UNIVERSITY OF CRETE

FORTH
FOUNDATION FOR RESEARCH AND TECHNOLOGY - HELLAS

Graduate Program in Molecular Biology and Biomedicine

Thesis

Analysis of insecticide resistance mechanisms in enemies of public health and agriculture with the use of -omics technologies and bioinformatics

Student: Jason Charamis

Supervisor: Professor John Vontas

Three-member Examination Committee:

Professor John Vontas, Agricultural University of Athens

Researcher B' Pavlos Pavlidis, ICS-FORTH

Researcher B' Christoforos Nikolaou, BRSC Alexander Fleming

Heraklion, July 2020

Aknowledgements

Master thesis is the most important course of a post-graduate programme, as it provides the students with invaluable knowledge and research experience on a practical level, thereby preparing them for the challenges of a career in science.

Completing my thesis at the Molecular Entomology Lab was demanding but also very productive. I learned a variety of computational methods and approaches and acquired a plethora of programming skills, while it also contributed into cultivating my scientific thinking.

For that reason, I would like to thank some people who had a decisive contribution into carrying out my thesis. First of all, I would like to thank Professor John Vontas, who trusted me, supported me and guided me throughout the course of this year. Further, I would like to specially thank Dr. Panagiotis Ioannidis for his inspiring and productive day-to-day supervising and guidance, as well as for his enjoyable and engaging mentoring. In addition, I would like to thank Mary Kefi, Aris Ilias and all the members of the Molecular Entomology Lab for their companionship and their contribution to a nice working environment, as well as Evangelia Morou for her willingness and decisiviness to always helping me with the bureucratic issues.

Finally I would like to thank my partner, Andriani, my friends and my family for always being there for me and for their unquestionable support throughout the course of my studies in Crete.

Thank you very much!

Summary

Insecticide resistance imposes great challenges in public health and global food security. This study focused on investigating mechanisms of insecticide resistance by using genomic and transcriptomic data. More specifically, we studied the possible association of gene transcripts expressed in the legs of resistant mosquitoes with insecticide resistance. The transcriptomic data suggest that the constitutive resistance phenotype is associated with the over-expression of cytochrome P450s and other detoxification enzymes, while there is also a possible contribution of cuticular resistance mechanisms. Moreover, to uncover genes participating in deltamethrin induction we examined the leg-specific transcriptomic profile of this highly resistant strain after exposure to deltamethrin. The results of the transcriptomic analysis suggest that deltamethrin induces the expression of genes coding for detoxification enzymes and chemosensory proteins with a possible pyrethroid-protecting role. In addition, it also indicates the involvement of GPCRs in the development of resistance after exposure to the insecticide. Further, we studied the evolution of cytochrome P450s in the olive fly *Bactrocera oleae* and performed phylogenetic comparisons with the P450omes of *Drosophila melanogaster* and *Ceratitis capitata*. This analysis demonstrated that *B. oleae* bears expansions in the CYP12A and CYP12E subfamilies belonging to the mitochondrial clan and the CYP6A and CYP6G subfamilies belonging to the CYP3 clan of cytochrome P450s. In summary, this study demonstrates the complex mechanisms controlling constitutive resistance and deltamethrin induction in *A. coluzzii* and has generated a valuable genomic resource for experimental and computational studies in *B. oleae*.

Table of Contents

1. Introduction

1.1 Why study insecticide resistance?

Insecta is the most diverse and species-rich animal group in the world, constituting an estimated 53% of all species living today (Robinson *et al.*, 2011). Insects have an extremely important role for the maintenance and productivity of the natural and agricultural ecosystems (Robinson *et al.*, 2011). Throughout human history they have been integral to the development of human societies, by serving as pollinators, to food producers and more recently as model organisms for biological research.

However, a small proportion of the total insect species imposes enormous challenges to human sustainability. Public health is under constant threat of several insect-borne diseases that cause millions of deaths every year, mostly in the developing countries (Wilson *et al.*, 2020). Mosquitoes are the leading vectors of insect-borne diseases, transmitting malaria, dengue virus, West Nile virus, Zika virus, yellow fever, chikungunya fever and lymphatic filariasis (David *et al.*, 2013).

