την αύξηση του μεταβολισμού, την εμβρυϊκή ανάπτυξη, την ενεργοποίηση της μετάφρασης και των μεταμεταφραστικών τροποποιήσεων, τη μυϊκή δραστηριότητα του αναπαραγωγικού σωλήνα και άλλες βιολογικές διαδικασίες. Ακόμα βρέθηκε μία αντιστοιχία ενός μεταγράφου με ένα άγνωστης λειτουργίας ACP μετάγραφο της μεσογειακής μύγας.

Αξίζει να μελετηθεί περαιτέρω το μεταγραφικό προφίλ των αντίστοιχων γονιδίων και πιθανόν να παρατηρηθούν διαφωτιστικοί και χρήσιμοι φαινότυποι από σίγηση κάποιων από αυτά.

Το γεγονός ότι ο χαρακτηρισμός του γονιδιώματος του δάκου δεν είχε ολοκληρωθεί μέχρι την ολοκλήρωση της παρούσας εργασίας και ότι η προσπάθεια ταυτοποίησης των πεπτιδίων έγινε με τη χρήση της τρανσκριπτομικής

βιβλιοθήκης από προηγούμενα RNA-seq και με το σύνολο των πρωτεϊνών των *Drosophila*, αποτελεί έναν σημαντικό παράγοντα που επηρεάζει την ποσότητα και την ποιότητα των αποτελεσμάτων. Αφενός, για την κατασκευή της τρανσκριπτομικής βιβλιοθήκης λαμβάνονται υπ' όψιν οι περιοχές που εμφανίζουν σχετικά καλό coverage, επομένως μικρά reads, που είναι πιθανόν να αφορούν σε μικρές ποσότητες μεταγράφων ή σε μικρά ORFs (sORFs), δεν αλληλοσυμπληρώνονται από άλλα reads, δε σχηματίζουν configs και δε συμπεριλαμβάνονται. Από την άλλη πλευρά, η πρωτεϊνική βάση δεδομένων των *Drosophila* καταλαμβάνεται κυρίως από μεγάλου μεγέθους πρωτεΐνες και η ύπαρξη εξελικτικής απόστασης με τις αλληλουχίες του δάκου δυσκολεύουν την ταυτοποίηση.

Είναι βέβαιο πως υπάρχουν ακόμα στα δεδομένα της φασματοσκοπικής ανάλυσης πεπτιδία που δεν έχουν ταυτοποιηθεί και, μέχρι να αποκτηθεί μία βιβλιοθήκη που να είναι πιο πολύ εμπλουτισμένη σε μικρές αλληλουχίες, ο εντοπισμός και ο περαιτέρω χαρακτηρισμός τους θα παραμείνει δύσκολος. Όταν αυτό συμβεί, θα είναι πολύ πληροφοριακό το να επεκταθεί ο πεπτιδομικός χαρακτηρισμός της αναπαραγωγής, τόσο στα αρσενικά αναπαραγωγικά όργανα, όσο και σε άλλες συνθήκες, όπως μετασυζευκτικά χρονικά διαστήματα και ωοαπόθεση.

## 4. Παράρτημα: Πίνακες

Πίνακας 1. Το σύνολο των πεπτιδίων που έγιναν αποδεκτά με χειροκίνητη επαλήθευση (manual validation)							
Ασύζευκτα - Τρανσκριπτομική Βιβλιοθήκη		Συζευγμένα – Τρανσκριπτομική Βιβλιοθήκη		Ασύζευκτα – Πρωτεϊνική Βιβλιοθήκη (SWISSPROT)		Συζευγμένα – Πρωτεϊνική Βιβλιοθήκη (SWISSPROT)	
Κωδικός Μεταγράφου	Πεπτίδιο	Κωδικός Μεταγράφου	Πεπτίδιο	Κωδικός Πρωτεΐνης	Πεπτίδιο	Κωδικός Πρωτεΐνης	Πεπτίδιο
c12447_g1_i1	EDSAENSVLVKRRINAAKILKVVKP[VV/T]KA[L/I]AP[L/I][L/I][L/I]	c12447_g1_i1	EDSAENSVLVKR	OR35A_DROME	QAGLKILQ	OR35A_DROME	QAGLKILQ
c12447_g1_i1	QAAIIRAVQP	c12447_g1_i1	INAAK[I/L][K/Q]VV[K/Q]PVVKA LAP	THR_DROVI	NAAKLIQV	THR_DROVI	NAAKLIQV
c12447_g1_i1	NIALAPILL	c56147_g1_i1	KVVAIAP	SPNE_DROAN	KA AVKII	SPNE_DROAN	KA AVKII
c53521_g2_i1	MGYNKGASQSGHG GFGNTRHM	c51247_g1_i1	PTKAIAPILL	NAA25_DROME	QAVALLQ	BRM_DROME	VGARQRITAA
c97331_g1_i1	LDELNLDHFLAHVN	c57872_g1_i1	KVAACKLL	FBRL_DROER	KLAAAVLGGV	Y7065_DROME	SASAGGGGGVVGA
c58251_g1_i1	VIKSVNFKDASGGK VTKAAEKATKGKK	c57633_g2_i2	KLAIKVNIP	RS3A_DROSE	SEGAVIDRPEGYE PPVQE	TRX_DROVI	QLGLAQIAR
c57492_g2_i1	[K/Q]VAA[K/Q][L/I]			RS3A_DROSE	DIDPERSF	TRX_DROME	ACALVSPGGSSQGG
c56988_g2_i1	QAALARIGV			C28A5_DROME	QLLVINP	TRX_DROME	PQPTATP
c55040_g1_i1	KAVAKII			C28A5_DROME	KPTPIMS	MCM5_DROME	VVGVRAP
c55040_g1_i1	NIISLPLSKLQHSIV			ACT3_DROME	ISKEEYDESGPGIV HRK	HEAT1_DROME	KPTAQQLI
c55040_g1_i1	DRLADYRTYR			MCM5_DROME	VVGVRAP	DAAF1_DROPE	EAASGDVDSIVK
c75091_g1_i1	KVAGAAVGVKLLK			MCM5_DROME	ITAPRPEH	MCTS1_DROME	DLILPKK
c51099_g2_i2	ARSPDDAEVIDVPR			SIF1_DROME	VVVGGLGVAKP	FACR2_DROME	KVVPVV
c51099_g2_i2	NIDPEVKY			Y3800_DROME	GGGGGPGGVGGG GGGGGGMRGN DGGGMRR	DSCL_DROME	LLQVKVP
c41479_g1_i2	PKIASQEDEDPTY LF			AIMP2_DROME	NLAKVPVNPALP K	DSCL_DROME	LHTLIVQVP
c102929_g1_i1	KAAINIL			BCD_DROME	ASACRVLVK	TITIN_DROME	KVREKVVKT
c51296_g1_i1	AGSNKGATQAGQNI GASRKILLGK			ROP_DROME	RNIVPILL	TITIN_DROME	IDNVGGIH
c32983_g1_i1	AGAVAQLQLV			PUR4_DROME	KPAPKDLEQ	WHITE_DROME	QVLAVVP
c84026_g1_i1	FGGGGGPTGGGG GGGPIGGGGGGGPI GGGGGPIGGGGG RG			DOT1L_DROME	KQLKNLPE	KLP68_DROPS	MPNVRNI
c50896_g1_i1	AKTPSVKTKKAAVN KKPVTKKAPTKK			DOT1L_DROME	IEVPPP	MMS22_DROME	KLVARP
c57633_g2_i2	KLAIKVNIP			DOT1L_DROME	LRGPRL	ATPA_DROME	KGIRPAINV
c40025_g1_i1	KILVVKPV			ANKHM_DROME	KVQVPVNAI	BX42_DROME	RRLVARGAP
c28791_g1_i1	QLPAVQLPNLVQL			WNT4_DROME	KLGIIVPGGQLP	EIF3C_DROAN	TIELVLQY
c42884_g1_i1	FLWRKYHQYRYPS LP			GAWKY_DROME	QPTSQQQP		
c42884_g1_i1	ESPISTNTPLVPRKG EGEFA			CP9F2_DROME	LKGLNKILKV		
c49224_g1_i1	KRILVPLLI			WDR48_DROGR	IRGGAAIK		
c49413_g1_i1	HSHGGIPGENLPFGL NNRYKLTAY			RTBS_DROME	LKVNPKQ		
c19149_g1_i1	SEGAVVERPEGYEP PVQE			MON2_DROME	KVARAKPQ		
c57771_g2_i1	NIDFH			MON2_DROME	QVLFPLLDNVRAL SS		
c40343_g1_i1	DVGVGGSLSANPRG GAD			CLH_DROME	AALAPKAIL		
c49805_g2_i1	ISKEEYDESGPGIVH RK			SUCA_DROME	KLVGGISPKKGGT QHLGL		
				CAZ_DROME	GGGGGGGRYDRG GGGGGGGGN VQPR		
				PDFR_DROME	VRAAIVLLPL		
				SUV39_DROME	KDVPKP		
				RBGPR_DROME	PLQAAAKLIQKVG R		



				<b>MSH6_DROME</b>	VRTVLGGILKEPV P		
				<b>C5210_DROME</b>	AKELPNK		
				<b>2AAA_DROME</b>	KLLLPTVLL		
				<b>BOP1_DROPE</b>	KVVPVV		
				<b>MED15_DROME</b>	PTQRVPL		
				<b>MED15_DROME</b>	APVPGGPGTA		
				<b>MED15_DROME</b>	TLQSPVANHTL		
				<b>AP3D_DROME</b>	GRLIAEQLLDVAI RVPV		
				<b>GBB2_DROME</b>	NKVQIIPL		
				<b>RL7A_DROME</b>	QLFEKRPK		
				<b>PCAT1_DROME</b>	VQPVLLK		

Πίνακας 2. Όροι GO για τα έγκυρα πεπτιδία προσυζευκτικά. Οι καταχωρίσεις με σκούρο γκρι στο τέλος αφορούν blast hits των μεταγράφων.

Κωδικός	Όνομασία	Gene ontology (biological process)	Gene ontology (molecular function)	Gene ontology (cellular component)
Q9V3Q2	OR35A_DROME	detection of chemical stimulus involved in sensory perception of smell [GO:0050911]	odorant binding [GO:0005549]; olfactory receptor activity [GO:0004984]	dendrite membrane [GO:0032590]; integral component of membrane [GO:0016021]; plasma membrane [GO:0005886]
Q6V3V8	THR_DROVI	cell division [GO:0051301]; <b>Malpighian tubule morphogenesis [GO:0007443]</b> ; mitotic sister chromatid segregation [GO:0000070]		cytoplasm [GO:0005737]
B3M383	SPNE_DROAN	gene silencing by RNA [GO:0031047]; meiotic cell cycle [GO:0051321]; multicellular organismal development [GO:0007275]; <b>oogenesis [GO:0048477]</b> ; <b>spermatogenesis [GO:0007283]</b>	ATP binding [GO:0005524]; helicase activity [GO:0004386]; metal ion binding [GO:0046872]; nucleic acid binding [GO:0003676]	cytoplasm [GO:0005737]
Q9VDQ7	NAA25_DROME	activation of immune response [GO:0002253]; border follicle cell migration [GO:0007298]; immune response [GO:0006955]; innate immune response [GO:0045087]; negative regulation of neuron apoptotic process [GO:0043524]; N-terminal peptidyl-methionine acetylation [GO:0017196]; phagocytosis, engulfment [GO:0006911]; positive regulation of lamellipodium assembly [GO:0010592]; sensory neuron axon guidance [GO:0097374]		cytoplasm [GO:0005737]; lysosome [GO:0005764]; NatB complex [GO:0031416]
Q81F4	FBRL_DROER	RNA processing [GO:0006396]; rRNA processing [GO:0006364]; tRNA processing [GO:0008033]	methyltransferase activity [GO:0008168]; RNA binding [GO:0003723]	dense fibrillar component [GO:0001651]; ribonucleoprotein complex [GO:0030529]
B4IL76	RS3A_DROSE	multicellular organismal development [GO:0007275]; <b>oogenesis [GO:0048477]</b> ; translation [GO:0006412]	structural constituent of ribosome [GO:0003735]	cytosolic small ribosomal subunit [GO:0022627]
Q9V419	C28A5_DROME	insecticide catabolic process [GO:0046701]; response to DDT [GO:0046680]	electron carrier activity [GO:0009055]; heme binding [GO:0020037]; iron ion binding [GO:0005506]; monooxygenase activity [GO:0004497]; oxidoreductase activity [GO:0016491]; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen [GO:0016705]	endoplasmic reticulum membrane [GO:0005789]
P53501	ACT3_DROME	centrosome organization [GO:0051297]; cytoskeleton organization [GO:0007010]; heart development [GO:0007507]; mitotic cytokinesis [GO:0000281]; neuromuscular junction development, skeletal muscle fiber [GO:0098529]; positive regulation of NFAT protein import into nucleus [GO:0051533]; regulation of synapse organization [GO:0050807]	ATP binding [GO:0005524]; structural constituent of cytoskeleton [GO:0005200]	actin filament [GO:0005884]; cytoplasm [GO:0005737]; microtubule associated complex [GO:0005875]; plasma membrane [GO:0005886]
Q9VGW6	MCM5_DROME	chromosome condensation [GO:0030261]; DNA duplex unwinding [GO:0032508]; DNA endoreduplication [GO:0042023]; DNA replication initiation [GO:0006270]; meiotic nuclear division [GO:0007126]; resolution of meiotic recombination intermediates [GO:0000712]	ATP binding [GO:0005524]; chromatin binding [GO:0003682]; DNA helicase activity [GO:0003678]; DNA replication origin binding [GO:0003688]	MCM complex [GO:0042555]; nucleus [GO:0005634]
P91621	SIF1_DROME	actin cytoskeleton organization [GO:0030036]; positive regulation of filopodium assembly [GO:0051491]; positive regulation of GTPase activity [GO:0043547]; regulation of axonogenesis [GO:0050770]; regulation of GTPase activity [GO:0043087]; regulation of Rho protein signal transduction [GO:0035023]; regulation of synapse structure or activity [GO:0050803]	Rac guanyl-nucleotide exchange factor activity [GO:0030676]; receptor signaling protein activity [GO:0005057]	cell junction [GO:0030054]; intracellular [GO:0005622]; synapse [GO:0045202]
Q8T8R1	Y3800_DROME		nucleic acid binding [GO:0003676]; zinc ion binding [GO:0008270]	
Q9VUR3	AIMP2_DROME	translation [GO:0006412]		cytosol [GO:0005829]; nucleus [GO:0005634]

P09081	BCD_DROME	anterior/posterior axis specification [GO:0009948]; anterior/posterior axis specification, <b>embryo</b> [GO:0008595]; anterior region determination [GO:0007355]; <b>maternal determination of anterior/posterior axis, embryo</b> [GO:0008358]; negative regulation of translation [GO:0017148]; <b>oogenesis</b> [GO:0048477]; positive regulation of transcription from RNA polymerase II promoter [GO:0045944]; regulation of transcription, DNA-templated [GO:0006355]; regulation of transcription from RNA polymerase II promoter [GO:0006357]; regulation of translation [GO:0006417]; segment polarity determination [GO:0007367]; transcription, DNA-templated [GO:0006351]	morphogen activity [GO:0016015]; mRNA 3'-UTR binding [GO:0003730]; sequence-specific DNA binding [GO:0043565]; transcription factor activity, RNA polymerase II distal enhancer sequence-specific binding [GO:0003705]; transcription factor activity, sequence-specific DNA binding [GO:0003700]; translation regulator activity [GO:0045182]; translation repressor activity [GO:0030371]	nucleus [GO:0005634]
Q07327	ROP_DROME	exocytosis [GO:0006887]; mitotic cytokinesis [GO:000281]; neurotransmitter secretion [GO:0007269]; protein transport [GO:0015031]; <b>regulation of pole plasm oskar mRNA localization</b> [GO:0007317]; response to light stimulus [GO:0009416]; secretion by cell [GO:0032940]; synaptic transmission [GO:0007268]; synaptic vesicle priming [GO:0016082]; synaptic vesicle transport [GO:0048489]; vesicle docking involved in exocytosis [GO:0006904]; vesicle-mediated transport [GO:0016192]	SNARE binding [GO:0000149]	cytoplasm [GO:0005737]; cytosol [GO:0005829]; extrinsic component of membrane [GO:0019898]; membrane [GO:0016020]; terminal bouton [GO:0043195]
P35421	PUR4_DROME	'de novo' IMP biosynthetic process [GO:0006189]; glutamine metabolic process [GO:0006541]; purine nucleotide biosynthetic process [GO:0006164]; ribonucleoside monophosphate biosynthetic process [GO:0009156]	ATP binding [GO:0005524]; metal ion binding [GO:0046872]; phosphoribosylformylglycinamide synthase activity [GO:0004642]	cytoplasm [GO:0005737]
Q8INR6	DOT1L_DROME	chromatin-mediated maintenance of transcription [GO:0048096]; chromatin silencing [GO:0006342]; chromatin silencing at telomere [GO:0006348]; histone H3-K79 methylation [GO:0034729]; histone methylation [GO:0016571]; lateral inhibition [GO:0046331]; positive regulation of gene silencing by miRNA [GO:2000637]; regulation of cell cycle [GO:0051726]	histone methyltransferase activity (H3-K79 specific) [GO:0031151]	histone methyltransferase complex [GO:0035097]; nucleus [GO:0005634]
Q9VCA8	ANKHM_DROME	<b>antimicrobial humoral response</b> [GO:0019730]; centrosome duplication [GO:0051298]; compound eye photoreceptor cell differentiation [GO:0001751]; negative regulation of apoptotic process [GO:0043066]; positive regulation of JAK-STAT cascade [GO:0046427]; transmembrane receptor protein tyrosine kinase signaling pathway [GO:0007169]	RNA binding [GO:0003723]	cytoplasm [GO:0005737]; cytosol [GO:0005829]; lipid particle [GO:0005811]
P40589	WNT4_DROME	cell fate commitment [GO:0045165]; cell migration [GO:0016477]; establishment of imaginal disc-derived wing hair orientation [GO:0001737]; <b>female gonad development</b> [GO:0008585]; motor neuron axon guidance [GO:0008045]; neural retina development [GO:0003407]; neuron differentiation [GO:0030182]; non-canonical Wnt signaling pathway [GO:0035567]; regulation of hemocyte proliferation [GO:0035206]; retinal ganglion cell axon guidance [GO:0031290]; salivary gland morphogenesis [GO:0007435]; synaptic target inhibition [GO:0016201]; Wnt signaling pathway [GO:0016055]; Wnt signaling pathway, planar cell polarity pathway [GO:0060071]	frizzled binding [GO:0005109]; signal transducer activity [GO:0004871]	cell surface [GO:0009986]; extracellular region [GO:0005576]; extracellular space [GO:0005615]; proteinaceous extracellular matrix [GO:0005578]
Q8SY33	GAWKY_DROME	<b>embryonic development via the syncytial blastoderm</b> [GO:0001700]; gene silencing by miRNA [GO:0035195]; locomotor rhythm [GO:0045475]; mRNA catabolic process [GO:0006402]; regulation of protein localization [GO:0032880]; regulation of translation [GO:0006417]	nucleotide binding [GO:0000166]; RNA binding [GO:0003723]	cytoplasmic mRNA processing body [GO:0000932]
Q9VG82	CP9F2_DROME	wing disc development [GO:0035220]	heme binding [GO:0020037]; iron ion binding [GO:0005506]; monooxygenase activity [GO:0004497]; oxidoreductase activity [GO:0016491]; oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen [GO:0016705]	endoplasmic reticulum membrane [GO:0005789]
B4JH6	WDR48_DROGR			

<b>Q95SX7</b>	RTBS_DROME	RNA-dependent DNA replication [GO:0006278]; transposition, DNA-mediated [GO:0006313]	RNA-directed DNA polymerase activity [GO:0003964]	
<b>Q9VLT1</b>	MON2_DROME	actin cytoskeleton reorganization [GO:0031532]; <b>pole plasm protein localization [GO:0007318]</b> ; positive regulation of GTPase activity [GO:0043547]; protein transport [GO:0015031]; vesicle-mediated transport [GO:0016192]	ARF guanyl-nucleotide exchange factor activity [GO:0005086]; cytoskeletal regulatory protein binding [GO:0005519]	endosome [GO:0005768]; Golgi apparatus [GO:0005794]; trans-Golgi network [GO:0005802]
<b>P29742</b>	CLH_DROME	compound eye development [GO:0048749]; compound eye retinal cell programmed cell death [GO:0046667]; dsRNA transport [GO:0033227]; endocytosis [GO:0006897]; extracellular matrix organization [GO:0030198]; intracellular protein transport [GO:0006886]; liquid clearance, open tracheal system [GO:0035002]; neurotransmitter secretion [GO:0007269]; <b>oocyte microtubule cytoskeleton polarization [GO:0008103]</b> ; <b>pole plasm oskar mRNA localization [GO:0045451]</b> ; positive regulation of endocytosis [GO:0045807]; positive regulation of Notch signaling pathway [GO:0045747]; puparial adhesion [GO:0007594]; regulation of growth [GO:0040008]; regulation of tube length, open tracheal system [GO:0035159]; secretory granule organization [GO:0033363]; <b>sperm individualization [GO:0007291]</b> ; synaptic transmission [GO:0007268]; synaptic vesicle coating [GO:0016183]; synaptic vesicle exocytosis [GO:0016079]	clathrin light chain binding [GO:0032051]; structural molecule activity [GO:0005198]	cell cortex [GO:0005938]; clathrin-coated vesicle [GO:0030136]; clathrin coat of coated pit [GO:0030132]; clathrin coat of trans-Golgi network vesicle [GO:0030130]; clathrin complex [GO:0071439]; clathrin vesicle coat [GO:0030125]; coated pit [GO:0005905]; coated vesicle [GO:0030135]; cytoplasmic vesicle [GO:0031410]; lipid particle [GO:0005811]; perinuclear region of cytoplasm [GO:0048471]; plasma membrane [GO:0005886]; secretory granule [GO:0030141]; synaptic vesicle [GO:0008021]; trans-Golgi network [GO:0005802]
<b>Q94522</b>	SUCA_DROME	neurogenesis [GO:0022008]; nucleoside triphosphate biosynthetic process [GO:0009142]; succinate metabolic process [GO:0006105]; succinyl-CoA metabolic process [GO:0006104]; tricarboxylic acid cycle [GO:0006099]	cofactor binding [GO:0048037]; GTP binding [GO:0005525]; succinate-CoA ligase (ADP-forming) activity [GO:0004775]; succinate-CoA ligase (GDP-forming) activity [GO:0004776]	mitochondrial matrix [GO:0005759]; mitochondrion [GO:0005739]
<b>Q27294</b>	CAZ_DROME	adult locomotory behavior [GO:0008344]; compound eye development [GO:0048749]; larval locomotory behavior [GO:0008345]; mRNA splicing, via spliceosome [GO:0000398]; positive regulation of synaptic growth at neuromuscular junction [GO:0045887]; synaptic growth at neuromuscular junction [GO:0051124]; transcription initiation from RNA polymerase II promoter [GO:0006367]	mRNA binding [GO:0003729]; nucleotide binding [GO:0000166]; zinc ion binding [GO:0008270]	catalytic step 2 spliceosome [GO:0071013]; nucleoplasm [GO:0005654]; nucleus [GO:0005634]; transcription factor TFIID complex [GO:0005669]
<b>Q9W4Y2</b>	PDFR_DROME	cell surface receptor signaling pathway [GO:0007166]; circadian behavior [GO:0048512]; circadian rhythm [GO:0007623]; circadian sleep/wake cycle [GO:0042745]; circadian temperature homeostasis [GO:0060086]; copper ion homeostasis [GO:0055070]; G-protein coupled receptor signaling pathway [GO:0007186]; gravitaxis [GO:0042332]; locomotor rhythm [GO:0045475]; neuropeptide signaling pathway [GO:0007218]; positive regulation of adenylate cyclase activity involved in G-protein coupled receptor signaling pathway [GO:0010579]; regulation of circadian sleep/wake cycle [GO:0042749]	calcitonin receptor activity [GO:0004948]; neuropeptide receptor activity [GO:0008188]	integral component of membrane [GO:0016021]; integral component of plasma membrane [GO:0005887]; neuronal cell body [GO:0043025]; neuron projection [GO:0043005]
<b>P45975</b>	SUV39_DROME	chromatin organization [GO:0006325]; chromatin silencing [GO:0006342]; chromatin silencing at centromere [GO:0030702]; chromatin silencing at telomere [GO:0006348]; chromosome organization [GO:0051276]; DNA methylation [GO:0006306]; <b>female germ-line stem cell asymmetric division [GO:0048132]</b> ; <b>gene silencing [GO:0016458]</b> ; heterochromatin assembly [GO:0031507]; heterochromatin organization involved in chromatin silencing [GO:0070868]; histone H3-K9 methylation [GO:0051567]; histone methylation [GO:0016571]; <b>histone modification [GO:0016570]</b> ; negative regulation of response to gamma radiation [GO:2001229]; <b>oogenesis [GO:0048477]</b> ; transcription, DNA-templated [GO:0006351]	histone methyltransferase activity [GO:0042054]; histone methyltransferase activity (H3-K9 specific) [GO:0046974]; zinc ion binding [GO:0008270]	chromosome, centromeric region [GO:0000775]; heterochromatin [GO:0000792]; nucleus [GO:0005634]; polytene chromosome chromocenter [GO:0005701]
<b>Q9VKB9</b>	RBGPR_DROME	border follicle cell migration [GO:0007298]; neurotransmitter secretion [GO:0007269]; regulation of GTPase activity [GO:0043087]; regulation of synaptic activity [GO:0060025]; vesicle-mediated transport [GO:0016192]	enzyme regulator activity [GO:0030234]; GTPase activator activity [GO:0005096]; protein heterodimerization activity [GO:0046982]	synaptic vesicle [GO:0008021]

Q9VUM0	MSH6_DROME	DNA repair [GO:0006281]; intrinsic apoptotic signaling pathway in response to DNA damage [GO:0008630]; maintenance of DNA repeat elements [GO:0043570]; meiotic mismatch repair [GO:0000710]; negative regulation of DNA recombination [GO:0045910]; reciprocal meiotic recombination [GO:0007131]; response to UV [GO:0009411]	ATP binding [GO:0005524]; damaged DNA binding [GO:0003684]; DNA-dependent ATPase activity [GO:0008094]; four-way junction DNA binding [GO:0000400]; guanine/thymine mispair binding [GO:0032137]; single base insertion or deletion binding [GO:0032138]	MutSalpa complex [GO:0032301]; nuclear chromosome [GO:0000228]
Q23997	C5210_DROME	carbohydrate metabolic process [GO:0005975]; cuticle chitin catabolic process [GO:0006036]; multicellular organismal development [GO:0007275]	chitinase activity [GO:0004568]; chitin binding [GO:0008061]	extracellular region [GO:0005576]; extracellular space [GO:0005615]
P36179	2AAA_DROME	autophagy [GO:0006914]; centriole replication [GO:0007099]; centrosome cycle [GO:0007098]; centrosome duplication [GO:0051298]; centrosome organization [GO:0051297]; chromosome segregation [GO:0007059]; mitotic spindle assembly [GO:0090307]; mitotic spindle organization [GO:0007052]; neurogenesis [GO:0022008]; phagocytosis [GO:0006909]; protein dephosphorylation [GO:0006470]; regulation of protein phosphatase type 2A activity [GO:0034047]; spindle assembly [GO:0051225]	protein phosphatase type 2A regulator activity [GO:0008601]	centriole [GO:0005814]; cytoplasm [GO:0005737]; protein phosphatase type 2A complex [GO:0000159]
B4GIU9	BOP1_DROPE	maturation of 5.8S rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) [GO:0000466]; maturation of LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) [GO:0000463]	ribonucleoprotein complex binding [GO:0043021]	nucleolus [GO:0005730]; nucleoplasm [GO:0005654]; preribosome, large subunit precursor [GO:0030687]
Q9Y149	MED15_DROME	transcription from RNA polymerase II promoter [GO:0006366]; transcription initiation from RNA polymerase II promoter [GO:0006367]	RNA polymerase II transcription cofactor activity [GO:0001104]	mediator complex [GO:0016592]; nucleus [GO:0005634]
P54362	AP3D_DROME	compound eye pigmentation [GO:0048072]; determination of adult lifespan [GO:0008340]; endocytosis [GO:0006897]; exocytosis [GO:0006887]; eye pigment biosynthetic process [GO:0006726]; eye pigment granule organization [GO:0008057]; Golgi to endosome transport [GO:0006895]; intracellular protein transport [GO:0006886]; intracellular transport [GO:0046907]; lysosomal transport [GO:0007041]; lysosome organization [GO:0007040]; negative regulation of gene silencing by RNA [GO:0060967]; Notch receptor processing [GO:0007220]; ocellus pigment biosynthetic process [GO:0008055]; ommochrome biosynthetic process [GO:0006727]	protein transporter activity [GO:0008565]	AP-3 adaptor complex [GO:0030123]; clathrin-coated vesicle membrane [GO:0030665]; endoplasmic reticulum [GO:0005783]; Golgi-associated vesicle [GO:0005798]; Golgi stack [GO:0005795]; late endosome [GO:0005770]
P29829	GBB2_DROME	activation of phospholipase C activity [GO:0007202]; deactivation of rhodopsin mediated signaling [GO:0016059]; G-protein coupled receptor signaling pathway [GO:0007186]; phototransduction [GO:0007602]; rhodopsin mediated signaling pathway [GO:0016056]; visual perception [GO:0007601]	GTPase activity [GO:0003924]; protein heterodimerization activity [GO:0046982]; signal transducer activity [GO:0004871]	cytoplasm [GO:0005737]; heterotrimeric G-protein complex [GO:0005834]; plasma membrane [GO:0005886]; rhabdomere [GO:0016028]
P46223	RL7A_DROME	centrosome duplication [GO:0051298]; centrosome organization [GO:0051297]; mitotic spindle elongation [GO:0000022]; mitotic spindle organization [GO:0007052]; ribosome biogenesis [GO:0042254]	structural constituent of ribosome [GO:0003735]	ribosome [GO:0005840]
Q0KHU5	PCAT1_DROME	lipid modification [GO:0030258]; phospholipid biosynthetic process [GO:0008654]	1-acylglycerophosphocholine O-acyltransferase activity [GO:0047184]; calcium ion binding [GO:0005509]; lysophospholipid acyltransferase activity [GO:0071617]	endoplasmic reticulum [GO:0005783]; endoplasmic reticulum membrane [GO:0005789]; Golgi apparatus [GO:0005794]; Golgi membrane [GO:0000139]; integral component of membrane [GO:0016021]; lipid particle [GO:0005811]
Q29CU2	RS3A_DROPS	multicellular organismal development [GO:0007275]; oogenesis [GO:0048477]; translation [GO:0006412]	structural constituent of ribosome [GO:0003735]	cytosolic small ribosomal subunit [GO:0022627]
P08736	EF1A1_DROME	determination of adult lifespan [GO:0008340]; translational elongation [GO:0006414]	GTPase activity [GO:0003924]; GTP binding [GO:0005525]; translation elongation factor activity [GO:0003746]	cytosol [GO:0005829]; eukaryotic translation elongation factor 1 complex [GO:0005853]; lipid particle [GO:0005811]; microtubule associated complex [GO:0005875]
Q9V9A7	MCCB_DROME	leucine catabolic process [GO:0006552]; regulation of eclosion [GO:0007563]	ATP binding [GO:0005524]; methylcrotonoyl-CoA carboxylase activity [GO:0004485]	mitochondrial matrix [GO:0005759]
Q14667	K0100_HUMAN			extracellular region [GO:0005576]

Q5BLY4	ICA_APIME			extracellular region [GO:0005576]
Q9Y297	FBW1A_HUMAN	<p>anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process [GO:0031145]; branching involved in mammary gland duct morphogenesis [GO:0060444]; cellular response to organic cyclic compound [GO:0071407]; circadian rhythm [GO:0007623]; Fc-epsilon receptor signaling pathway [GO:0038095]; G2/M transition of mitotic cell cycle [GO:0000086]; innate immune response [GO:0045087]; <b>mammary gland epithelial cell proliferation [GO:0033598]</b>; mitotic cell cycle [GO:0000278]; MyD88-dependent toll-like receptor signaling pathway [GO:0002755]; MyD88-independent toll-like receptor signaling pathway [GO:0002756]; negative regulation of sequence-specific DNA binding transcription factor activity [GO:0043433]; negative regulation of smoothed signaling pathway [GO:0045879]; negative regulation of transcription, DNA-templated [GO:0045892]; NIK/NF-kappaB signaling [GO:0038061]; positive regulation of circadian rhythm [GO:0042753]; positive regulation of proteolysis [GO:0045862]; positive regulation of transcription, DNA-templated [GO:0045893]; positive regulation of ubiquitin-protein ligase activity involved in regulation of mitotic cell cycle transition [GO:0051437]; proteasome-mediated ubiquitin-dependent protein catabolic process [GO:0043161]; protein dephosphorylation [GO:0006470]; protein destabilization [GO:0031648]; protein polyubiquitination [GO:0000209]; protein ubiquitination [GO:0016567]; regulation of circadian rhythm [GO:0042752]; regulation of I-kappaB kinase/NF-kappaB signaling [GO:0043122]; regulation of proteasomal protein catabolic process [GO:0061136]; regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle [GO:0051439]; SCF-dependent proteasomal ubiquitin-dependent protein catabolic process [GO:0031146]; signal transduction [GO:0007165]; stimulatory C-type lectin receptor signaling pathway [GO:0002223]; stress-activated MAPK cascade [GO:0051403]; T cell receptor signaling pathway [GO:0050852]; toll-like receptor 10 signaling pathway [GO:0034166]; toll-like receptor 2 signaling pathway [GO:0034134]; toll-like receptor 3 signaling pathway [GO:0034138]; toll-like receptor 4 signaling pathway [GO:0034142]; toll-like receptor 5 signaling pathway [GO:0034146]; toll-like receptor 9 signaling pathway [GO:0034162]; toll-like receptor signaling pathway [GO:0002224]; toll-like receptor TLR1:TLR2 signaling pathway [GO:0038123]; toll-like receptor TLR6:TLR2 signaling pathway [GO:0038124]; TRIF-dependent toll-like receptor signaling pathway [GO:0035666]; tumor necrosis factor-mediated signaling pathway [GO:0033209]; ubiquitin-dependent protein catabolic process [GO:0006511]; viral process [GO:0016032]; Wnt signaling pathway [GO:0016055]</p>	<p>ligase activity [GO:0016874]; ubiquitin protein ligase activity [GO:0061630]</p>	<p>cytosol [GO:0005829]; nucleoplasm [GO:0005654]; SCF ubiquitin ligase complex [GO:0019005]</p>
P14318	MP20_DROME	<p>cell adhesion [GO:0007155]; myoblast fusion [GO:0007520]; regulation of cell shape [GO:0008360]; regulation of myoblast fusion [GO:1901739]</p>	<p>actin binding [GO:0003779]</p>	

Πίνακας 3. Όροι GO για τα έγκυρα πεπτιδία μετασυζευκτικά. Οι καταχωρίσεις με σκούρο γκρι στο τέλος αφορούν blast hits των μεταγράφων.

Κωδικός	Όνομασία	Gene ontology (biological process)	Gene ontology (molecular function)	Gene ontology (cellular component)
Q9V3Q2	OR35A_DROME	<b>detection of chemical stimulus involved in sensory perception of smell [GO:0050911]</b>	<b>odorant binding [GO:0005549]; olfactory receptor activity [GO:0004984]</b>	dendrite membrane [GO:0032590]; integral component of membrane [GO:0016021]; plasma membrane [GO:0005886]
Q6V3V8	THR_DROVI	cell division [GO:0051301]; <b>Malpighian tubule morphogenesis [GO:0007443]</b> ; mitotic sister chromatid segregation [GO:0000070]		cytoplasm [GO:0005737]
B3M383	SPNE_DROAN	<b>gene silencing by RNA [GO:0031047]</b> ; meiotic cell cycle [GO:0051321]; multicellular organismal development [GO:0007275]; <b>oogenesis [GO:0048477]; spermatogenesis [GO:0007283]</b>	ATP binding [GO:0005524]; helicase activity [GO:0004386]; metal ion binding [GO:0046872]; nucleic acid binding [GO:0003676]	cytoplasm [GO:0005737]
P25439	BRM_DROME	ATP-dependent chromatin remodeling [GO:0043044]; axonogenesis [GO:0007409]; dendrite guidance [GO:0070983]; dendrite morphogenesis [GO:0048813]; hemocyte proliferation [GO:0035172]; hippo signaling [GO:0035329]; histone H3-K27 acetylation [GO:0043974]; imaginal disc-derived leg morphogenesis [GO:0007480]; imaginal disc-derived wing margin morphogenesis [GO:0008587]; imaginal disc-derived wing vein morphogenesis [GO:0008586]; imaginal disc-derived wing vein specification [GO:0007474]; intestinal stem cell homeostasis [GO:0036335]; muscle organ development [GO:0007517]; negative regulation of G1/S transition of mitotic cell cycle [GO:2000134]; negative regulation of neuroblast proliferation [GO:0007406]; neurogenesis [GO:0022008]; neuron development [GO:0048666]; oogenesis [GO:0048477]; phagocytosis [GO:0006909]; positive regulation of epidermal growth factor receptor signaling pathway [GO:0045742]; positive regulation of stem cell proliferation [GO:2000648]; positive regulation of transcription, DNA-templated [GO:0045893]; regulation of innate immune response [GO:0045088]; regulation of transcription from RNA polymerase II promoter [GO:0006357]; transcription, DNA-templated [GO:0006351]	ATPase activity [GO:0016887]; ATP binding [GO:0005524]; helicase activity [GO:0004386]	brahma complex [GO:0035060]; nucleus [GO:0005634]; PBAF complex [GO:0070604]
Q7YZA2	Y7065_DROME			
Q24742	TRX_DROVI	multicellular organismal development [GO:0007275]; transcription, DNA-templated [GO:0006351]	histone methyltransferase activity (H3-K4 specific) [GO:0042800]; sequence-specific DNA binding [GO:0043565]; transcription factor activity, sequence-specific DNA binding [GO:0003700]; zinc ion binding [GO:0008270]	histone methyltransferase complex [GO:0035097]
P20659	TRX_DROME	axon guidance [GO:0007411]; <b>germ cell migration [GO:0008354]; histone acetylation [GO:0016573]</b> ; histone H3 acetylation [GO:0043966]; histone H3-K4 methylation [GO:0051568]; <b>histone methylation [GO:0016571]</b> ; positive regulation of transcription elongation from RNA polymerase II promoter [GO:0032968]; regulation of response to DNA damage stimulus [GO:2001020]; transcription, DNA-templated [GO:0006351]	chromatin binding [GO:0003682]; histone methyltransferase activity (H3-K4 specific) [GO:0042800]; identical protein binding [GO:0042802]; protein homodimerization activity [GO:0042803]; protein phosphatase 1 binding [GO:0008157]; sequence-specific DNA binding [GO:0043565]; transcription factor activity, sequence-specific DNA binding [GO:0003700]; transcription regulatory region DNA binding [GO:0044212]; zinc ion binding [GO:0008270]	histone acetyltransferase complex [GO:0000123]; histone methyltransferase complex [GO:0035097]; microtubule associated complex [GO:0005875]; MLL1/2 complex [GO:0044665]; nucleus [GO:0005634]; polytene chromosome [GO:0005700]; transcription elongation factor complex [GO:0008023]
Q9V6W6	MCM5_DROME	chromosome condensation [GO:0030261]; DNA duplex unwinding [GO:0032508]; DNA endoreduplication [GO:0042023]; DNA replication initiation [GO:0006270]; meiotic nuclear division [GO:0007126]; resolution of meiotic recombination intermediates [GO:0000712]	ATP binding [GO:0005524]; chromatin binding [GO:0003682]; DNA helicase activity [GO:0003678]; DNA replication origin binding [GO:0003688]	MCM complex [GO:0042555]; nucleus [GO:0005634]
Q9VM75	HEAT1_DROME	cellular response to starvation [GO:0009267]; maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) [GO:0000462]; neurogenesis [GO:0022008]; positive regulation of transcription from RNA polymerase I promoter [GO:0045943]	snoRNA binding [GO:0030515]	90S preribosome [GO:0030686]; small-subunit processome [GO:0032040]; t-UTP complex [GO:0034455]
B4GT53	DAAF1_DROPE	cilium morphogenesis [GO:0060271]	dynein binding [GO:0045502]	axoneme [GO:0005930]
Q9W445	MCTS1_DROME	positive regulation of translational initiation [GO:0045948]; translation reinitiation [GO:0002188]	RNA binding [GO:0003723]	cytoplasm [GO:0005737]