Furthermore, global food safety is at risk from many insect species. Herbivorous insects are responsible for destroying up to one-fifth of the total agricultural and livestock production worldwide, thus destabilizing global food production and inflicting major losses in global economy (Robinson *et al.*, 2011).

Insecticides are synthetic chemical compounds that are used to contain insect pest populations by targeting core molecules and/or biological processes of insect physiology. There are four main classes of insecticides: organochlorines, organophosphates, carbamates and pyrethroids (Perry *et al.*, 2011). The latter, are the most commonly used insecticides for vector control due to their high biological activity, limited side effects to mammals and fast biodegradation rate (Antwi & Reddy, 2015).

Since the introduction of dichlorodiphenyltrichloroethane (DDT) in the 1940s, insecticides were established as the first line of defence against vector-borne diseases and agricultural pests and they still remain the most cost-effective method of dealing with these challenges (David *et al.*, 2013). However, the massive use of insecticides has caused intense selection in insect populations throughout the world, thus leading to the emergence of resistance to insecticides globally (Balabanidou *et al.*, 2018).

According to the World Health Organization (WHO), insecticide resistance is defined as "*changes in an insect that increase its ability to withstand or overcome the effects of one or more insecticides. When the frequency of resistant insects in a population increases (e.g. through resistance traits being passed on from one generation to another), the efficacy of an insecticidal intervention can be compromised. Increased frequency of resistant insects may be detected through assays that measure insect mortality in response to a particular insecticide, or through genetic tests that detect resistance mechanisms in individual insects.*" (Morand & Lajaunie, 2017, pp. 86). The development of insecticide resistance has a huge impact on the efficiency of insecticide-based interventions to control target insect populations (Hemingway *et al.*,

2002). Unraveling insecticide resistance mechanisms is crucial for the development of Integrated Pest Management (IPM) strategies as well as for the design of novel and more effective insecticides (Hemingway *et al.*, 2002), thus critically contributing on the sustainability of insecticide-based interventions to ensure global health and prosperity. To this end, the recent advances in genomics and transcriptomics have produced massive data resources and have offered unprecedented capabilities towards delineating the adaptations that underlie the complex resistance phenotypes. Nowadays, novel candidate genes can be identified extremely rapidly even in poorly studied insects (ffrench-Constant, 2014) and their role in resistance can be further investigated using genetic experiments and functional expression biochemical assays.

1.2 Mechanisms of Insecticide Resistance

Insecticide resistance is caused by three main mechanisms: target-site resistance, metabolic resistance and cuticular resistance (Balabanidou *et al.*, 2018).

1.2.1 Target-site Resistance

Target-site resistance is caused by non-synonymous mutations that lead to target protein modifications, thus resulting in decreased sensitivity or total insensitivity to the insecticide (Feyereisen *et al.*, 2015). These include point mutations such as small insertions and deletions as well as transposable element insertions (Feyereisen *et al.*, 2015). The most well-known cases of target-site resistance involve mutations in the voltage-gated sodium channel (*vgsc*) gene (Dong, 2007), also known as knock-down (*kdr*) resistance mutations, the acetyl-cholinesterase (Mutero *et al.*, 1994) and the chitin synthase genes (Douris *et al.*, 2016; Fotakis *et al.*, 2020).

1.2.2 Metabolic Resistance

Metabolic resistance is caused by genetic changes that lead to over-expression or higher catalytic activity of detoxification enzymes, which act by metabolizing or sequestering insecticides (Feyereisen *et al.*, 2015). The enzyme families which are most commonly associated with this type of resistance are cytochrome P450s (CYPs), glutathione-S-transferases (GSTs) and carboxyl/cholinesterases (CCEs) (Ranson *et al.*, 2002). Among them, cytochrome have the leading role in insecticide detoxification.
Cytochrome P450s are 45-55 kDa heme-thiolate enzymes that are capable of catalyzing more than 60 reactions (Feyereisen, 2012), but are mostly known for catalyzing the monooxygenase reaction, an oxidation-reduction process which is mediated by the NADPH-cytochrome P450 reductase (CPR) and is characterized by high substrate specificity (Feyereisen, 2012). Cytochrome P450 metabolism has a