A1ZA13	FACR2_DROME	lipid metabolic process [GO:0006629]; long-chain fatty-acyl-CoA metabolic process [GO:0035336]; wax biosynthetic process [GO:0010025]	fatty-acyl-CoA reductase (alcohol-forming) activity [GO:0080019]; long-chain-fatty-acyl-CoA reductase activity [GO:0050062]	integral component of membrane [GO:0016021]; peroxisome [GO:0005777]
Q9VS29	DSCL_DROME	homophilic cell adhesion via plasma membrane adhesion molecules [GO:0007156]; neuron projection morphogenesis [GO:0048812]	identical protein binding [GO:0042802]	integral component of plasma membrane [GO:0005887]
Q9I7U4	TITIN_DROME	cell division [GO:0051301]; <b>locomotion [GO:0040011]</b> ; mesoderm development [GO:0007498]; mitotic chromosome condensation [GO:0007076]; muscle attachment [GO:0016203]; muscle organ development [GO:0007517]; myoblast fusion [GO:0007520]; regulation of hemocyte proliferation [GO:0035206]; sarcomere organization [GO:0045214]; sister chromatid cohesion [GO:0007062]; skeletal muscle tissue development [GO:0007519]; somatic muscle development [GO:0007525]; visceral muscle development [GO:0007522]	actin binding [GO:0003779]; <b>structural constituent of muscle [GO:0008307]</b>	condensed nuclear chromosome [GO:0000794]; I band [GO:0031674]; microtubule associated complex [GO:0005875]; sarcomere [GO:0030017]; striated muscle myosin thick filament [GO:0005863]; Z disc [GO:0030018]
P10090	WHITE_DROME	cellular biogenic amine biosynthetic process [GO:0042401]; cGMP transport [GO:0070731]; compound eye pigmentation [GO:0048072]; eye pigment metabolic process [GO:0042441]; eye pigment precursor transport [GO:0006856]; gravitaxis [GO:0042332]; male courtship behavior [GO:0008049]; memory [GO:0007613]; ommochrome biosynthetic process [GO:0006727]; signal transduction [GO:0007165]; transmembrane transport [GO:0055085]	ATPase activity, coupled to transmembrane movement of substances [GO:0042626]; ATP binding [GO:0005524]; pigment binding [GO:0031409]; transmembrane signaling receptor activity [GO:0004888]	cytoplasmic vesicle [GO:0031410]; integral component of membrane [GO:0016021]; plasma membrane [GO:0005886]
Q29DY1	KLP68_DROPS	<b>microtubule-based movement [GO:0007018]</b>	ATP binding [GO:0005524]; microtubule motor activity [GO:0003777]; motor activity [GO:0003774]	cytoplasm [GO:0005737]; kinesin complex [GO:0005871]; microtubule [GO:0005874]; microtubule associated complex [GO:0005875]
O46089	MMS22_DROME	double-strand break repair via homologous recombination [GO:0000724]; replication fork processing [GO:0031297]		nuclear replication fork [GO:0043596]
P35381	ATPA_DROME	ATP hydrolysis coupled proton transport [GO:0015991]; ATP synthesis coupled proton transport [GO:0015986]; electron transport chain [GO:0022900]; growth [GO:0040007]; lipid storage [GO:0019915]; phagocytosis [GO:0006909]; regulation of choline O-acetyltransferase activity [GO:1902769]; <b>spermatid development [GO:0007286]</b>	ATP binding [GO:0005524]; proton-transporting ATPase activity, rotational mechanism [GO:0046961]; proton-transporting ATP synthase activity, rotational mechanism [GO:0046933]	lipid particle [GO:0005811]; microtubule associated complex [GO:0005875]; mitochondrial inner membrane [GO:0005743]; mitochondrion [GO:0005739]; proton-transporting ATP synthase complex, catalytic core F1 [GO:0045261]
P39736	BX42_DROME	<b>embryonic development via the syncytial blastoderm [GO:0001700]</b> ; eye-antennal disc development [GO:0035214]; mRNA splicing, via spliceosome [GO:0000398]	DNA binding [GO:0003677]	catalytic step 2 spliceosome [GO:0071013]; precatalytic spliceosome [GO:0071011]; Prp19 complex [GO:0000974]
B3MIF1	EIF3C_DROAN	formation of translation preinitiation complex [GO:0001731]; regulation of translational initiation [GO:0006446]	translation initiation factor activity [GO:0003743]	eukaryotic 43S preinitiation complex [GO:0016282]; eukaryotic 48S preinitiation complex [GO:0033290]; eukaryotic translation initiation factor 3 complex [GO:0005852]
Q503E1	INT11_DANRE		hydrolase activity [GO:0016787]	nucleus [GO:0005634]
Q8BK75	ELP6_MOUSE	positive regulation of cell migration [GO:0030335]; regulation of transcription, DNA-templated [GO:0006355]; transcription, DNA-templated [GO:0006351]		Elongator holoenzyme complex [GO:0033588]
Q9W3V8	ALG8_DROME	<b>apical constriction involved in gastrulation [GO:0003384]</b> ; <b>germ-band extension [GO:0007377]</b> ; oligosaccharide-lipid intermediate biosynthetic process [GO:0006490]; protein glycosylation [GO:0006486]; protein N-linked glycosylation [GO:0006487]	dolichyl pyrophosphate Glc1Man9GlcNAc2 alpha-1,3-glucosyltransferase activity [GO:0042283]; dolichyl pyrophosphate Man9GlcNAc2 alpha-1,3-glucosyltransferase activity [GO:0042281]	endoplasmic reticulum [GO:0005783]; endoplasmic reticulum membrane [GO:0005789]; integral component of membrane [GO:0016021]



Πίνακας 4. Σχετικές εκφράσεις (log2) των μεταγράφων που αντιστοιχούν σε πεπτιδία στους αναπαραγωγικούς ιστούς σε σύγκριση με τον ιστό ελέγχου GUT (σκόιαση όπου  $p < 0.005$ )

Μεταγραφο	Συνθήκη	blast_hit Evalue bitscore	FAGs προσυζευκτικά – GUT		FAGs μετασυζευκτικά – GUT		GUT – MAGs προσυζευκτικά		GUT – MAGs μετασυζευκτικά		Όργανο Προ/Μετα-συζευκτικά	
			log2FC	p-value	log2FC	p-value	log2FC	p-value	log2FC	p-value	log2FC	p-value
c28791_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	<b>-3.263</b>	0.000291
c19149_g1_i1	Before Mating	sp Q29CU2 RS3A_DROPS 7e-168 475	<b>1.536435218</b>	0.00443	-0.8167261	0.1772538	0.252956	0.69065	0.9026975	0.1364054	NA	NA
c40343_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c53521_g2_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c50896_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c12447_g1_i1	Before & After Mating	-	<b>-8.502215253</b>	6E-21	<b>-8.9611243</b>	3.649E-22	<b>-4.62403</b>	3.39E-06	-1.765412	0.0131762	NA	NA
c58251_g1_i1	Before Mating	sp P08736 EF1A1_DROME 0.0 903	1.385727301	0.00965	-0.5134955	0.3966015	0.196243	0.762755	0.9878088	0.1027242	NA	NA
c32983_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c51247_g1_i1	After Mating	sp Q503E1 INT11_DANRE 0.0 202	<b>1.631339643</b>	0.00355	-0.527151	0.3921002	-0.22774	0.657737	0.6327157	0.3035571	NA	NA
c84026_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c49413_g1_i1	Before Mating	sp Q9V9A7 MCCB_DROME 0.0 1028	<b>2.054201127</b>	0.00025	-0.1646686	0.811019	-0.50808	0.349434	-0.279179	0.5978193	NA	NA
c41479_g1_i2	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c102929_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c75091_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	3.4406427	0.6209677	NA	NA
c97331_g1_i1	Before Mating	-	<b>-8.204655912</b>	6.3E-19	<b>-8.5007165</b>	1.504E-19	<b>-6.63781</b>	0.000597	<b>-2.26822</b>	0.0293231	NA	NA
c42884_g1_i1	Before Mating	sp Q9W4L1 RM33_DROME 5e-30 116	<b>-5.317237281</b>	3E-11	<b>-6.0841074</b>	2.469E-13	1.784256	0.015257	1.7002683	0.0200415	NA	NA
c57771_g2_i1	Before Mating	sp Q14667 K0100_HUMAN 7.0 36.2	0.829018881	0.32898	-0.4528858	0.563052	0.145505	0.890081	0.8058135	0.2657612	NA	NA
c49224_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c51099_g2_i2	Before Mating	sp Q5BLY4 ICA_APIME 1e-07 56.6	<b>-3.885308728</b>	3.5E-08	<b>-5.1205748</b>	1.167E-11	<b>3.163243</b>	3.03E-06	<b>3.18304</b>	2.693E-06	NA	NA
c56147_g1_i1	After Mating	sp Q8BK75 ELP6_MOUSE 0.045 39.7	0.990219731	0.11542	-0.4028543	0.554157	0.34825	0.604189	0.4856146	0.4694121	NA	NA
c57633_g2_i2	Before & After Mating	-	-3.538541646	0.62097	NA	NA	NA	NA	3.4406427	0.6209677	NA	NA
c57492_g2_i1	Before Mating	sp A1L162 ERIC2_HUMAN 2e-08 60.1	-3.538541646	0.62097	-6.6153232	0.0163992	4.054783	0.432796	5.4200061	0.0847412	NA	NA
c49805_g2_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c57872_g1_i1	After Mating	sp Q9W3V8 ALG8_DROME 9e-179 545	<b>3.541372161</b>	0.00296	0.42340166	0.5710039	-0.4882	0.536992	-0.62423	0.4271957	NA	NA
c40025_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c55040_g1_i1	Before Mating	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
c56988_g2_i1	Before Mating	sp Q9Y297 FBW1A_HUMAN 0.0 821	0.539658645	0.33871	-1.138163	0.0662676	0.573288	0.357919	0.8815623	0.1542955	NA	NA
c51296_g1_i1	Before Mating	sp P14318 MP20_DROME 3e-106 330	-0.094127882	0.90157	<b>-2.1428943</b>	0.0007702	0.335999	0.587026	1.3613925	0.0265389	NA	NA

Πίνακας 5. Ανάλυση ORF των μεταγράφων που αντιστοιγήθηκαν φασματοσκοπικά με πεπτίδια (* : Το πεπτίδιο είναι παρόν σε άλλο, μικρότερο ORF του μεταγράφου που δε δίνει BLASTP hits)							
Κωδικός μεταγράφου	Εμφάνιση πεπτιδίου στα FAGs στη συνθήκη:	Εμφάνιση μεταγράφου στους ιστούς:	Πεπτίδιο παρόν στο μεταφρασμένο ORF	ORF		ORF BLASTP	
				Αναγνωστικό πλαίσιο	Μέγεθος μεταφρασμένου ORF	Hit	ID
c12447_g1_i1	Before & After Mating	MAGs/FAGs	+	1	192	Ceratoxin A  <i>Ceratitidis capitata</i>	XP_004523341.1
c97331_g1_i1	Before Mating	MAGs/FAGs	+	3	84	ompA  <i>Maritimibacter alkaliphilus</i>	WP_008329170.1
c19149_g1_i1	Before Mating	MAGs/FAGs	+	1	269	RS3A  <i>Drosophila yakuba</i>	XP_002099572.1
c49413_g1_i1	Before Mating	MAGs/FAGs	-	1	583	carboxylase b-chain methylcrotonoyl-CoA  <i>Bactrocera oleae</i>	XP_014103629.1
c51247_g1_i1	After Mating	MAGs/FAGs	-	3	1033	sry H-1  <i>Bactrocera oleae</i>	XP_014093312.1
c56147_g1_i1	After Mating	MAGs/FAGs	+*	3	268	ELP6  <i>Ceratitidis capitata</i>	XP_004530004.1
c56988_g2_i1	Before Mating	MAGs/FAGs	-	1	587	UBAP1  <i>Bactrocera dorsalis</i>	XP_011211222.1
c57771_g2_i1	Before Mating	MAGs/FAGs	-	1	2287	KIA0100  <i>Bactrocera oleae</i>	XP_014094229.1
c57872_g1_i1	After Mating	MAGs/FAGs	+	3	1599	Dolichyl pyrophosphate alpha-1,3-glucosyltransferase  <i>Bactrocera oleae</i>	XP_014101855.1
c58251_g1_i1	Before Mating	MAGs/FAGs	+	2	464	Elongation Factor 1-alpha 1  <i>Bactrocera oleae</i>	XP_014093560.1
c42884_g1_i1	Before Mating	MAGs/FAGs	-	3	65	L33 ribosomal protein  <i>Bactrocera cucurbitae</i>	XP_011188950.1
c51099_g2_i2	Before Mating	MAGs/FAGs	+	3	210	Icparin-like  <i>Bactrocera oleae</i>	XP_014101837.1
c57492_g2_i1	Before Mating	MAGs/FAGs	-	2	448	DDB_G0288805  <i>Bactrocera oleae</i>	XP_014092398.1
c51296_g1_i1	Before Mating	MAGs/FAGs	+	3	185	Muscle-specific Protein 20   <i>Bactrocera oleae</i>	XP_014099975.1
c57633_g2_i2	Before & After Mating	MAGs/FAGs	-	2	48	-	-
c75091_g1_i1	Before Mating	MAGs/FAGs	-	3	38	-	-
c28791_g1_i1	Before Mating	Testes	-	1	97	-	-
c40343_g1_i1	Before Mating	-	-	2	39	-	-
c53521_g2_i1	Before Mating	-	-	1	100	-	-
c50896_g1_i1	Before Mating	-	-	2	30	-	-
c32983_g1_i1	Before Mating	-	-	2	58	-	-
c84026_g1_i1	Before Mating	-	-	2	136	-	-
c41479_g1_i2	Before Mating	-	-	3	157	-	-
c102929_g1_i1	Before Mating	-	-	2	27	-	-
c49224_g1_i1	Before Mating	-	-	2	38	-	-
c49805_g2_i1	Before Mating	-	-	2	91	-	-
c40025_g1_i1	Before Mating	-	-	3	30	-	-
c55040_g1_i1	Before Mating	-	-	2	95	-	-

Πίνακας 6. Accessory gland proteins από τα είδη <i>D.melanogaster</i> , <i>A.gambiae</i> , <i>C.capitata</i> που χρησιμοποιήθηκαν στην αναζήτηση ACP								
<i>Drosophila melanogaster</i>				<i>Anopheles gambiae</i>				<i>Ceratitis capitata</i>
Hdc	Cyp304a1	CG4774	CG31705	AGAP009001	AGAP012706	AGAP000918	AGAP005504	DQ406805
Eps-15	Cyp6a2	G0196	CG8172	AGAP009189	AGAP012807	AGAP000920	AGAP005684	DQ406817
sens-2	Cyp12b2	CG15439	CG15065	AGAP009190	AGAP012855	AGAP000926	AGAP005859	DQ406806
gsb	Cyp6a2	CG4688	CG2137	AGAP009191	AGAP012986	AGAP000927	AGAP006362	DQ406807
prd	Acp70, SP	CG7872		AGAP009339	AGAP013714	AGAP001039	AGAP006425	DQ406809
Toll6	spi	CG5644		AGAP009377	AGAP013731	AGAP001510	AGAP006589	DQ406808
pom1	sqh	CG30053		AGAP009699	AGAP013734	AGAP002181	AGAP007339	DQ406810
pyx	spo	CG6296		AGAP010358	AGAP013776	AGAP002824	AGAP008019	DQ406811
pAbp	cwo	CG8539		AGAP010359	AGAP000107	AGAP003067	AGAP008070	DQ406812
scramb1	mwh	CG13699		AGAP010637	AGAP000139	AGAP003736	AGAP008071	DQ406813
TrnC47D	Mkrn1	CG34290		AGAP010720	AGAP000355	AGAP003749	AGAP008116	DQ406814
Rab3	tam	CG34290		AGAP010861	AGAP000754	AGAP003844	AGAP008203	DQ406815
Cul-4	CG16936	CG8323		AGAP012315	AGAP000831	AGAP004761	AGAP008439	DQ406816
Syx17	CG6654	CG8539		AGAP012680	AGAP000843	AGAP005239	AGAP008997	
Takr86C	CG5973	CG9701		AGAP012681	AGAP000916	AGAP005255	AGAP009352	

## Βιβλιογραφία

- A programme for the eradication of the New World Screwworm from North Africa. (1990). Vienna: Prepared by a Consultants Group.
- Apger-McGlaughon, J. W. (2013). Post-mating change in excretion by mated *Drosophila melanogaster* females is a long-term response that depends on sex peptide and sperm. *J. Insect Physiol.* 59, 1024-30.
- Avila, F., Mattei, A., & Wolfner, M. (2010). Sex peptide is required for the efficient release of stored sperm in mated *Drosophila* females. *Genetics*, 595-600.
- Avila, F., Mattei, A., & Wolfner, M. (2015). Sex peptide receptor is required for the release of stored sperm by mated *Drosophila melanogaster* females. *Journal of Insect Physiology*.
- Aykanat, T., & Dincturk, H. (2007). An outer membrane protein A (ompA) homologue from the photosynthetic purple sulfur bacterium *Allochrochromatium vinosum*. *Microbiological Research*.
- Baer, B. H. (2009). The seminal fluid proteome of the honeybee *Apis mellifera*. *Proteomics* 9, 2085-97.
- Baker, D. A. (2007). Variable sexually dimorphic gene expression in laboratory strains of *Drosophila melanogaster*. *BMC Genomics*.
- Baldini, F., Gabrieli, P., Rogers, D., & Catteruccia, F. (2012). Function and composition of male accessory gland secretions in *Anopheles gambiae*: a comparison with other insect vectors of infectious diseases. *Pathog Glob Health* 106(2), 82-93.
- Bardou, P., Mariette, J., Escudié, F., Djemiel, C., & Klopp, C. (2014). jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics*.
- Bartlett, J. M., & Stirling, D. (2003). A Short History of the Polymerase Chain Reaction. *PCR Protocols. Methods in Molecular Biology* 226 (2nd ed.).
- Begun, D. J.-D. (2000). Molecular population genetics of Male accessory glands in *Drosophila*. *Genetics*.
- Bono, J., Matzkin, L., Hoang, K., & Brandsmeier, L. (2015 ). Molecular evolution of candidate genes involved in post-mating-prezygotic reproductive isolation. *J Evol Biol. Feb;28(2)*, 403-14.
- Burda, P., & Markus, A. (1999). The dolichol pathway of N-linked glycosylation. *Biochimica et Biophysica Acta* 1426, 239-257.
- Bushland, R., & Hopkins, D. (1953). Sterilization of Screw-Worm Flies with X-Rays. *Journal of Economic Entomology*.
- Caceres, C. J. (2004). Comparison of Mediterranean fruit fly (*Ceratitidis capitata*) (Tephritidae) bisexual and genetic sexing strains: development, evaluation and economics. *6th International Symposium on Fruit Flies of Economic Importance, 6–10 May 2002* (σσ. 367-381). Stellenbosch, South Africa: Isteg Scientific Publications, Irene, South Africa.
- Carvalho, G., Kapahi, P., Anderson, D., & Benzer, S. (2006). Allocrine modulation of feeding behavior by the Sex Peptide of *Drosophila*. *Curr Biol* 16(7), 692-696.
- Cayol, J. P. (2004). Fruit fly control and monitoring in the Near East: shared concern in a regional transboundary problem. *6th International Symposium on Fruit Flies of Economic Importance, 6–10 May 2002* (σσ. 155–171). Stellenbosch, South Africa: Isteg Scientific Publications, Irene, South Africa.
- Chapman. (2004). Functions and analysis of the seminal fluid proteins of male *Drosophila melanogaster* fruit flies. *Peptides*, 1477–1490.

- Chapman, T., Bangham, J., Vinti, G., B. Seifried, O. L., Wolfner, M., Smith, H., & Partridge, L. (2003). The sex peptide of *Drosophila melanogaster*: Female post-mating responses analyzed by using RNA interference. *PNAS*.
- Chen, P., Stumm-Zollinger, E., Aigaki, T., Balmer, J., Bienz, M., & Böhlen, P. (1988). A male accessory gland peptide that regulates reproductive behavior of female *D. melanogaster*. *Cell*, 291-8.
- Corrigan. (2014). BMP-regulated exosomes from *Drosophila* male reproductive glands reprogram female behavior. *J Cell Biol* 206(5), 671–688.
- Dalton. (2010). Dynamic, mating-induced gene expression changes in female head and brain tissues of *Drosophila melanogaster*. *BMC Genomics*, 11:541.
- Dame, D. A. (1981). Assessment of released sterile *Anopheles albimanus* and *Glossina morsitans morsitans*. *Cytogenetics and genetics of vectors. Proceedings of XVI International Congress of Entomology* (σσ. 231-248). Kyoto, Japan: Elsevier Science Publishers, Amsterdam, The Netherlands.
- Dame, D. A., Woodward, D. B., Ford, H. R., & Weidhaas, D. E. (1964). Field behavior of sexually sterile *Anopheles quadrimaculatus* males. *Mosquito News*, 6-16.
- Davies, S. J. (2006). Identification of genes expressed in the accessory glands of male Mediterranean Fruit Flies (*Ceratitis capitata*). *Insect Biochemistry and Molecular Biology*, 846-56.
- De Novo PSM Rules*. (n.d.). Ανάκτηση από <http://www.ionsource.com/tutorial/DeNovo/rules.htm>
- Dobzhansky. (1973). Nothing in Biology Makes Sense except in the Light of Evolution. *The American Biology Teacher* Vol. 35, No. 3, pp. 125-129.
- Dobzhansky, T. (1973). Nothing in Biology Makes Sense except in the Light of Evolution. *The American Biology Teacher*, 125-129.
- Dottorini, T. e. (2007). A genome-wide analysis in *Anopheles gambiae* mosquitoes reveals 46 male accessory gland genes, possible modulators of female behavior. *P. Natl. Acad. Sci. USA* 104, 16215–16220.
- Findlay, G., Yi, X., Maccoss, M., & Swanson, W. (2008). Proteomics reveals novel *Drosophila* seminal fluid proteins transferred at mating. *PLoS Biology*.
- Fisher, K. (1996). Queensland fruit fly (*Bactrocera tryoni*): eradication from Western Australia. Στο *Fruit fly pests. A world assessment of their biology and management* (σσ. 535– 541). Delray Beach, FL: St. Lucie Press.
- Fletcher, B. (1989). Life history strategies of tephritid fruit flies. Στο A. Robinson, & G. Hooper, *Fruit flies, their biology, natural enemies and control, vol 3B* (σσ. 195–208). Elsevier Science Publishers, Amsterdam,.
- Frank, A., & Pevzner, P. (2005). PepNovo: de novo peptide sequencing via probabilistic network modeling. *Analytical Chemistry*.
- Franz, A., Sanchez-Vargas, I., Adelman, Z., Blair, C., Beaty, B., & James, A. (2006). Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes aegypti*. *Proc Natl Acad Sci USA* 103, 4198-4203.
- Global Strategic Framework For Integrated Vector Management . (2004). WHO.
- Gomulski, L., Dimopoulos, G., Xi, Z., Scolari, F., Gabrieli, P., Siciliano, P., . . . Gasperi, G. (2012). Transcriptome Profiling of Sexual Maturation and Mating in the Mediterranean Fruit Fly, *Ceratitis capitata*,. *PLoS One*.
- Harris, A., Nimmo, D., McKemey, A., Kelly, N., Scaife, S., & Donnelly, C. (2011). Field performance of engineered male mosquitoes. *Nat Biotechnology* 29, 1034-1037.
- Heifetz, Y. L. (2000). The *Drosophila* seminal fluid protein Acp26Aa stimulates

- release of oocytes by the ovary. *Curr Biol* 10(2), 99–102.
- Heifetz, Y., Lung, O., Frongillo, E., & Wolfner, M. (2000). The *Drosophila* seminal fluid protein Acp26Aa stimulates release of oocytes by the ovary. *Curr Biol*, 10(2), 99-102.
- Hodin, J. (2009). She shapes events as they come: Plasticity in Female Insect Reproduction. Στο D. W. Whitman, *Phenotypic plasticity in insects, Mechanisms and Consequences*. Science Publishers.
- Höltkä, M., Zetterberg, H., Mirgorodskaya, E., Mattsson, N., Blennow, K., & Gobom, J. (2012). Peptidome Analysis of Cerebrospinal Fluid by LC-MALDI MS. *PLoS One*.
- Isaac, R. L. (2010). *Drosophila* male sex peptide inhibits siesta sleep and promotes locomotor activity in the post-mated female. *Proc. R. Soc. B* 277, 65-70.
- Kelly, T., & Smith, H. (1970). A restriction enzyme from *Hemophilus influenzae*. II. *Journal of Molecular Biology* 51 (2):, 393–409.
- Knipling, E. (1959). Sterile-Male Method of Population Control. *Science*, 902-4.
- Kocher, S., Richard, F., Tarpy, D., & Grozinger, C. (2008). Genomic analysis of post-mating changes in the honey bee queen (*Apis mellifera*). *BMC Genomics*.
- Lazareva AA, R. G. (2007). A role for the adult fat body in *Drosophila* male courtship behavior. *PLoS Genet*,3(1):e16 .
- Lederberg, J. (1952). Cell genetics and hereditary symbiosis. *Physiological Reviews*.
- Lee, K., Choo, Y., Giovannoni, S., & Cho, J. (2007 ). *Maritimibacter alkaliphilus* gen. nov., sp. nov., a genome-sequenced marine bacterium of the Roseobacter clade in the order Rhodobacterales. *Int J Syst Evol Microbiol. Jul;57(Pt 7)*, 1653-8.
- Liu, Y., H.Liu, Liu, S., Wang, S., Jiang, R., & Li, S. (2009). Hormonal and nutritional regulation of insect fat body development and function. *Arch Insect Biochem Physiol*, 71(1), 16-30.
- Lodish, H., Berk, A., & Zipursky, S. (2000). Overview of Extracellular Signaling. Στο *Molecular Cell Biology. 4th edition*.
- Lofgren, C. S., Dame, D., Breeland, S. G., Weidhaas, D. E., Jeffery, G., Kaiser, R., . . . Baldwin, K. (1974). Release of chemosterilized males for the control of *Anopheles albimanus* in El Salvador. III. Field methods and population control. *American Journal of Tropical Medicine and Hygiene*, 288-297.
- Manning, A. (1962). A Sperm Factor Affecting the Receptivity of *Drosophila Melanogaster* Females. *Nature*, 252-3.
- Marchini, D., Manetti, A., Rosetto, M., Bernini, L., Telford, J., Baldari, C., & Dallai, R. (1995). cDNA Sequence and Expression of the Ceratotoxin Gene Encoding an Antibacterial Sex-specific Peptide from the Medfly *Ceratitis capitata* (diptera). *The journal of biomedical chemistry*, 6199-6204.
- Marec, F. N. (2005). Development of genetic sexing strains in Lepidoptera: from traditional to transgenic approaches. *J Econ Entomol., Apr;98(2)*, 248-59.
- Marygold, S. J. (2007). The ribosomal protein genes and Minute loci of *Drosophila melanogaster*. *Genome Biol.* 8(10): R216.
- Mattei, A., Riccio, M., Avila, F., & Wolfner, M. (2015). Integrated 3D view of postmating responses by the *Drosophila melanogaster* female reproductive tract, obtained by micro-computed tomography scanning. *PNAS*.
- McGraw EA, O. S. (2013). Beyond insecticides: new thinking on an ancient problem. *Nat Rev Microbiol, Mar;11(3)*, 181-93.
- Miller. (1953). Production of Amino Acids Under Possible Primitive Earth Conditions. *Science* 117 (3046), 528–9.



- Miller, S. L. (1953). Production of Amino Acids Under Possible Primitive Earth Conditions. *Science*, 528–9.
- Monnard P.A., W. P. (2015). Current Ideas about Prebiological Compartmentalization. *Life (Basel)*, 1239-63.
- Monnard, P. W. (2015). Current Ideas about Prebiological Compartmentalization. *Life (Basel) Apr 10;5(2)*, 1239-63.
- Mooney C, H. N. (2012). Towards the Improved Discovery and Design of Functional Peptides: Common Features of Diverse Classes Permit Generalized Prediction of Bioactivity. *PLoS ONE*.
- Msangi, A. R. (2000). Success in Zanzibar: eradication of tsetse. *Area-Wide Control of Fruit Flies and Other Insect Pests. International Conference on Area-Wide Control of Insect Pests, and the 5th International Symposium on Fruit Flies of Economic Importance*, (σσ. 57-66). Penang, Malaysia.
- Muller, H. (1927). Artificial transmutation of the gene. *Science*, 84-87.
- Nesvizhskii, A., Roos, F., Grossmann, J., Vogelzang, M., Eddes, J., Gruissem, W., . . . Aebersold, R. (2006). Dynamic Spectrum Quality Assessment and Iterative Computational Analysis of Shotgun Proteomic Data. *Mol Cell Proteomics*.
- Neubaum, D., & Wolfner, M. (1999). Mated *Drosophila melanogaster* females require a seminal fluid protein, Acp36DE, to store sperm efficiently. *Genetics*, 153(2), 845-857.
- Oliva, C., Jacquet, M., Gilles, J., Lemperiere, G., & P.O.Maquart. (2012). The Sterile Insect Technique for Controlling Populations of *Aedes albopictus* (Diptera: Culicidae) on Reunion Island: Mating Vigour of Sterilized Male. *PLoS ONE*.
- Oloo, F. P. (2000). Integration of the sterile insect technique (SIT) and autosterilisation lethality in the eradication of *Glossina fuscipes fuscipes* Newst. (Diptera: Glossinidae) in Buvuma Islands in Lake Victoria, Uganda. *Area-Wide Control of Fruit Flies and Other Insect Pests. International Conference on Area-Wide Control of Insect Pests, and the 5th International Symposium on Fruit Flies of Economic Importance*, (σσ. 311-314). Penang, Malaysia.
- Patterson, R. S. (1970). Suppression and elimination of an island population of *Culex quinquefasciatus* with sterile males. *Science*, 1368-1370.
- Peiren, N., Graaf, D. d., Brunain, M., Bridts, C., Ebo, D., Stevens, W., & F.J.Jacobs. (2006). Molecular cloning and expression of icarapin, a novel IgE-binding bee venom protein. *FEBS Lett.*, 4895-9.
- Peng, J. C. (2005). Gradual release of sperm bound sex-peptide controls female postmating behavior in *Drosophila*. *Curr. Biol.*, 207-13.
- PepNovo+ Readme File. (n.d.).
- Picó, Y. (2015). Advanced Mass Spectrometry for Food Safety and Quality, Volume 68. *Elsevier*.
- Pilpel, N. N. (2008). Mating-increases trypsin in female *Drosophila* hemolymph. *Insect Biochem*, 320-30.
- Politzar, H. D. (1984). An integrated campaign against riverine tsetse, *Glossina palpalis gambiensis* and *Glossina tachinoides* by trapping and the release of sterile males. . *Insect Science and its Application*, 439–442.
- Qazi, M. B., & Wolfner, M. (2003). An early role for the *Drosophila melanogaster* male seminal protein Acp36DE in female sperm storage. *J. Exp. Biol.* 206, 3521–28.
- Ram, K., & Wolfner, M. (2007). Sustained post-mating response in *Drosophila melanogaster* requires multiple seminal fluid proteins. *PLoS Genetics* 3(12).

- Rogers, D., Whitten, M., Thailayil, J., Soichot, J., Levashina, E., & Catteruccia, F. (2008). Molecular and cellular components of the mating machinery in *Anopheles gambiae* females. *Proc Natl Acad Sci U.S.A.*
- Russell, P., Hertz, P., & McMillan, B. (2011). *Biology: The Dynamic Science.*
- Sagri, E., Reczko, M., Tsoumani, K., Gregoriou, M., Harokopos, V., Mavridou, A., . . . Mathiopoulos, K. (2014). The molecular biology of the olive fly comes of age. *BMC Genetics.*
- Sanger, F. N. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463–7.
- Scolari, F. e. (2012). Transcriptional profiles of mating-responsive genes from testes and male accessory glands of the Mediterranean fruit fly, *Ceratitis capitata*. *PLoS ONE* 7, e46812 .
- Simmons, L. T. (2013). Sperm and seminal fluid proteomes of the field cricket *Teleogryllus oceanicus*: identification of novel proteins transferred to females at mating. *Insect Mol. Biol.* 22, 115-30.
- Sun, J. S. (2013). Ovulation in *Drosophila* is controlled by secretory cells of the female reproductive tract. *eLife* 2:e00415.
- Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of Gene Ontology terms. *PLoS ONE.*
- Swanson, W. J.-D. (2001). Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *PNAS.*
- Takken, W. M. (1986). The eradication of *Glossina palpalis palpalis* (Robineau-Desvoidy) (Diptera: Glossinidae) using traps, insecticide-impregnated targets and the sterile insect technique in central Nigeria. *Bulletin of Entomological Research*, 275–286.
- Tessera. (2009). Life began when evolution began: a lipidic vesicle-based scenario. *Orig Life Evol Biosph. Dec;39(6)*, 559-64.
- Tessera, M. (2009). Life began when evolution began: a lipidic vesicle-based scenario. *Orig Life Evol Biosph*, 559-64.
- Traver, B., Anderson, M., & Adelman, Z. (2009). Homing endonucleases catalyze double-stranded DNA breaks and somatic transgene excision in *Aedes aegypti*. *Insect Mol Biol* 18, 623–633.
- (2008). *United States, Mexico, and Guatemala Fruit Fly Emergence and Release Facilities Review.* United States Department of Agriculture.
- Villaseñor, A. J. (2000). Current progress in the medfly program Mexico-Guatemala. *Area-Wide Control of Fruit Flies and Other Insect Pests. International* (σσ. 361-368). In K. H. Tan (ed.).
- Vincent, A., Kejzlarová-Lepesant, J., Segalat, L., Yanicostas, C., & Lepesant, J. (1988). sry h-1, a new *Drosophila melanogaster* multifingered protein gene showing maternal and zygotic expression. *Mol Cell Biol.*
- Vreysen, M. R. (2007). *Area-Wide Control of Insect Pests: From Research to Field Implementation.* 121.
- Weiss, B., & Richardson, C. C. (1967). Enzymatic breakage and joining of deoxyribonucleic acid, I. Repair of single-strand breaks in DNA by an enzyme system from *Escherichia coli* infected with T4 bacteriophage. *Proceedings of the National Academy of Sciences .*
- Whitten, M. (1969). Automated sexing of pupae and its usefulness in control by sterile insects. *J Econ Entomol* 62, 272-3.
- Whyard, S., Erdelyan, C., Partridge, A., Singh, A., Beebe, N., & Capina, R. (2015). Silencing the buzz: a new approach to population suppression of mosquitoes

- by feeding larvae double-stranded RNAs. *Parasit Vectors*.
- Wilke, A., Nimmo, D., John, O., Kojin, B., Capurro, M., & Marrelli, M. (2009). Mini-review: Genetic enhancements to the sterile insect technique to control mosquito populations. *Asia Pac J Mol Biol Biotechnol*. 17:, 65–74.
- Windbichler, N., Menichelli, M., Papathanos, P., Thyme, S., Li, H., & Ulge, U. (2011). A synthetic homing endonuclease-based gene drive system in the human malaria mosquito. *Nature* 473, 212–217.
- Winskill, P., Harris, A., Morgan, S., Stevenson, J., Raduan, N., & Alphey, L. (2014). Genetic control of *Aedes aegypti*: data-driven modelling to assess the effect of releasing different life stages and the potential for long-term suppression. *Parasit Vectors* 7:68.
- Wong, A. A. (2008). A role for Acp29AB, a predicted seminal fluid lectin, in female sperm storage in *Drosophila melanogaster*. *Genetics* 180, 921–31.
- Yasuno, M. W. (1978). A control experiment with chemosterilised male *Culex pipiens fatigans* in a village near Delhi surrounded by a breeding-free zone. *Japanese Journal of Sanitary Zoology*, 325-343.
- Yosiaki, I., Kakinohana, H., Yamagishi, M., & Kohama, T. (2003). Eradication of the Melon Fly, *Bactrocera cucurbitae*, from Okinawa, Japan, by Means of the Sterile Insect Technique, with Special Emphasis on the Role of Basic Studies. *Journal of Asia-Pacific Entomology*, 119-129.
- Zygouridis, N. E. (2014). *Population analysis of the major olive insect pest, Bactrocera oleae, with the use of molecular markers.*

**Δημοσίευση:**

RESEARCH

Open Access

# The molecular biology of the olive fly comes of age

Efthimia Sagri<sup>1</sup>, Martin Reczko<sup>2</sup>, Konstantina T Tsoumani<sup>1</sup>, Maria-Eleni Gregoriou<sup>1</sup>, Vaggelis Harokopos<sup>2</sup>, Anna-Maria Mavridou<sup>1</sup>, Spyros Tastsoglou<sup>1</sup>, Konstantinos Athanasiadis<sup>1</sup>, Jiannis Ragoussis<sup>2†</sup>, Kostas D Mathiopoulos<sup>1\*</sup>

## Abstract

**Background:** Olive cultivation blends with the history of the Mediterranean countries since ancient times. Even today, activities around the olive tree constitute major engagements of several people in the countryside of both sides of the Mediterranean basin. The olive fly is, beyond doubt, the most destructive pest of cultivated olives. The female fly leaves its eggs in the olive fruit. Upon emergence, the larvae feed on the olive sap, thus destroying the fruit. If untreated, practically all olives get infected. The use of chemical insecticides constitutes the principal olive fly control approach. The Sterile Insect Technique (SIT), an environmentally friendly alternative control method, had been tried in pilot field applications in the 1970's, albeit with no practical success. This was mainly attributed to the low, non-antagonistic quality of the mixed-sex released insects. Many years of experience from successful SIT applications in related species, primarily the Mediterranean fruit fly, *Ceratitis capitata*, demonstrated that efficient SIT protocols require the availability of fundamental genetic and molecular information.

**Results:** Among the primary systems whose understanding can contribute towards novel SIT approaches (or its recently developed alternative RIDL: Release of Insects carrying a Dominant Lethal) is the reproductive, since the ability to manipulate the reproductive system would directly affect the insect's fertility. In addition, the analysis of early embryonic promoters and apoptotic genes would provide tools that confer dominant early-embryonic lethality during mass-rearing. Here we report the identification of several genes involved in these systems through whole transcriptome analysis of female accessory glands (FAGs) and spermathecae, as well as male testes. Indeed, analysis of differentially expressed genes in these tissues revealed higher metabolic activity in testes than in FAGs/spermathecae. Furthermore, at least five olfactory-related genes were shown to be differentially expressed in the female and male reproductive systems analyzed. Finally, the expression profile of the embryonic *serendipity-α* locus and the pre-apoptotic *head involution defective* gene were analyzed during embryonic developmental stages.

**Conclusions:** Several years of molecular studies on the olive fly can now be combined with new information from whole transcriptome analyses and lead to a deep understanding of the biology of this notorious insect pest. This is a prerequisite for the development of novel embryonic lethality female sexing strains for successful SIT efforts which, combined with improved mass-reared conditions, give new hope for efficient SIT applications for the olive fly.

## Background

When Athena, the goddess of peace and wisdom, offered an olive tree to the people of Attica to sway them into choosing her name for their city - and not that of her brother's Poseidon - neither she nor the people of Attica

were aware of the 'worm' that could destroy the precious fruit of that tree. That was described much later in the 3<sup>rd</sup> century AD, by the botanist Theophrastus who, in his works "*Enquiry into Plants*" and "*Causes of Plants*" [1], talked about the 'worm underneath the skin of the olive that destroys the fruit'. Indeed, the female olive fly (*Bactrocera oleae*, Rossi) lays her eggs in an olive fruit and the resulting larva feeds on the olive sap, opening channels inside it, thus destroying it. In this way, a female fly can damage more than 300 olives in her lifetime. Given the

\* Correspondence: kmathiop@bio.uth.gr

† Contributed equally

<sup>1</sup>Department of Biochemistry and Biotechnology, University of Thessaly, Larissa, Greece

Full list of author information is available at the end of the article

fact that during the summer and fall months about five generations of these flies are born, one can imagine the cumulative damage that can take place in an olive orchard. If untreated, practically every single olive will get infested. It is estimated that due to olive fly infestation olive oil production is reduced by more than 30% annually [2].

Control of these flies is traditionally based on cover or bait sprays with chemical insecticides. During the last 40-50 years, organophosphate insecticides have been extensively used against the olive fly, mainly dimethoate and fenthion. More recently, pyrethroids as well as the naturalyte spinosad have been added in the arsenal against the olive fly. The use of chemical pesticides, however, entails many known hazards. Among these are ecological disturbances, the development and spread of insecticide resistance, harmful toxicological effects on human health [3]. Many of these risks are apparent not only to scientists but also to growers and consumers who require a cleaner and safer environment as well as products of high quality. Alternative, environmentally friendly control methods against insect pests, such as the Sterile Insect Technique (SIT) have been experimented in the past with considerable success [4]. The SIT involves the mass production, sterilization and subsequent release of the sterilized insects [5]. The sterilized males will mate with wild females, whose unfertilized eggs will never hatch, thus reducing the numbers of the following generation. In theory, if continued releases are performed over several consecutive generations, the population will progressively be reduced and, eventually, a total eradication could occur.