leading role in insecticide detoxification (David *et al.*, 2013) and the exact mechanisms of P450 activity have been previously described in detail (Feyereisen, 2005; Feyereisen, 2012).
Cytochrome P450s comprise one of the largest super-gene families and are found in all living organisms, including bacteria, fungi, plants, insects and mammals (Feyereisen, 2012). Insect cytochrome P450s are typically distributed within four major clans: CYP2, CYP3, CYP4 and mitochondrial (Feyereisen, 2012). These include gene families with a conserved role in insect physiology and development as well as rapidly expanding gene families associated with the metabolism of pesticides and other xenobiotic compounds (Ranson *et al.*, 2002). As a result, cytochrome P450 monooxygenases are extremely important for the development of insecticide resistance as well as for host-plant adaptation in phytophagous insects (Calla *et al.*, 2017; Wang *et al.*, 2018; Zhu *et al.*, 2016).

1.2.3 Cuticular Resistance

The insect leg is the entry point of the insecticides. Cuticular resistance is developed by genetic changes that lead to the thickening or modification of the leg cuticle, thereby reducing or slowing the uptake of insecticides which allows their effective clearance by detoxification mechanisms (Balabanidou *et al.*, 2018). This type of resistance is particularly relevant because it confers extreme multi-resistance phenotypes (Balabanidou *et al.*, 2016). Cuticular resistance is developed by complex mechanisms involving genes associated with cuticle structure formation, such as cuticular proteins, as well as genes participating in cuticular hydrocarbon biosynthesis and transport, including cytochrome P450s belonging to the CYP4G subfamily, ABC transporters and odorant-binding proteins (OBPs) (Balabanidou *et al.*, 2019).

1.2.4 Synergism between different resistance mechanisms

Resistance phenotypes are complex traits which tend to be polygenic and are developed by combining distinct resistance mechanisms (Hawkins *et al.*, 2018). Combining distinct resistance mechanisms has important ramifications in insecticide resistance management (IRM). A recent study has functionally validated that the combination of target-site resistance and metabolic resistance leads to synergistic, almost multiplicative effects in resistance capacity (Samantsidis & Panteleri *et al.*, 2020).

1.3 Insecticide resistance in malaria vectors

1.3.1 Malaria is resurging due to insecticide resistance

Malaria is one of the most important life-threatening diseases in the world, causing more than 400,000 deaths annually, two-thirds of which occur in children aged under five years old (WHO, 2019). Africa carries the highest burden of this disease, bearing an estimated 93% of all malaria cases and 94% of all malaria deaths worldwide (WHO, 2019).

Human malaria is an insect-borne disease, caused by *Plasmodium* parasites (mainly *P. falciparum*) that are transmitted through the bites of infected female *Anopheles* mosquitoes (WHO, 2019). Although there are many Anopheline species that are capable of transmitting the disease, the major malaria vectors in sub-saharan Africa, the region with the highest incidence of malaria in the world, are *A. gambiae* and its closely related species *A. coluzzii* (The *Anopheles* gambiae 1000 Genomes Consortium, 2017).

Due to the recurrent failure of producing an effective vaccine against *Plasmodium* parasites (Wykes, 2013), as well as the drug efficacy limitations of the available antimalarial treatment (Tse *et al.*, 2019), mosquito vector control is proven to be the most effective means of reducing malaria incidence and the number of deaths associated with this disease (Bhatt *et al.*, 2015). More specifically, malaria incidence has halved since 2000 and up to 80% of this reduction is attributed to insecticide-based interventions, such as the widespread use of long-lasting pyrethroid-impregnated bednets and indoor residual spraying (IRS) (Bhatt *et al.*, 2015).

However, the extensive use of these interventions exerted a massive selection pressure on mosquito populations that has led to the development and spread of resistance to the most commonly used insecticides, mostly pyrethroids. The escalation of pyrethroid resistance in African malaria vectors is now at a critical point and poses a serious threat in reversing the gains that have been recently made by malaria vector control (Heminway *et al.*, 2016). In this context, the monitoring and investigation of the molecular basis of insecticide resistance in *Anopheles* mosquitoes is particularly relevant to avoid a global health disaster caused by the failure of malaria vector control (Heminway *et al.*, 2016).

1.3.2 Mechanisms of pyrethroid resistance in *