Given the substantial economic burden of the olive fly in olive producing countries and the concerns raised about the heavy use of insecticides to control the flies, the SIT was proposed [6] and implemented in two pilot efforts. In the early 1970s, about 150,000 laboratory-reared male and female flies were sterilized by gamma-irradiation and subsequently released in the environment [7]. Although initially the releases seemed to contribute to low infestation levels, by the end of the season olives were as highly infested as in the two nearby control plantations. The sterilized flies were proven ineffective to reduce infestation. Similar results were obtained in a second pilot SIT effort that took place in the late '70s in a small Greek island. These unsuccessful pilot experiments led to funding suspension and the eventual abandonment of the program [8-10]. Apart from the high cost and labor-intensive rearing of the olive fly, extensive research that followed these first pilot efforts revealed several key issues of olive fly biology that should have been sorted out before a successful SIT could be implemented. The first issue regarded assortative mating of the released and wild populations. Laboratory-reared flies mated several hours before

scotophase whereas wild flies mated at the end of the photophase [11]. Apparently, mass-laboratory rearing caused substantial alterations in the genetic makeup of the flies due to selective pressures in the artificial laboratory environment [12,13]. The second issue regarded the quality of the radiation-sterilized mass-reared flies. Radiation did not leave the vigor of the flies unaffected [14]. Another factor that probably exacerbated the low fitness of the laboratory reared flies was the use of antibiotics in the flies' diet that destroyed the endosymbiotic bacteria that are now known to play a very important role in the organism's fitness [15-19]. Finally, but equally importantly, extensive stinging of the olive fruits from the released females led to further fungal infestation [7].

Since those early years, several molecular and genetic studies have changed *B. oleae's* research landscape. First, the development of microsatellite markers [20] and the analysis of the mitochondrial genome [21] have offered tools for a fairly detailed analysis of population structure and dynamics in the Mediterranean basin [22-26]. Second, cytogenetic analysis, including *in situ* hybridization of several molecular markers, established the details of the chromosomal complement [27-31]. Third, isolation and characterization of various genes has shed light on important processes such as insecticide resistance [32-35], female germline differentiation and morphogenesis of epidermal cells [36], enzyme catalytic mechanisms [37], sex-determining cascades [38,39]. Fourth, an initial assessment of the genome of the olive fly was gained by an accurate estimate of its size [40] and the characterization and analysis of centromeric repeats [41] and several EST loci [42]. This was followed by a whole transcriptome analysis with 454 pyrosequencing [43]. Fifth, *B. oleae* was successfully transformed with the use of a *Minos*-based transposon [44]. Transformation efforts recently led to the development of *piggyBac*-based conditional female-lethal olive fly strains that provide highly penetrant female specific lethality, dominant fluorescent marking and genetic sterility [45]. Sixth, *B. oleae* was recently trans-infected with a cherry fly *Wolbachia* strain and shown to induce complete cytoplasmic incompatibility in the fly [46]. Finally, the experience gained during the first two pilot SIT efforts and the relevant research that followed, underlined a few key requirements for the maintenance of high quality and well-fit mass-reared olive flies (reviewed in [47]). Among them were changes in larval and adult diets (eg removal of antibiotics) that would preserve the endosymbiotic flora (that is now known to improve fitness) and occasional enrichments of the long-term laboratory colonies with wild individuals (that provide natural vigor). These achievements have renewed the interest in using SIT for olive fly control. In fact, there is a large international effort led by the Joint Division of the Food and Agricultural Organization and the International Atomic Energy Agency



(FAO/IAEA) to develop a vigorous laboratory olive fly strain that could be used in such new SIT efforts.

Further scientific and technological developments, in addition to successful SIT applications in other insects, point to the direction olive fly research could go. Indeed, SIT has proven particularly effective in the medfly, the prototype Tephritid species where most genetic and molecular tools have been developed. One of the most active medfly research areas in recent years has been the development of the RIDL technology. RIDL (Release of Insects carrying a Dominant Lethal; [48,49]) is a variant of the conventional SIT, in which sterilization of the released insects is induced not by irradiation but by homozygosity for a dominant lethal gene. Mating with wild individuals results in offspring that are heterozygous for the lethal gene leading to the death of all progeny [50,51]. This dominant lethal gene can be placed under the control of an inducible early embryonic female promoter [51,52] that could achieve genetic sexing at a very early developmental stage. In this way, both genetic sexing and sterilization can be accomplished by the same construct. One other active research area regards the analysis of biological systems with relevance to SIT. Of particular interest are those that regard reproduction and olfaction. The first one is involved in successful mating and egg development, while the second in food and mate localization. A possible manipulation of either or both of these systems would severely affect the destructive ability of the flies. In that sense, transgenic flies could be developed in which genes regulating food and mate recognition or fertility are knocked-down, over-expressed or mis-expressed (depending on the case). Such flies would be safer and more efficient to be released in control programs in an SIT context.

The falling prices of next generation sequencing make it now possible to sequence the entire transcriptome of non-model organisms under different settings and identify differentially expressed genes relevant to the chosen conditions. Subsequently, these genes can be manipulated *in vitro* and re-introduced into the genome of the organism through well-established transgenic technologies. In a first attempt to explore the relevant-to-SIT transcriptome of the olive fly, we present differences observed in female and male reproductive systems and we examine the differential expression of olfactory genes in the same tissues. Finally, we assess the developmental expression of two of the most commonly used early embryonic genes.

## Results and discussion

### 1. Sequencing and annotation

#### 1.1. Solid ABI sequencing and reads assembly

In order to explore differentially expressed genes in the transcriptome of reproductive organs of the olive fly that could be useful in SIT development, the entire

transcriptomes from female accessory glands and spermathecae were compared to male testes. For transcriptome assembly, the sequences from these two libraries (FEMALE and MALE) were combined with two more obtained from heads of spinosad-sensitive (LAB) and spinosad-resistant (SPIN) olive flies [53]. Paired-end sequencing with 35nt and 50nt read sizes was performed for each library and a total of 122,623,894 read pairs was obtained. All reads of the libraries were pooled to obtain a reference transcriptome assembly using SOAPdenovo assembler [54].

#### 1.2. Sequence annotation

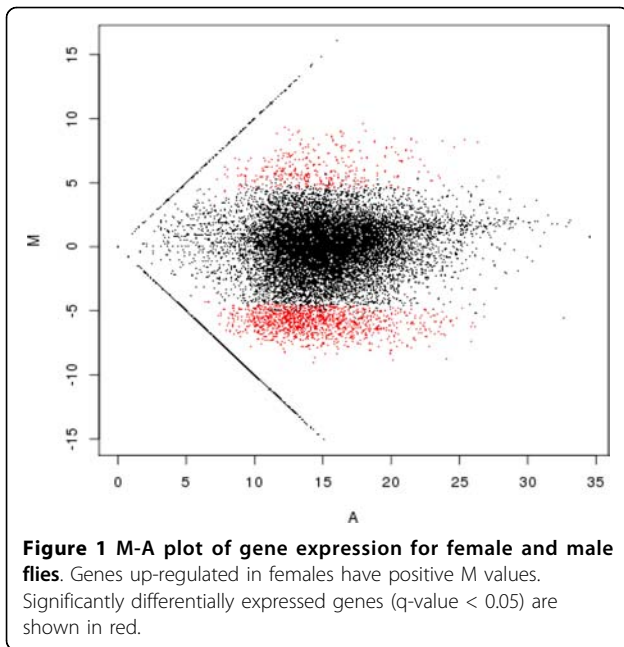
Annotation of the assembled sequences was obtained by aligning the 69,359 assembled *B. oleae* sequences against the NCBI non-redundant (Nr) protein database using blastx and collecting the annotations with the BLAST2GO tool [55]. Using an E-value threshold of  $\leq 1e^{-6}$ , 20207 (29.13 %) of the contigs were aligned. Of the 69,359 contigs, 23,042 (33.22%) have almost exact hits in the *B. oleae* transcriptome of Pavlidi et al [43] (E-value  $\leq 1e^{-6}$ ).

### 2. Female vs male differential expression

The Cuffdiff [56] tool was used in order to reveal the differentially expressed genes between the reproductive systems of female and male flies, a stringent cutoff (p value adjusted for multiple testing, called q value  $< 0.05$ ) was used. This resulted in 1568 differentially expressed transcripts in the FEMALE vs. MALE comparison. Three hundred and thirty of these transcripts were up-regulated in FEMALE, while 1238 were up-regulated in MALE *B. oleae* flies. The top 40 up-regulated genes in each category are listed in Table S1. The entire lists of all significantly ( $q < 0.05$ ) up-regulated genes in FEMALE and MALE are given in Tables S3 and S4, respectively.

An M-A plot was constructed for comparison of the genes for FEMALE vs MALE flies with q value  $< 0.05$ . In Figure 1 the de-regulated genes are depicted in red.

Functional annotation was made for the assembled sequences of the significantly differentially expressed female- and male- specific genes mentioned in Table S1, based on gene ontology (GO) categorization obtained using BLAST2GO. The FEMALE and MALE GO analysis performed for biological process of the top 40 female and male expressed genes is shown in Figure 2. In general, more GO terms appear in female tissues than in male (16 vs 12), a point that holds even in deeper GO-term analysis. This can be attributed to the fact that the FEMALE library was comprised of both FAGs and spermathecae, while the MALE from testes only. Furthermore, there were more male- than female-specific genes involved in metabolism and development, a fact that can be attributed to sperm activity in the MALE tissue. Finally, the presence of three immune system process genes in the female list should be noted. In fact, increased

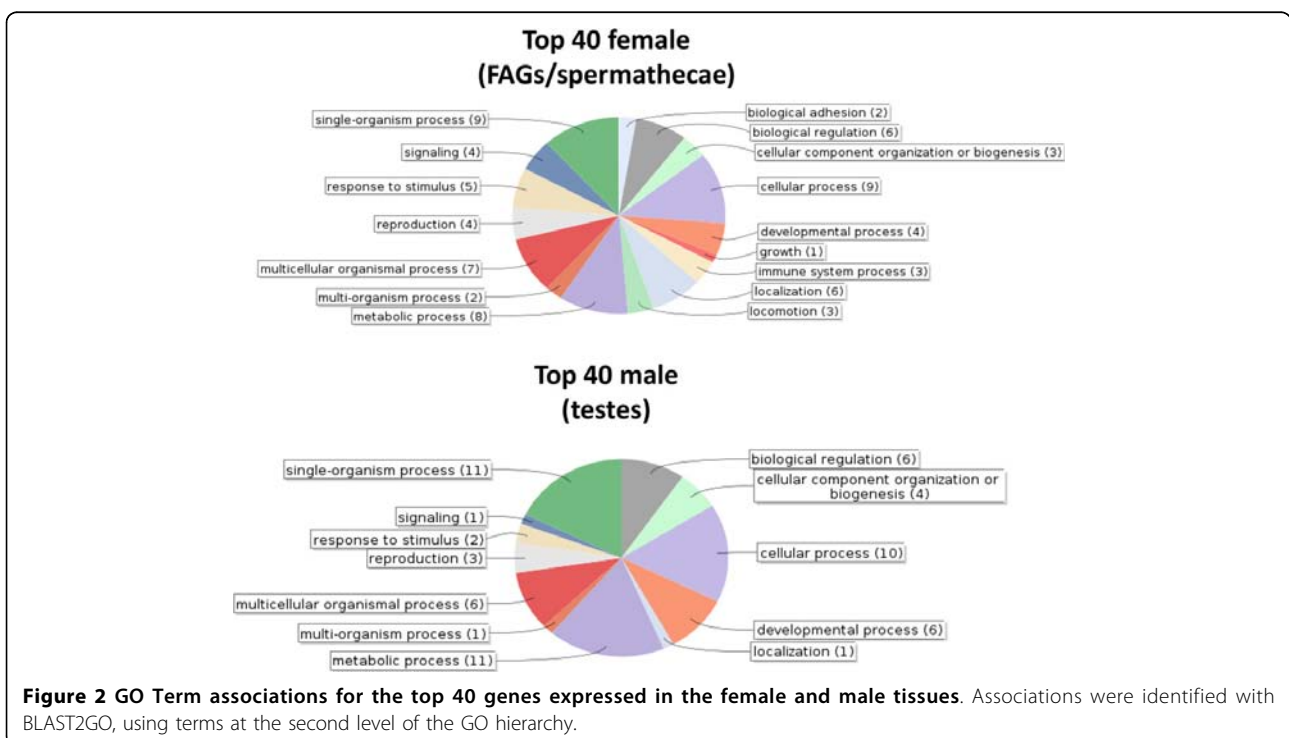


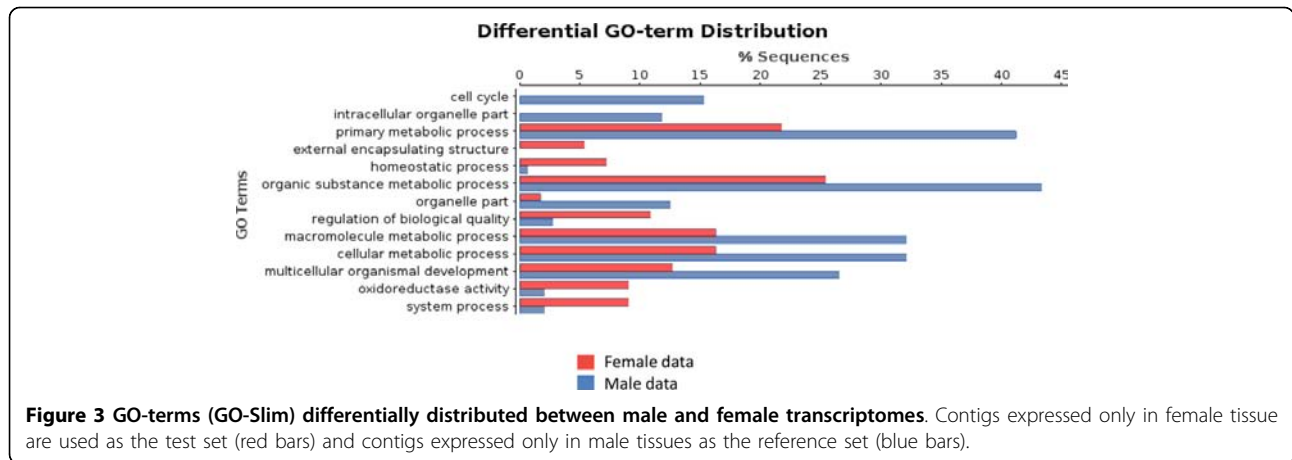
levels of immune response genes have been found in transcriptome analyses of insect female reproductive systems, particularly after mating [57,58]. Upregulation of these genes may assist females to combat pathogens introduced during copulation. Alternatively, it could be a result of female's perception of sperm as non-self molecules.

A more direct comparison between FEMALE-only and MALE-only GO-term distribution is shown in Figure 3. Interestingly, numbers of GO-terms for biological process appear different in the two datasets, suggesting a different complexity of the studied female and male reproductive tissues. In most terms, there are more male- than female-specific transcripts that are differentially expressed. Many of these terms (cell cycle, intracellular organelle part, primary metabolic process, organic substance metabolic process, macromolecule metabolic process, cellular metabolic process, multicellular organismal development) refer to higher metabolic processes. This could be attributed to higher metabolic and cellular activity that takes place in the testes before mating.

### 3. Genes that might be implicated in sexual differentiation in *B. oleae*

In order to validate the differential expression of various genes observed after the RNAseq analysis of reproductive tissues of female and male olive flies, further functional analysis was performed for twelve genes that were differentially expressed in female accessory glands and spermathecae, on one hand, and male testes, on the other (Figure 4). These genes were selected on the basis of known involvement in sexual differentiation in other insects. Seven of them were selected from the 1238 significantly up-regulated in MALE (Table S4): *kl2* (male fertility factor *kl2*), *kl3* (male fertility factor *kl3*), *kl5*





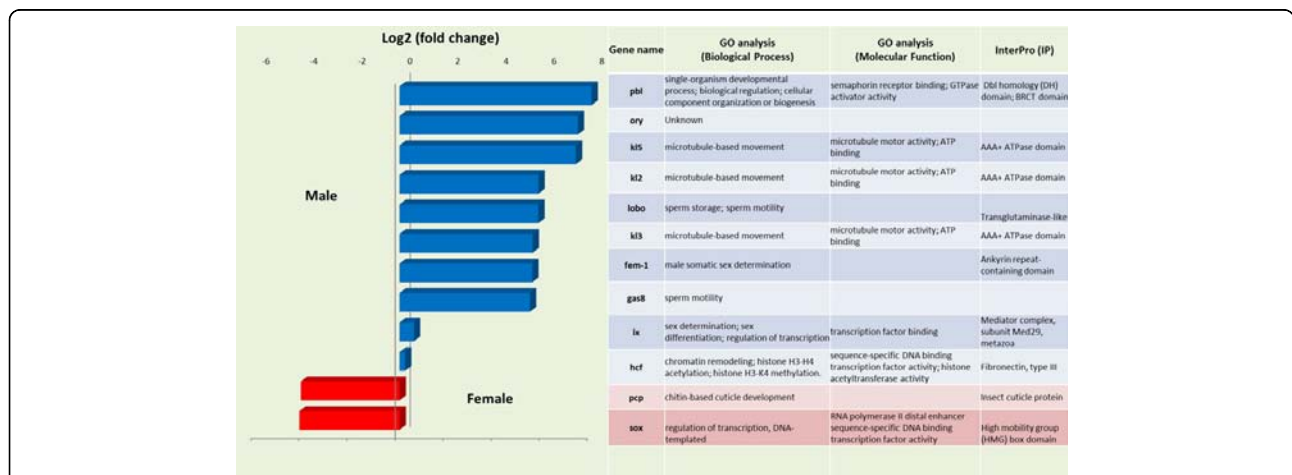
(male fertility factor *kl5*), *ory* (occludin-related Y protein), *fem-1* (sex-determining protein fem-1), *gas8* (growth arrest specific protein 8) and *lobo* (lost boys). Three more genes that were up-regulated in MALE [*ix* (*intersex*), *pbl* (*pebble*) and *hcf* (*host cell factor C1*)] and two that were up-regulated in FEMALE [*sox* and *pcp* (*pupal cuticle protein 78E*)], albeit with lower statistical power (i.e.,  $q > 0.05$ ) were also selected for further validation.

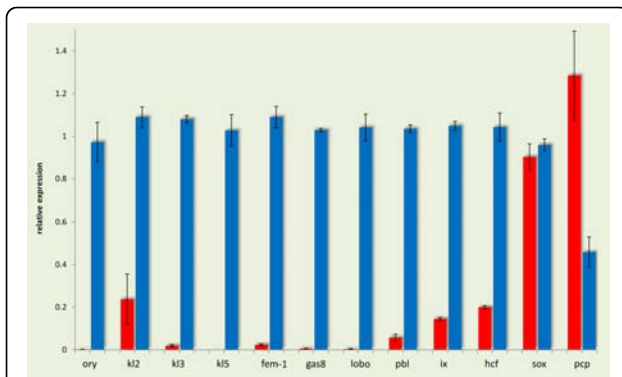
### 3.1. Drosophila Y-linked genes *kl3*, *kl5* and *ory*

Quantitative RT-PCR confirmed the elevated expression of *kl2*, *kl3*, *kl5* and *ory* in male testes of the olive fly (Figure 5). In *Drosophila melanogaster*, *kl3* and *kl5* (along with *kl2*) are known Y-linked fertility factors. The lack of *kl3* or *kl5* causes the loss of the outer arm of the sperm tail axoneme [59], a structure known to contain the molecular motor protein dynein in other organisms [60].

Indeed, Goldstein et al. showed in 1982 that sperm from mutant *kl3<sup>-</sup>* and *kl5<sup>-</sup>* males lack three discrete high molecular weight proteins with mobility similar to dynein heavy chains of *Chlamydomonas reinhardtii* and proposed that these fertility factors are the structural genes of three different dynein heavy chain proteins [61]. In 1993, Gepner and Hays sequenced part of *kl5* and showed that it encodes an axonemal  $\beta$ -dynein heavy chain that is expressed in the testes [62].

*ory* is also Y-linked in *D. melanogaster*, although details on this gene are scarce. *kl3*, *kl5* and *ory* are Y-linked in 12 different sequenced Drosophila genomes [63]. In Drosophila, the closest paralogs of *kl2*, *kl3*, and *kl5* are autosomal and not X-linked, suggesting that the evolution of the Drosophila Y chromosome has been driven by an accumulation of male-related genes arising *de novo* from the autosomes [64]. While the most likely function of the





**Figure 5 Validation profiles of differentially expressed sex-differentiation genes.** Differentially expressed sex-differentiation genes of Figure 4 were further validated by qRT-PCR. Expression in male testes is depicted in blue color columns and expression in female accessory glands and spermathecae in red. Standard error of the mean of the two biological replicates is shown in bars. In all genes, except *sox* and *pcp*, expression in FEMALE and MALE was significantly different, as determined by t-test ( $p < 0.05$ ).

three genes in the olive fly might be similar to that of *Drosophila*, we have no indication with regard to their chromosomal localization in the olive fly. Such information could shed some light to the evolutionary origin of the olive fly's Y chromosome.

### 3.2. Spermatogenesis and sperm motility genes

One spermatogenesis and two sperm motility genes were shown to be differentially over-expressed in male olive fly tissues both in the transcriptome analysis and after q-RT PCR (Figure 4 and 5). The first locus, **sex-determining protein fem-1 (*fem-1*)**, encodes an essential spermatogenesis product in *Caenorhabditis elegans*. Three *fem* genes, *fem-1*, *fem-2*, and *fem-3*, have been shown to be essential for male development [65]. Loss-of-function mutations in any one of the *fem* genes prevent all aspects of male development and transform the animals that are genetically males into females [66,67]. The predicted product of the *fem-1* gene is an intracellular protein that contains ankyrin repeats, which in many other proteins mediate specific protein-protein interaction [67]. In *D. melanogaster*, a *fem-1* homolog with similar structure has been found [68]. The second locus, **growth arrest-specific protein 8 (*Gas8*)** is a microtubule-binding protein localized to regions of dynein regulation in mammalian cells. In mouse, *Gas8* is predominantly a testicular protein, whose expression is developmentally regulated during puberty and spermatogenesis. In humans, it is absent in infertile males who lack the ability to generate gametes [69]. *Gas8* has not been studied in insects. Finally, **lost boys (*lobo*)**, has been shown to affect sperm entry movement into the female seminal receptacle and does not affect sperm exit movement from the seminal vesicle of *D. melanogaster*

[70]. Given a similar function of these two loci in the olive fly, over-expression in male testes is expected.

### 3.3. Sex determination genes

In *D. melanogaster*, **intersex (*ix*)** controls somatic sexual differentiation only in females, acting near the end of the sex determination hierarchy. Its product does not have a known DNA-binding domain and, therefore, it is thought to act as a transcriptional co-factor for the female variant of Doublesex protein (*DSX<sup>F</sup>*), a key gene of the sexual determination cascade in *D. melanogaster* [71]. Minimal differences were observed in *ix* expression between the two sexes of the olive flies.

Transcriptome analysis also showed a four-fold over-expression of *sox* in female tissues, a result that was not confirmed after validation. The *sox* gene family is a group of related transcription factors that play critical roles in embryonic development. This family was originally identified in mammals based on sequence similarity to SRY, the sex-determining region Y chromosome [72]. In the honeybee, as SOX proteins play key roles in gonad differentiation, the SoxE group orthologues were up-regulated in the drone testes [73]. In *Drosophila* SoxN is a new group B Sox gene expressed in the developing CNS and is one of the earliest transcription factors to be expressed in a pan-neuroectodermal manner [74].

### 3.4. Other genes

The **Pebble (*pbl*)** gene belongs to a family of GTP exchange factors that are essential for the construction of a contractile ring and the initiation of cytokinesis during the embryonic division cycles of the somatic cells in *D. melanogaster* [75,76]. Its role in spermatogenesis has not been elucidated yet. Expression of *pbl* in *D. melanogaster* testes is low [68]. On the other hand, expression in olive fly testes was found elevated in comparison to its expression in female accessory glands/spermathecae (Figure 4 and 5).

**Host cell factor C1 (*Hcf*)** is involved in a wide variety of cellular functions, including regulation of transcription, cytokinesis, cell cycle progression and chromatin remodeling [77]. The protein is essential for cellular viability and demonstrates similar activity among a broad range of species. A single *hcf* homolog is also present in *Drosophila* (called dHCF) and is expressed in all tissues, although at relatively low levels [68]. The transcriptome analysis in the olive fly tissues showed a ~0,2-fold higher expression in the male tissues. This result was confirmed after qRT-PCR in the same tissues, where higher levels of expression in testes were observed in comparison with female accessory glands/spermathecae (Figure 4 and 5).

Quantitation by RT-PCR confirmed the over-expression of **pupal cuticle protein (*pcp*)** in female accessory glands/spermathecae as compared to male testes. Cuticle proteins, along with chitin, are the two components of insect cuticle. The cuticular proteins seem to be specific

to the type of cuticle that occurs at stages of the insect development. Flexible proteins are found in the flexible cuticle of larva and pupa, but can also be found in the soft endocuticle of adult insects [78].

Female insects require the steroid hormone 20-hydroxyecdysone (20E) in order to activate vitellogenesis, a process required for egg development. In *Anopheles gambiae* mosquitoes, large amounts of 20E are produced and stored in male accessory glands and subsequently delivered to female mosquitoes during mating [79]. Pupal cuticle proteins, on the other hand, are known to accumulate in response to a pulse of 20E [80]. However, given that FAGs/spermathecae collected were from unmated females, we cannot offer a plausible explanation for the over-expression of *pcps*.

#### 4. Validation of olfactory gene differential expression

Insects possess very sensitive chemosensory systems that can detect and discriminate among a diverse array of odors. These systems play a crucial role in insect survival and reproductive success, mediating responses to food detection, mating and oviposition. Odor recognition is a coordinated process requiring the combined specificities contributed by odorant-binding proteins (OBPs) and chemosensory proteins (CSPs) as well as odorant receptors (ORs) (Reviewed in [81]). Insect odorant-binding proteins (OBPs) are soluble proteins surrounding the extracellular lymph of olfactory neurons [82]. OBPs are capable of binding and solubilizing small hydrophobic molecules from the environment and therefore transport them to the underlying ORs, which are expressed on peripheral olfactory receptor neurons. Insect ORs are either ionotropic receptors (IRs) or seven-transmembrane proteins (ORs) with an inverse topology compared to GPCRs, that form heterodimers of a ligand-binding OR and an ubiquitous highly conserved co-receptor named Orco [83]. These complexes are suggested to constitute ligand-gated nonselective cation channels triggering the olfactory signaling [81].

While OR expression in olfactory tissues is obvious and well-established, the distribution of ORs beyond the olfactory system has also been documented in different mammalian species [84-86], suggesting that ORs may play an important role in the ectopic expression of non-chemosensory tissues. Interestingly, OR expression has been documented in human and mouse germ cells [87-91] and recently in mosquitoes [92]. Similarly, other non-olfactory functions have been reported for OBP-like proteins including the B proteins of *Tenebrio molitor* accessory glands [93], the male specific serum proteins of *Ceratitis capitata* [94], and the heme-binding protein of *Rhodnius prolixus* [95]. These demonstrate that OBPs are not restricted to olfaction and are likely to be involved in broader physiological functions, suggesting that their

roles may be restricted to general carrier capabilities with broad specificity for lipophilic compounds [96].

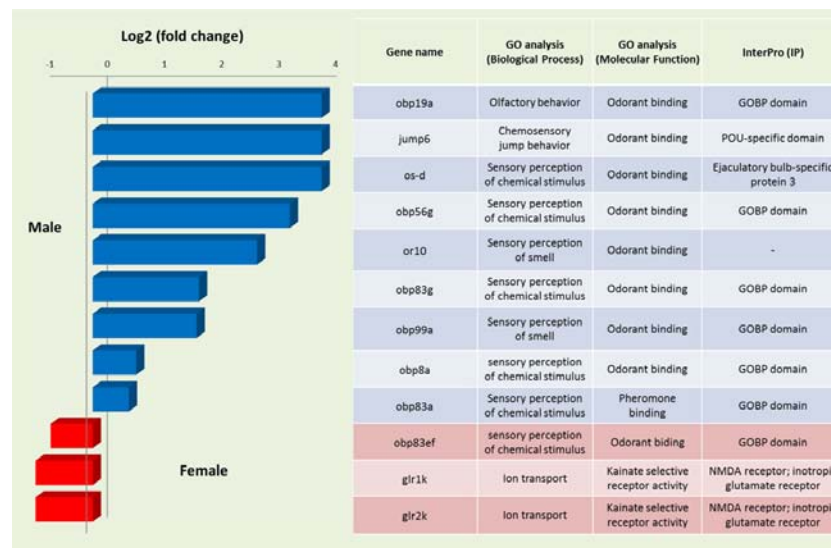
With that in mind, we opted to explore the expression of various olfactory-related genes in the reproductive systems under investigation. Twelve olfactory-related genes were present in the annotated list that resulted from the transcriptome assembly of the FEMALE and MALE olive fly tissues (Figure 6), nine of which presented various levels of over-expression in MALE, whereas the remaining three in FEMALE. In order to get a deeper insight, the relative expression of five of these genes was further analyzed in female FAGs/spermathecae, male testes and male accessory glands (MAGs), before and after mating.

*obp83a*, *obp8a* and *obp19a* genes are over-expressed in MALE tissue (Figure 6). qRT-PCR revealed that these genes share the same expression pattern in MAGs. *obp83a* and *obp8a* are over-expressed before mating in testes while *obp83a* and *obp19a* are over-expressed after mating in FAGs/spermathecae (Figure 7). All three genes are characterized by a GOBP (general odorant binding protein) domain that is also found in their orthologues in *Drosophila melanogaster*. This structural domain is found in pheromone binding proteins, which exist in extracellular fluid surrounding odorant receptors [97]. The presence of these OBPs in the reproductive tissues implicates their interaction with other substrates except the olfactory system as transporters in the post-mating events in the male reproductive system. In fact, *D. melanogaster*'s *obp8a* shows the highest levels of expression in male accessory glands [98,99] and has been associated with non-olfactory functions such as RNA transcription [100].

*os-d* is over-expressed in MALE tissue (Figure 6) while qRT-PCR showed similar expression patterns in mature FAGs/spermathecae, MAGs and testes, but no expression in MAGs before mating (Figure 7). Os-D is a chemosensory protein (CSP) that encodes the antennal protein 10 in *D. melanogaster*. CSPs are secreted in the sensillum lymph of insect chemosensory sensilla and some OS-D-like proteins bind short to medium chain length fatty acid derivatives with low specificity [101,102]. Their specific function remains uncertain [103], suggesting a more general physiological function relating to the transport/solubility of hydrophobic ligands in various tissues.

*or10* showed expression in male tissues (Figure 6) while qRT-PCR detected same transcriptional profiles in all three tissues before and after mating (Figure 7). *or10* encodes an olfactory receptor protein and has a G-protein coupled receptor activity. The expression of ORs in testes has been reported for a number of species [90,104]. ORs' function in mammalian sperm is thought to regulate motility in response to exogenous signals derived from the existence of sperm-egg chemotaxis in invertebrates. The small peptides, speract and resact, are secreted by sea urchin eggs and attract spermatozoa in a species-specific





**Figure 6 Functional annotation of differentially expressed olfactory genes.** At the left part of the figure, the expression levels of the differentially expressed olfactory genes (Log2, fold change) are shown, as resulted from the RNA-seq analysis. The up-regulated genes in males are depicted in blue bars and the up-regulated genes in females in red bars. At the right part of the figure, the Gene Ontology (GO) classification of the same genes for the ontologies: Biological Process (BP), Molecular Function (MF) and Interpro (IP) protein domains is listed. Gene names are based on the nomenclature of the *Drosophila melanogaster* homologues [68].

manner by stimulating sperm motility and respiration [105,106]. The presence of a similar chemoreceptor may be essential in female spermatheca in order to establish a concentration gradient of a putative chemo-attractant. Since female accessory glands and spermatheca were dissected together, we are not able at this point to establish which exact tissue is the source of the observed expression of *or10*.

### 5. Early embryonic gene expression in the olive fly

As mentioned in the Background, promoters of early embryonic genes in combination with pro-apoptotic cell death genes are very important tools in inducing dominant early-embryonic lethality during insect transgenesis [107]. In that regard, the *serendipity-α* (*sry-α*) and *head involution defective* (*hid*) genes were selected for expression evaluation during embryonic development in the olive fly.

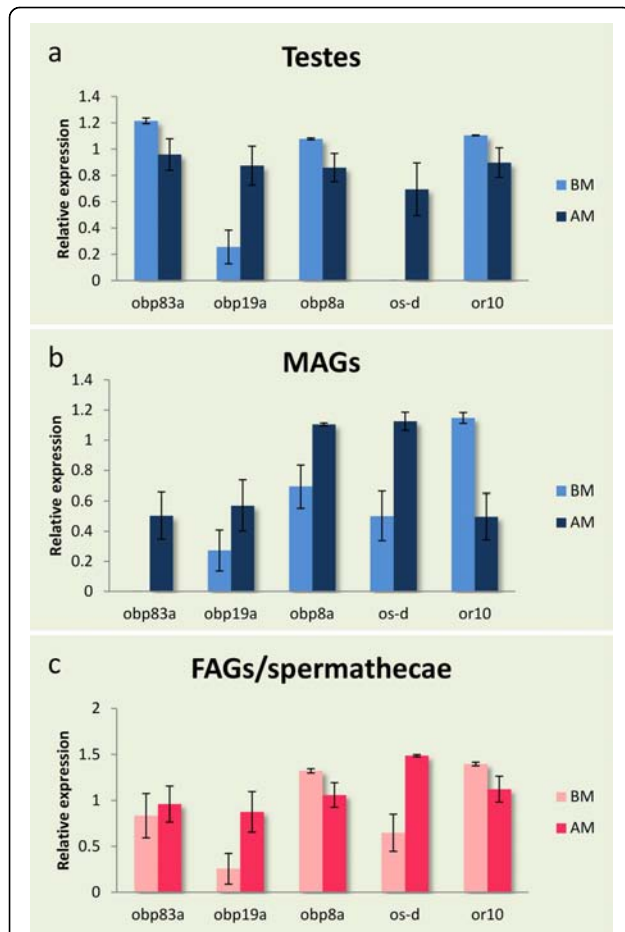
The embryonic developmental progress begins with the egg maturation and formation of the zygote, then enters the stage of blastoderm formation and gastrulation and ultimately ends with the organogenesis. Accordingly, three stages of embryogenesis have been also designated in *B. oleae*, whose average duration is 65-70h at 25 ± 1°C under standard laboratory conditions [108]. Microscopy morphological observations in living embryos report that cellularization of the blastoderm begins 6h after oviposition and lasts until 10h. During the third stage of organogenesis, the ventral furrow formation starts by 22h and the head and abdominal lobe masses become visible by 46h. Gut and mouth hook formation can be identified by

52h, whereas the development of other systems are distinct by 60h.

In *Drosophila melanogaster*, *sry-α* gene is specifically transcribed at the blastoderm stage in all somatic nuclei, from nuclear cycle 11 to the onset of gastrulation [109]. The gene product is required for the complete reorganization of the microfilaments at the onset of membrane invagination [110]. *sry-α* is fast evolving even within the Drosophilidae [111] and extensive divergence of many developmental genes within dipterans has also been reported [112-114]. This was most likely the reason for the unsuccessful efforts in *C. capitata* to obtain *sry-α* by degenerate PCR on the basis of sequence similarity with the homologous *D. melanogaster* [115]. Given the availability of both *D. melanogaster* and *C. capitata* *sry-α* sequences in the NCBI database, a homology search in the *B. oleae* transcriptome identified the relevant *B. oleae* *sry-α* gene homologue.

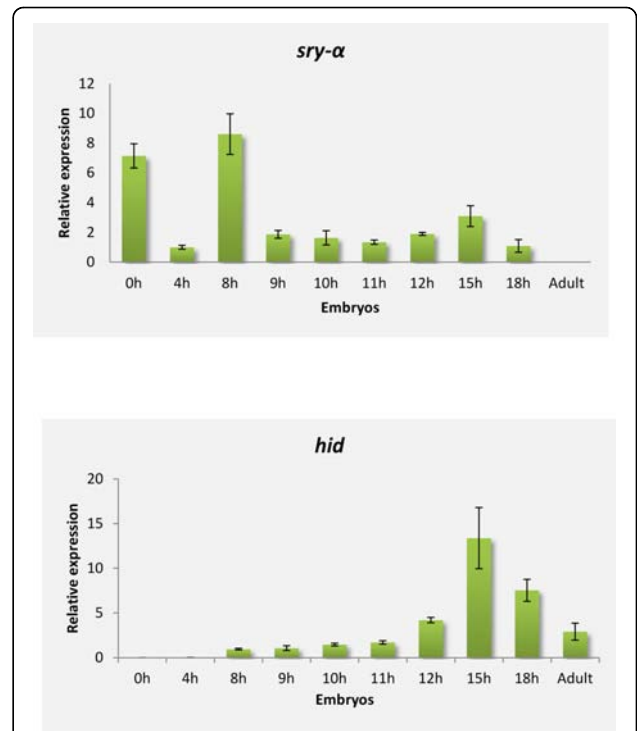
Based on this sequence, *B. oleae*-specific primers were designed and the expression profiles of *sry-α* mRNA were studied by qRT-PCR analysis at different stages of *B. oleae* embryonic development. Eggs were collected throughout embryogenesis from the time of egg laying to larval hatching. The selected time points represented embryos at 0h, 4h, 8h, 9h, 10h, 11h, 12h, 15h and 18h after oviposition (Figure 8, panel A). This analysis revealed that *sry-α* mRNA is developmentally regulated during the second major event in the first stage of embryogenesis. It is initially present in large amounts just after oviposition (0h embryos), following a reduction in 4h embryos. The larger





**Figure 7** Relative expression profiles of differentially expressed olfactory genes. Expression profiles of five olfactory genes [odorant binding proteins *obp83a*, *obp19a*, *obp8a*, chemosensory protein, *os-d*, and odorant receptor 10, *or10*] as determined by qRT-PCR in three different tissues: Testes (a), MAGs (b) and FAGs/spermatheca (c) before (BM) and after (AM) mating. Standard error of the mean of five biological replicates is depicted in bars. No significant difference (for  $P < 0.05$ ) was detected.

amounts of the transcripts among all time points examined were detected in 8h embryos. This suggests the presence of maternal mature transcripts which in turn are eliminated probably in the first event of maternal-to-zygotic transition (MZT). The subsequent wave of ‘zygotic’ activity requires zygotically synthesized transcripts [116]. In *D. melanogaster* as well as in *C. capitata*, *sry-α* is expressed only in the zygote [117]. However the retrieved *B. oleae* transcript shared greater amino acid similarity to the *D. melanogaster* CG8247 gene than to *sry-α*, as was also reported for the *Csry-α like* gene [118]. The orthologous CG8247 in *D. melanogaster* is characterized as a *sry-α*-like gene being also involved in cellular blastoderm formation. However, it is maternally inherited in contrast to *sry-α*, demonstrating a different mechanism of molecular control of transcription. In our case *Bosry-α like* gene



**Figure 8** Expression profile analysis during the early stages of embryogenesis. Expression levels of A) *Bosry-α* and B) *Bohid* in individual eggs collected at different time points during embryonic development, as determined by qRT-PCR. Standard error of the mean of two biological replicates per time point is depicted in bars.

seems to be maternally supplied in the embryos as mature transcripts. Previous studies have designated that the cellular blastoderm formation in *C. capitata* occurs within 9 h and 11 h after oviposition [115]. In accordance with *C. capitata*, a relative Tephritid species, we suggest that the cellularization process in *B. oleae* during embryogenesis also occurs at 8h, since the *sry-α* transcripts were detected at higher levels during this time.

### 6. Apoptotic gene expression

At the same time, *head involution defective (hid)*, known to have a central role in apoptosis pathway, was also selected for further study. Apoptosis is a genetically controlled mechanism of cytological events that results in programmed cell death. During development, programmed cell death plays a key role by eliminating unwanted cells from a variety of tissues, such as, for example, larval tissues during insect metamorphosis (Reviewed in [119]). A series of caspases, a family of cysteine proteases, play a central role during apoptosis. Once activated, caspases can cleave more than 100 different cell target proteins, bringing about ultimately the cell death [120]. Regulators of caspase activation may either promote apoptosis (pro-apoptotic) or inhibit

apoptosis (anti-apoptotic). *Drosophila Hid* belongs to a family of pro-apoptotic proteins which act as antagonists of IAPs (Inhibitor of Apoptosis Proteins), thus resulting in caspase activation and apoptosis [119,121,122]. Such pro-apoptotic genes have been used in transgenic control systems for pest insects. In tetracycline-suppressible systems for female-specific lethality and conditional embryonic expression of a *Drosophila hid*-containing transgene, for example, 100% lethality was observed in *Drosophila* [123], as well as in the Tephritid flies *Ceratitis capitata* [117] and *Anastrepha suspensa* [124].

The developmental regulation of *Bohid* was explored by determining the transcript levels during embryogenesis. A qRT-PCR approach with species-specific primers was used to evaluate the expression pattern of *hid* in embryos at 0 h, 4 h, 8 h, 9 h, 10 h, 11 h, 12 h, 15 h, 18 h after oviposition. Based on *D. melanogaster hid* expression pattern, no expression was expected in embryos prior to formation of the syncytial blastoderm [125]. Indeed, until 8h no transcripts were detected. *hid* expression was first detected at 12h and peaked during 15h (Figure 8, panel B).

It is noteworthy that most developmental programmed cell death occurs during the gastrulation process of *D. melanogaster* embryonic development [126], suggesting that the onset of this period in *B. oleae* could be defined approximately at 12h, occurring mainly within 15-18h.

However, further examination of the pro-apoptotic function of *hid* gene is required in order to explore its ability of inducing apoptosis in *B. oleae* cells. Specific lethal embryonic phenotypes need to be obtained to characterize its role in the cell-death pathway. Ongoing analysis for the isolation of the complete gene will provide the essential tools for the generation of an endogenous effective lethal effector system.

## Conclusions

In serious agricultural pests (like the olive fly) which are not model experimental organisms (unlike the medfly), the major focus of most scientific research is, in the end, directed towards control of the pest. Old and new environmental concerns and sensibilities, that regard mostly insecticide use, drive science to the quest of alternative, environmentally friendlier methods of pest control. Time and again it has been shown that such methods go through thorough understanding of the biology and ecology of the target organism. Since the initial unsuccessful SIT efforts, molecular and genetic studies in the olive fly have focused on genetic analyses of natural populations, cytogenetics, isolation and characterization of genes that control important biological processes, as well as the identification and mapping of several microsatellite loci. Just a few years ago, *B. oleae* was successfully transformed, an

achievement that gave new perspective towards the efficient use of the SIT. Lately, this is being coupled with genomics studies and transcriptomics analyses of various important systems, as well as efforts in advancing olive fly mass-rearing, that are setting the ground for the application of modern control approaches through the genetic manipulation of the insect.

## Methods

### Ethics statement

The study was carried out on laboratory reared olive flies. No specific permissions are required for these experiments, since these studies did not involve endangered or protected species.

### Fly culture and stocks

#### Laboratory strain

The laboratory strain of the olive fly (LAB) is part from the original stock from the Department of Biology, 'Demokritos' Nuclear Research Centre, Athens, Greece, and has been reared in our laboratory for over 15 years. The flies are reared at 25°C with a 12h light/12h dark photoperiod in 30x30x30cm<sup>3</sup> cages, as described by [127-129].

#### Egg collection

For embryo analysis, eggs were collected from 10-day old mated females maintained in our laboratory, which were fed with artificial adult diet to ensure high oviposition rates and embryo viability. Adults were exposed to paraffin oviposition domes for 10 minutes and the eggs were obtained with a 0.3% propionic acid solution, assigning this as the start time point. Eggs were maintained in an incubator according to the standard rearing conditions.

### RNA isolation for library preparation and functional analysis

Total RNA was isolated from female accessory glands (FAGs) and spermathecae of ~300 female flies and from testes of ~150 male flies. Four-day old sexually immature unmated insects were used. For RNA isolation, the TRIzol<sup>®</sup> Reagent (Ambion-Invitrogen) was used, following the instructions of the manufacturer with minor modifications. RNA extraction was followed by an additional DNA removal using the TURBO DNA-free Kit (Ambion-Invitrogen), according to manufacturer's instructions. The integrity of RNA was assessed by 1% agarose gel electrophoresis and the purity of all RNA samples was evaluated at Fleming Institute (Greece) with the use of (Agilent 2100 Bioanalyzer) and NanoDrop (2000).

### Whole transcriptome library preparation for next-generation sequencing with the SOLiD 4 Sequencing System

RNA transcripts from olive fly FAGs/spermathecae (FEMALE) and testes (MALE) were used to construct

two cDNA libraries for sequencing analysis on the SOLiD 4 Sequencing System. More specifically, polyadenylated RNA (polyA-RNA) was isolated from 5 µg of total RNA using the Dynabeads Oligo(dT) kit (Ambion, Life Technologies Corporation). The isolated polyA-RNA was randomly fragmented by chemical hydrolysis at 94°C for 5 minutes and was then treated with antarctic phosphatase to remove phosphate groups from the fragments' ends, followed by treatment with T4 polynucleotide kinase to add a Pi at the 5' end of each fragment. The resulting RNA fragments were hybridized and ligated to the P1 and P2 adaptor sequences specifically designed for sequencing with the SOLiD system (SOLiD Total RNA-Seq Kit, Life Technologies Corporation). The RNA produced was reverse transcribed to cDNA which was then amplified in a 15-cycle PCR. At this step, the use of different barcoded 3' PCR primers from the selection included in the SOLiD barcoding kit allowed the preparation of cDNA libraries for multiplex sequencing. From the cDNA produced, only fragments of average size 200-300 bp were selected with two rounds of magnetic bead purification (Agencourt AMPure XP Reagent, Beckman Coulter).

The quality and size of the purified cDNA library was assessed on the Agilent Bioanalyzer 2100 (Agilent Technologies Inc.) and with quantitative PCR using the Library Quant Kit ABI Solid (KAPA Biosystems). A multiplex library mix (500pM) was used to prepare a full-slide for analysis on the SOLiD 4 Sequencing System (Applied Biosystems) with 35+50 bp PE-chemistry.

#### RNA isolation and expression analysis of selected genes

*RNA extraction for expression analysis of sexually differentially expressed genes.* For the validation of the differential expression of sexually differentially expressed genes, RNA was extracted from two pools of 40 pairs of spermathecae/FAGs and 40 pairs of testes (two biological pool replicates), dissected from an equivalent number of female and male adult laboratory flies, respectively.

*RNA extraction for expression analysis of olfactory and early embryonic developmental genes.* For the validation of the olfactory genes expression, RNA was extracted from five female and five male individual insects (five biological replicates, respectively) before and after mating of the aforementioned laboratory strain. Two groups of insects were considered. Firstly, unmated insects, i.e., sexually mature 7-day old unmated insects (before mating, BM). Secondly, mated insects, i.e., sexually mature 7-day old insects that were allowed to mate on the seventh day and were dissected 12 hours after mating (after mating, AM). For the validation of the sexually differentially expressed genes, the RNA isolated for the construction of the two libraries was used. RNA was extracted using TriZol reagent according to manufacturer's protocol.

For the validation of the early embryonic genes, eggs were removed from the incubator at different time intervals throughout embryonic development and total RNA was extracted from each egg using TriZol reagent according to the manufacturer's protocol. Two individual eggs (two biological replicates) from the various time points during the embryonic developmental stages were used for the extractions.

Following extraction, the RNA was treated with 1.0 unit of DNase I (Invitrogen) according to manufacturer's instructions. In all of the above cases, the total amount of DNA-free RNA obtained from each tissue (between 400 to 700 ng) was converted into cDNA using 300ng Random hexamer primers (equimolar mix of N<sub>5</sub>A, N<sub>5</sub>G, N<sub>5</sub>C and N<sub>5</sub>T), 200 units MMLV Reverse Transcriptase (Geneon), 5X reaction buffer, 40mM dNTP mix and 40 units RNase Inhibitor (GeneOn) according to the manufacturer's instructions. Reverse transcription was conducted at 42°C for 50 min and 70°C for 15 min. The resulting cDNA was used in the subsequent qPCR reactions.

Specific primers for the amplification of selected differentially expressed genes revealed by the transcriptome analysis were designed by Primer-BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast>) (Table S2). To identify sequences with homology to the genes *sry-α* and *hid*, the orthologous genes of *C. capitata* and *An. suspensa* were used as queries to search for *B. oleae* transcripts using tBLASTX in the TSA Database. Species-specific Blast hits for each of the query sequences were retrieved (Genbank: GAKB01005111.1, GAKB01003654.1) and used to design primers (Table S2) for the subsequent amplification of gene-specific sequences by quantitative real-time PCR (qRT-PCR).

Relative quantitation was used to analyze changes in expression levels of the selected genes using a Real-time PCR approach. Expression values were calculated relatively to the housekeeping *rpl19* gene. *Rpl19* and *14-3-3z* genes were used as reference in MAGs and testes while *actin3* and *a-tubulin* in FAGs/spermathecae. The qRT-PCR conditions were: polymerase activation and DNA denaturation step at 95 °C for 4 min, followed by 40 cycles of denaturation at 95 °C for 30 s, annealing/extension and plate read at 56 °C for 30 s and finally, a step of melting curve analysis at a gradual increase of temperature over the range 55 °C → 95 °C. In this step, the detection of one gene specific peak and the absence of primer dimer peaks was assured. Each reaction was performed in a total volume of 15 µl, containing 5 µl from a dilution 1:10 of the cDNA template, 1X iTaq Universal SYBR Green Supermix (Biorad, Gaithersburg, MD) and 400nM of each primer. The reactions were carried out on Bio-Rad Real-Time thermal cycler CFX96 (Bio-Rad, Hercules, CA, USA) and data analysed using the CFX Manager™ software. All qRT-PCRs were performed in triplicate (i.e., three technical replicates).

## Bioinformatics analysis

All paired and unpaired reads of the libraries were assembled to construct the reference transcriptome using the SOAPdenovo assembler [54] with a word size of 25 nt. Annotation of the assembled sequences was obtained by comparing to the NCBI non-redundant (Nr) protein database (May 7<sup>th</sup>, 2014 version) using blastx [130] and collecting the annotations with the BLAST2GO tool [55]. TopHat [131] was used to generate a spliced alignment to the reference transcriptome. Transcripts were assembled using Cufflinks and differentially expressed genes were identified using Cuffdiff [56]. GO-term enrichment between male and female transcriptomes was analyzed using the using the GOSSIP [132] application embedded in BLAST2GO.

## Availability of supporting data

The data sets supporting the results of this article are included within the article and its additional files. Additional File 1, Additional File 2, Additional File 3 and Additional File 4

## Additional material

Additional File 1:

Additional File 1:

Additional File 1:

Additional File 1:

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

ES was involved in the transcriptome library construction and performed the analysis of the sex-determination genes; MR performed the bioinformatics analysis of the transcriptome; VH constructed the transcriptome libraries and analysed the sequencing data; KTT and AMM analyzed the embryonic and apoptotic genes; MEG, ST and KA analysed the olfactory genes; JR directed the bioinformatics analysis; KDM designed and coordinated the study. All authors participated in drafting the manuscript and read and approved the final document.

## Acknowledgements

This research has been co-financed by: the Actions *Heracleitus II* and *ARISTEIA* ("OLFLY SMELL & SEX") of the "Operational programme Education and Life Long Learning", co-funded by the European Social Fund and Greek National Resources"; and the two postgraduate programs of the Department of Biochemistry and Biotechnology of the University of Thessaly ("Biotechnology - Nutrition and Environment" and "Molecular Biology and Genetics applications"). Special acknowledgements should also go the Joint FAO/IAEA Division of Nuclear Techniques in Food and Agriculture for their support in the organization of a Coordinated Research Project on "Development and evaluation of imported strains of insect pests for SIT. This article has been published as part of *BMC Genetics* Volume 15 Supplement 2, 2014: Development and evaluation of improved strains of insect pests for SIT. The full contents of the supplement are available online at <http://www.biomedcentral.com/bmcgenet/supplements/15/S2>. Publication of this supplement was funded by the International Atomic Energy Agency. The peer review process for articles published in this supplement was overseen by the Supplement Editors in accordance with

BioMed Central's peer review guidelines for supplements. The Supplement Editors declare that they have no competing interests.

## Authors' details

<sup>1</sup>Department of Biochemistry and Biotechnology, University of Thessaly, Larissa, Greece. <sup>2</sup>Institute of Molecular Biology and Genetics, Biomedical Sciences Research Centre "Alexander Fleming", Greece.

Published: 1 December 2014

## References

1. Theophrastus: *Enquiry into plants (History of plants HP), I & II (HORT, A. F., translator)*. London, Cambridge & Massachusetts: 1916, [in ancient Greek with English translation].
2. Daane KM, Johnson MW: **Olive fruit fly: managing an ancient pest in modern times.** *Annu Rev Entomol* 2010, **55**:151-69.
3. Pimentel D: **Ecological Effects of Pesticides on Non-target Species.** Washington, D.C.: Executive Office of the President, Office of Science and Technology; 1971, 220.
4. Baumhover A, Graham A, Bitter B, Hopkins D, New W, Dudleyandr F, Bushland C: **Screwworm control through release of sterilized flies.** *J Econ Entomol* 1955, **462**-466.
5. Knipling E: **Possibilities of insect control or eradication through the use of sexually sterile males.** *J Econ Entomol* 1955, **459**-462.
6. Greek Ministry of Agriculture: **Description of research organization for the control of the olive fruit fly.** 1961, **33**, (in Greek).
7. Economopoulos A, Avtzis N, Zervas G, Tsitsipis J, Haniotakis G, Tsiropoulos G, Manoukas A: **Control of the olive fly, *Dacus oleae* (Gmelin), by the combined effects of insecticides and release of gamma sterilized insects.** *J Appl Entomol* 1977, **201**-215.
8. Economopoulos AP, Haniotakis GE, Mathioudis J, Missis N, Kinigakis P: **Long-distance flight of wild and artificially-reared *Dacus oleae* (Gmelin) (Diptera, Tephritidae).** *Z Angew Entomol* 1978, **101**-108.
9. Economopoulos A, Zervas G: **The quality problem in olive flies produced for SIT experiments.** *IAEA STI/PUB* 1982.
10. Economopoulos A: **The olive fruit fly, *Bactrocera* (*Dacus*) *oleae* (Gmelin) (Diptera: Tephritidae): its importance and control; previous SIT research and pilot testing.** *Int At Energy Agency, Vienna, Austria* 2002.
11. Zervas GA, Economopoulos AP: **Mating frequency in caged populations of wild and artificially reared (normal or  $\gamma$ -sterilized) olive fruit flies.** *Environ Entomol* 1982, **17**-20.
12. Loukas M, Economopoulos AP, Zouros E, Vergini Y: **Genetic changes in artificially reared colonies of the olive fruit fly.** *Ann Ent Soc Amer* 1985, **159**-165.
13. Economopoulos A, Loukas M: **ADH allele frequency changes in olive fruit flies shift from olives to artificial larval food and vice versa, effect of temperature.** *Entomol Exp Appl* 1986, **215**-221.
14. Economopoulos A: **Sexual competitiveness of gamma-ray sterilized males of *Dacus oleae*. Mating frequency of artificially reared and wild females.** *Env Entomol* 1972, **490**-497.
15. Capuzzo C, Firrao G, Mazzon L, Squartini A, Girolami V: **"Candidatus *Erwinia dadicola*", a coevolved symbiotic bacterium of the olive fly *Bactrocera oleae* (Gmelin).** *Int J Syst Evol Microbiol* 2005, **55**(Pt 4):1641-7.
16. Sacchetti P, Granchietti A, Landini S, Viti C, Giovannetti L, Belcari A: **Relationships between the olive fly and bacteria.** *J Appl Entomol* 2008, **132**:682-689.
17. Estes AM, Hearn DJ, Bronstein JL, Pierson EA: **The olive fly endosymbiont, "Candidatus *Erwinia dadicola*," switches from an intracellular existence to an extracellular existence during host insect development.** *Appl Environ Microbiol* 2009, **75**:7097-106.
18. Ben-Yosef M, Aharon Y, Jurkevitch E, Yuval B: **Give us the tools and we will do the job: symbiotic bacteria affect olive fly fitness in a diet-dependent fashion.** *Proc Biol Sci* 2010, **277**:1545-52.
19. Kounatidis I, Crotti E, Sapountzis P, Sacchi L, Rizzi A, Chouaia B, Bandi C, Alma A, Daffonchio D, Mavragani-Tsipidou P, Bourtzis K: ***Acetobacter tropicalis* is a major symbiont of the olive fruit fly (*Bactrocera oleae*).** *Appl Environ Microbiol* 2009, **75**:3281-8.
20. Augustinos AA, Stratikopoulos EE, Zacharopoulou A, Mathiopoulos KD: **Polymorphic microsatellite markers in the olive fly, *Bactrocera oleae*.** *Mol Ecol Notes* 2002, **2**:278-280.
21. Nardi F, Carapelli A, Dallai R, Frati F: **The mitochondrial genome of the olive fly *Bactrocera oleae*: two haplotypes from distant geographical locations.** *Insect Mol Biol* 2003, **12**:605-611.

22. Augustinos AA, Mamuris Z, Stratikopoulos EE, D'Amelio S, Zacharopoulou A, Mathiopoulos KD: **Microsatellite analysis of olive fly populations in the Mediterranean indicates a westward expansion of the species.** *Genetica* 2005, **125**:231-41.
23. Nardi F, Carapelli A, Dallai R, Roderick GK, Frati F: **Population structure and colonization history of the olive fly, *Bactrocera oleae* (Diptera, Tephritidae).** *Mol Ecol* 2005, **14**:2729-38.
24. Nardi F, Carapelli A, Boore JL, Roderick GK, Dallai R, Frati F: **Domestication of olive fly through a multi-regional host shift to cultivated olives: comparative dating using complete mitochondrial genomes.** *Mol Phylogenet Evol* 2010, **57**:678-86.
25. Zygouridis NE, Augustinos AA, Zalom FG, Mathiopoulos KD: **Analysis of olive fly invasion in California based on microsatellite markers.** *Heredity (Edinb)* 2009, **102**:402-12.
26. Dogaç E, Kandemir İ, Taskin V: **The genetic polymorphisms and colonization process of olive fly populations in Turkey.** *PLoS One* 2013, **8**: e56067.
27. Mavragani-Tsipidou P: **Genetic and cytogenetic analysis of the olive fruit fly *Bactrocera oleae* (Diptera: Tephritidae).** *Genetica* 2002, **116**:45-57.
28. Mavragani-Tsipidou P, Karamanlidou G, Zacharopoulou A, Koliadis S, Kastritis C: **Mitotic and polytene chromosome analysis in *Dacus oleae* (Diptera: Tephritidae).** *Genome* 1992, **35**:373-8.
29. Zambetaki A, Kleanthous K, Mavragani-Tsipidou P: **Cytogenetic analysis of Malpighian tubule and salivary gland polytene chromosomes of *Bactrocera oleae* (*Dacus oleae*) (Diptera: Tephritidae).** *Genome* 1995, **38**:1070-81.
30. Drosopoulou E, Chrysopoulou A, Nikita V, Mavragani-Tsipidou P: **The heat shock 70 genes of the olive pest *Bactrocera oleae*: genomic organization and molecular characterization of a transcription unit and its proximal promoter region.** *Genome* 2009, **52**:210-4.
31. Drosopoulou E, Nakou I, Síchová J, Kubičková S, Marec F, Mavragani-Tsipidou P: **Sex chromosomes and associated rDNA form a heterochromatic network in the polytene nuclei of *Bactrocera oleae* (Diptera: Tephritidae).** *Genetica* 2012, **140**:169-80.
32. Vontas JG, Hejazi MJ, Hawkes NJ, Cosmidis N, Loukas M, Hemingway J, Janes RW: **Resistance-associated point mutations of organophosphate insensitive acetylcholinesterase, in the olive fruit fly *Bactrocera oleae*.** *Insect Mol Biol* 2002, **11**:329-336, April.
33. Vontas J, Blass C, Koutsos AC, David JP, Kafatos FC, Louis C, Hemingway J, Christophides GK, Ranson H: **Gene expression in insecticide resistant and susceptible *Anopheles gambiae* strains constitutively or after insecticide exposure.** *Insect Mol Biol* 2005, **14**:509-21.
34. Kakani EG, Mathiopoulos KD: **Organophosphosphate resistance-related mutations in the acetylcholinesterase gene of Tephritidae.** *J Appl Entomol* 2008, **132**:762-771.
35. Kakani EG, Bon S, Massoulié J, Mathiopoulos KD: **Altered GPI modification of insect AChE improves tolerance to organophosphate insecticides.** *Insect Biochem Mol Biol* 2011, **41**:150-8.
36. Khila A, El Haidani A, Vincent A, Payre F, Souda SI: **The dual function of ovo/shavenbaby in germline and epidermis differentiation is conserved between *Drosophila melanogaster* and the olive fruit fly *Bactrocera oleae*.** *Insect Biochem Mol Biol* 2003, **33**:691-9.
37. Benos P, Tavernarakis N, Brogna S, Thireos G, Savakis C: **Acquisition of a potential marker for insect transformation: isolation of a novel alcohol dehydrogenase gene from *Bactrocera oleae* by functional complementation in yeast.** *Mol Gen Genet* 2000, **263**:90-5.
38. Lagos D, Ruiz MF, Sánchez L, Komitopoulou K: **Isolation and characterization of the *Bactrocera oleae* genes orthologous to the sex determining Sex-lethal and doublesex genes of *Drosophila melanogaster*.** *Gene* 2005, **348**:111-21.
39. Lagos D, Koukidou M, Savakis C, Komitopoulou K: **The transformer gene in *Bactrocera oleae*: the genetic switch that determines its sex fate.** *Insect Mol Biol* 2007, **16**:221-30.
40. Tsoumani KT, Mathiopoulos KD: **Genome size estimation with quantitative real-time PCR in two Tephritidae species: *Ceratitis capitata* and *Bactrocera oleae*.** *J Appl Entomol* 2012, **136**:626-631.
41. Tsoumani KT, Drosopoulou E, Mavragani-Tsipidou P, Mathiopoulos KD: **Molecular characterization and chromosomal distribution of a species-specific transcribed centromeric satellite repeat from the olive fruit fly, *Bactrocera oleae*.** *PLoS One* 2013, **8**:e79393.
42. Tsoumani KT, Augustinos AA, Kakani EG, Drosopoulou E, Mavragani-Tsipidou P, Mathiopoulos KD: **Isolation, annotation and applications of expressed sequence tags from the olive fly, *Bactrocera oleae*.** *Mol Genet Genomics* 2011, **285**:33-45.
43. Pavlidi N, Dermauw W, Rombauts S, Chrisargiris A, Van Leeuwen T, Vontas J: **Analysis of the Olive Fruit Fly *Bactrocera oleae* Transcriptome and Phylogenetic Classification of the Major Detoxification Gene Families.** *PLoS One* 2013, **8**:e66533.
44. Koukidou M, Klinakis A, Reboulakis C, Zagoraiou L, Tavernarakis N, Livadaras I, Economopoulos A, Savakis C: **Germ line transformation of the olive fly *Bactrocera oleae* using a versatile transgenesis marker.** *Insect Mol Biol* 2006, **15**:95-103.
45. Ant T, Koukidou M, Rempoulakis P, Gong HF, Economopoulos A, Vontas J, Alphey L: **Control of the olive fruit fly using genetics-enhanced sterile insect technique.** *BMC Biol* 2012, **10**:51.
46. Apostolaki A, Livadaras I, Saridaki A, Chrysargiris A, Savakis C, Bourtzis K: **Transinfection of the olive fruit fly *Bactrocera oleae* with *Wolbachia*: towards a symbiont-based population control strategy.** *J Appl Entomol* 2011, **135**:546-553.
47. Estes AM, Hearn DJ, Burrack HJ, Rempoulakis P, Pierson EA: **Prevalence of *Candidatus Erwinia dadicola* in wild and laboratory olive fruit fly populations and across developmental stages.** *Environ Entomol* 2012, **41**:265-74.
48. Alphey L, Andreassen M: **Dominant lethality and insect population control.** *Mol Biochem Parasitol* 2002, **121**:173-8.
49. Alphey L, Beard C, Ben, Billingsley P, Coetzee M, Crisanti A, Curtis C, Eggleston P, Godfray C, Hemingway J, Jacobs-Lorena M, James AA, Kafatos FC, Mukwaya LG, Paton M, Powell JR, Schneider W, Scott TW, Sina B, Sinden R, Sinkins S, Spielman A, Touré Y, Collins FH: **Malaria control with genetically manipulated insect vectors.** *Science* 2002, **298**:119-21.
50. Heinrich JC, Scott MJ: **A repressible female-specific lethal genetic system for making transgenic insect strains suitable for a sterile-release program.** *Proc Natl Acad Sci USA* 2000, **97**:8229-32.
51. Thomas DD, Donnelly CA, Wood RJ, Alphey LS: **Insect population control using a dominant, repressible, lethal genetic system.** *Science* 2000, **287**:2474-6.
52. Gong P, Epton MJ, Fu G, Scaife S, Hiscox A, Condon KC, Condon GC, Morrison NI, Kelly DW, Dafa'alla T, Coleman PG, Alphey L: **A dominant lethal genetic system for autocidal control of the Mediterranean fruitfly.** *Nat Biotechnol* 2005, **23**:453-6.
53. Sagri E, Reczko M, Gregoriou M-E, Tsoumani KT, Zygouridis NE, Salpea KD, Zalom FG, Ragoussis J, Mathiopoulos KD: **Olive fly transcriptomics analysis implicates energy metabolism genes in spinosad resistance.** *BMC Genomics* 2014, **15**:714.
54. Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J: **De novo assembly of human genomes with massively parallel short read sequencing.** *Genome Res* 2010, **20**:265-72.
55. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A: **High-throughput functional annotation and data mining with the Blast2GO suite.** *Nucleic Acids Res* 2008, **36**:3420-35.
56. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L: **Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.** *Nat Biotechnol* 2010, **28**:511-5.
57. Domanitskaya E V, Liu H, Chen S, Kubli E: **The hydroxyproline motif of male sex peptide elicits the innate immune response in *Drosophila melanogaster*.** *FEBS J* 2007, **274**:5659-68.
58. McGraw LA, Clark AG, Wolfner MF: **Post-mating gene expression profiles of female *Drosophila melanogaster* in response to time and to four male accessory gland proteins.** *Genetics* 2008, **179**:1395-408.
59. Hardy RW, Tokuyasu KT, Lindsley DL: **Analysis of spermatogenesis in *Drosophila melanogaster* bearing deletions for Y-chromosome fertility genes.** *Chromosoma* 1981, **83**:593-617.
60. Gibbons IR: **Dynein family of motor proteins: present status and future questions.** *Cell Motil Cytoskeleton* 1995, **32**:136-44.
61. Goldstein LS, Hardy RW, Lindsley DL: **Structural genes on the Y chromosome of *Drosophila melanogaster*.** *Proc Natl Acad Sci USA* 1982, **79**:7405-9.

62. Gepner J, Hays TS: **A fertility region on the Y chromosome of *Drosophila melanogaster* encodes a dynein microtubule motor.** *Proc Natl Acad Sci USA* 1993, **90**:11132-6.
63. Koerich LB, Wang X, Clark AG, Carvalho AB: **Low conservation of gene content in the *Drosophila* Y chromosome.** *Nature* 2008, **456**:949-51.
64. Carvalho AB, Lazzaro BP, Clark AG: **Y chromosomal fertility factors kl-2 and kl-3 of *Drosophila melanogaster* encode dynein heavy chain polypeptides.** *Proc Natl Acad Sci USA* 2000, **97**:13239-44.
65. Kimble J, Edgar L, Hirsh D: **Specification of male development in *Caenorhabditis elegans*: the fem genes.** *Dev Biol* 1984, **105**:234-9.
66. Doniach T, Hodgkin J: **A sex-determining gene, fem-1, required for both male and hermaphrodite development in *Caenorhabditis elegans*.** *Dev Biol* 1984, **106**:223-35.
67. Spence AM, Coulson A, Hodgkin J: **The product of fem-1, a nematode sex-determining gene, contains a motif found in cell cycle control proteins and receptors for cell-cell interactions.** *Cell* 1990, **60**:981-90.
68. flybase. [http://www.flybase.org].
69. Yeh SD, Chen YJ, Chang AC, Ray R, She BR, Lee WS, Chiang HS, Cohen SN, Lin-Chao S: **Isolation and properties of Gas8, a growth arrest-specific gene regulated during male gametogenesis to produce a protein associated with the sperm motility apparatus.** *J Biol Chem* 2002, **277**:6311-7.
70. Yang Y, Cochran DA, Gargano MD, King I, Samhat NK, Burger BP, Sabourin KR, Hou Y, Awata J, Parry DAD, Marshall WF, Witman GB, Lu X: **Regulation of flagellar motility by the conserved flagellar protein CG34110/Ccdc135/FAP50.** *Mol Biol Cell* 2011, **22**:976-87.
71. Garrett-Engle CM, Siegal ML, Manoli DS, Williams BC, Li H, Baker BS: **intersex, a gene required for female sexual development in *Drosophila*, is expressed in both sexes and functions together with doublesex to regulate terminal differentiation.** *Development* 2002, **129**:4661-75.
72. Gubbay J, Collignon J, Koopman P, Capel B, Economou A, Münsterberg A, Vivian N, Goodfellow P, Lovell-Badge R: **A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes.** *Nature* 1990, **346**:245-50.
73. Wilson MJ, Dearden PK: **Evolution of the insect Sox genes.** *BMC Evol Biol* 2008, **8**:120.
74. Crémazy F, Berta P, Girard F: **Sox neuro, a new *Drosophila* Sox gene expressed in the developing central nervous system.** *Mech Dev* 2000, **93**:215-9.
75. Prokopenko SN, Brumby A, O'Keefe L, Prior L, He Y, Saint R, Bellen HJ: **A putative exchange factor for Rho1 GTPase is required for initiation of cytokinesis in *Drosophila*.** *Genes Dev* 1999, **13**:2301-14.
76. O'Keefe L, Somers WG, Harley A, Saint R: **The pebble GTP exchange factor and the control of cytokinesis.** *Cell Struct Funct* 2001, **26**:619-26.
77. Khurana B, Kristie TM: **A protein sequestering system reveals control of cellular programs by the transcriptional coactivator HCF-1.** *J Biol Chem* 2004, **279**:33673-83.
78. Talbo G, Højrup P, Rahbek-Nielsen H, Andersen SO, Roepstorff P: **Determination of the covalent structure of an N- and C-terminally blocked glycoprotein from endocuticle of *Locusta migratoria*. Combined use of plasma desorption mass spectrometry and Edman degradation to study post-translationally modified proteins.** *Eur J Biochem* 1991, **195**:495-504.
79. Pondeville E, Maria A, Jacques J-C, Bourgoignie C, Dauphin-Villemant C: **Anopheles gambiae males produce and transfer the vitellogenic steroid hormone 20-hydroxyecdysone to females during mating.** *Proc Natl Acad Sci USA* 2008, **105**:19631-6.
80. Doctor J, Fristrom D, Fristrom JW: **The pupal cuticle of *Drosophila*: biphasic synthesis of pupal cuticle proteins in vivo and in vitro in response to 20-hydroxyecdysone.** *J Cell Biol* 1985, **101**:189-200.
81. Leal WS: **Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes.** *Annu Rev Entomol* 2013, **58**:373-91.
82. Pelosi P, Maida R: **Odorant-binding proteins in insects.** *Comp Biochem Physiol B Biochem Mol Biol* 1995, **111**:503-14.
83. Vosshall LB, Hansson BS: **A unified nomenclature system for the insect olfactory coreceptor.** *Chem Senses* 2011, **36**:497-8.
84. Vanderhaeghen P, Schurmans S, Vassart G, Parmentier M: **Olfactory receptors are displayed on dog mature sperm cells.** *J Cell Biol* 1993, **123**(6 Pt 1):1441-52.
85. Vanderhaeghen P, Schurmans S, Vassart G, Parmentier M: **Specific repertoire of olfactory receptor genes in the male germ cells of several mammalian species.** *Genomics* 1997, **39**:239-46.
86. Kang N, Koo J: **Olfactory receptors in non-chemosensory tissues.** *BMB Rep* 2012, **45**:612-22.
87. Spehr M, Gisselmann G, Poplawski A, Riffell JA, Wetzel CH, Zimmer RK, Hatt H: **Identification of a testicular odorant receptor mediating human sperm chemotaxis.** *Science* 2003, **299**:2054-8.
88. Spehr M, Schwane K, Heilmann S, Gisselmann G, Hummel T, Hatt H: **Dual capacity of a human olfactory receptor.** *Curr Biol* 2004, **14**:R832-3.
89. Spehr M, Schwane K, Riffell JA, Zimmer RK, Hatt H: **Odorant receptors and olfactory-like signaling mechanisms in mammalian sperm.** *Mol Cell Endocrinol* 2006, **250**:128-36.
90. Fukuda N, Yomogida K, Okabe M, Touhara K: **Functional characterization of a mouse testicular olfactory receptor and its role in chemosensing and in regulation of sperm motility.** *J Cell Sci* 2004, **117**(Pt 24):5835-45.
91. Veitinger T, Riffell JR, Veitinger S, Nascimento JM, Triller A, Chandsawangbhuwana C, Schwane K, Geerts A, Wunder F, Berns MW, Neuhaus EM, Zimmer RK, Spehr M, Hatt H: **Chemosensory Ca<sup>2+</sup> dynamics correlate with diverse behavioral phenotypes in human sperm.** *J Biol Chem* 2011, **286**:17311-25.
92. Pitts RJ, Liu C, Zhou X, Malpartida JC, Zwiebel LJ: **Odorant receptor-mediated sperm activation in disease vector mosquitoes.** *Proc Natl Acad Sci USA* 2014, **111**:2566-71.
93. Paesen GC, Happ GM: **The B proteins secreted by the tubular accessory sex glands of the male mealworm beetle, *Tenebrio molitor*, have sequence similarity to moth pheromone-binding proteins.** *Insect Biochem Mol Biol* 1995, **25**:401-8.
94. Thymianou S, Mavroidis M, Kokolakis G, Komitopoulou K, Zacharopoulou A, Mintzas AC: **Cloning and characterization of a cDNA encoding a male-specific serum protein of the Mediterranean fruit fly, *Ceratitis capitata*, with sequence similarity to odorant-binding proteins.** *Insect Mol Biol* 1998, **7**:345-53.
95. Paiva-Silva GO, Sorgine MHF, Benedetti CE, Meneghini R, Almeida IC, Machado EA, Dansa-Petretski M, Yepiz-Plascencia G, Law JH, Oliveira PL, Masuda H: **On the biosynthesis of *Rhodnius prolixus* heme-binding protein.** *Insect Biochem Mol Biol* 2002, **32**:1533-41.
96. Forêt S, Maleszka R: **Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*).** *Genome Res* 2006, **16**:1404-13.
97. Vogt RG, Prestwich GD, Lerner MR: **Odorant-binding-protein subfamilies associate with distinct classes of olfactory receptor neurons in insects.** *J Neurobiol* 1991, **22**:74-84.
98. Arya GH, Weber AL, Wang P, Magwire MM, Negron YL, Mackay TF, Anholt RR: **Natural variation, functional pleiotropy and transcriptional contexts of odorant binding protein genes in *Drosophila melanogaster*.** *Genetics* 2010, **186**:1475-85.
99. Zhou S, Stone EA, Mackay TF, Anholt RR: **Plasticity of the chemoreceptor repertoire in *Drosophila melanogaster*.** *PLoS Genet* 2009, **5**:e1000681.
100. Kodrık D, Filippov VA, Sehnal F, Filippova MA: **Sericotropin: an insect neurohormonal factor affecting RNA transcription.** *Netherlands J Zool* 1995.
101. Nagnan-Le Meillour P, Cain AH, Jacquin-Joly E, François MC, Ramachandran S, Maida R, Steinbrecht RA: **Chemosensory proteins from the proboscis of mamestra brassicae.** *Chem Senses* 2000, **25**:541-53.
102. Jacquin-Joly E, Vogt RG, François MC, Nagnan-Le Meillour P: **Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae*.** *Chem Senses* 2001, **26**:833-44.
103. Wanner KW, Willis LG, Theilmann DA, Isman MB, Feng Q, Plettner E: **Analysis of the insect os-d-like gene family.** *J Chem Ecol* 2004, **30**:889-911.
104. Walensky LD, Ruat M, Bakin RE, Blackshaw S, Ronnett G V, Snyder SH: **Two novel odorant receptor families expressed in spermatids undergo 5'-splicing.** *J Biol Chem* 1998, **273**:9378-87.
105. Suzuki N, Garbers DL: **Stimulation of sperm respiration rates by speract and resact at alkaline extracellular pH.** *Biol Reprod* 1984, **30**:1167-74.
106. Parmentier M, Libert F, Schurmans S, Schiffrmann S, Lefort A, Eggerickx D, Ledent C, Mollereau C, Gérard C, Perret J, et al: **Expression of members of the putative olfactory receptor gene family in mammalian germ cells.** *Nature* 1992, **355**:453-5.
107. Ogaugwu CE, Schetelig MF, Wimmer EA: **Transgenic sexing system for *Ceratitis capitata* (Diptera: Tephritidae) based on female-specific embryonic lethality.** *Insect Biochem Mol Biol* 2013, **43**:1-8.
108. Hanife G: **Embryonic development of the olive fruit fly, *Bactrocera oleae* Rossi (Diptera: Tephritidae), in vivo.** *Turkish J Zool* 2014.



109. Schweisguth F, Lepesant JA, Vincent A: **The serendipity alpha gene encodes a membrane-associated protein required for the cellularization of the Drosophila embryo.** *Genes Dev* 1990, **4**:922-31.
110. Ibensouda S, Schweisguth F, de Billy G, Vincent A: **Relationship between expression of serendipity alpha and cellularisation of the Drosophila embryo as revealed by interspecific transformation.** *Development* 1993, **119**:471-83.
111. Schmid KJ, Tautz D: **A screen for fast evolving genes from Drosophila.** *Proc Natl Acad Sci USA* 1997, **94**:9746-50.
112. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, Anthouard V, Arensburg P, Atkinson PW, Baden H, de Berardinis V, Baldwin D, Benes V, Biedler J, Blass C, Bolanos R, Boscus D, Barnstead M, et al: **The genome sequence of the malaria mosquito Anopheles gambiae.** *Science* 2002, **298**:129-49.
113. Zou Z, Lopez DL, Kanost MR, Evans JD, Jiang H: **Comparative analysis of serine protease-related genes in the honey bee genome: possible involvement in embryonic development and innate immunity.** *Insect Mol Biol* 2006, **15**:603-14.
114. Haugen M, Flannery E, Tomchaney M, Mori A, Behura SK, Severson DW, Duman-Scheel M: **Semaphorin-1a is required for Aedes aegypti embryonic nerve cord development.** *PLoS One* 2011, **6**:e21694.
115. Schetelig MF, Horn C, Handler AM, Wimmer EA: **Development of an Embryonic Lethality System in Mediterranean Fruit Fly Ceratitis capitata.** In *Area-Wide Control Insect Pests* MJB Vreysen, AS Robinson J Hendrichs 2007, 85-93.
116. Tadros W, Lipshitz HD: **The maternal-to-zygotic transition: a play in two acts.** *Development* 2009, **136**:3033-42.
117. Schetelig MF, Caceres C, Zacharopoulou A, Franz G, Wimmer EA: **Conditional embryonic lethality to improve the sterile insect technique in Ceratitis capitata (Diptera: Tephritidae).** *BMC Biol* 2009, **7**:4.
118. Gabrieli P, Gornulski LM, Bonomi A, Siciliano P, Scolari F, Franz G, Jessup A, Malacrida AR, Gasperi G: **Interchromosomal duplications on the Bactrocera oleae Y chromosome imply a distinct evolutionary origin of the sex chromosomes compared to Drosophila.** *PLoS One* 2011, **6**:e17747.
119. Bilak A, Su TT: **Regulation of Drosophila melanogaster pro-apoptotic gene hid.** *Apoptosis* 2009, **14**:943-9.
120. Kornbluth S, White K: **Apoptosis in Drosophila: neither fish nor fowl (nor man, nor worm).** *J Cell Sci* 2005, **118**(Pt 9):1779-87.
121. Hay BA, Guo M: **Caspase-dependent cell death in Drosophila.** *Annu Rev Cell Dev Biol* 2006, **22**:623-50.
122. Steller H: **Regulation of apoptosis in Drosophila.** *Cell Death Differ* 2008, **15**:1132-8.
123. Horn C, Wimmer EA: **A transgene-based, embryo-specific lethality system for insect pest management.** *Nat Biotechnol* 2003, **21**:64-70.
124. Schetelig MF, Nirmala X, Handler AM: **Pro-apoptotic cell death genes, hid and reaper, from the tephritid pest species, Anastrepha suspensa.** *Apoptosis* 2011, **16**:759-68.
125. Grether ME, Abrams JM, Agapite J, White K, Steller H: **The head involution defective gene of Drosophila melanogaster functions in programmed cell death.** *Genes Dev* 1995, **9**:1694-708.
126. Abrams JM, White K, Fessler LI, Steller H: **Programmed cell death during Drosophila embryogenesis.** *Development* 1993, **117**:29-43.
127. M.E Economopoulos A, Tsitsipis J: **The importance of conditions during the adult stage in evaluating an artificial food for larvae of Dacus oleae (Gmel.) (Diptera, Tephritidae).** *Z Angew Entomol* 1967, **59**:127-130.
128. Tsitsipis J: **Development of a caging and egg system for mass rearing the olive fruit fly, Dacus oleae (Gmel.) (Diptera, Tephritidae).** *Ann Zool Ecol Anim* 1977, **9**:133-139.
129. Tsitsipis JA, Kontos A: **Improved solid adult diet for the olive fruit fly Dacus oleae.** *Entomol Hell* 1983, **1**:24-29.
130. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL: **BLAST+: architecture and applications.** *BMC Bioinformatics* 2009, **10**:421.
131. Trapnell C, Pachter L, Salzberg SL: **TopHat: discovering splice junctions with RNA-Seq.** *Bioinformatics* 2009, **25**:1105-11.
132. Blüthgen N, Brand K, Cajavec B, Swat M, Herzel H, Beule D: **Biological profiling of gene groups utilizing Gene Ontology.** *Genome Inform* 2005, **16**:106-15.

doi:10.1186/1471-2156-15-S2-S8

Cite this article as: Sagri et al.: The molecular biology of the olive fly comes of age. *BMC Genetics* 2014 **15**(Suppl 2):S8.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

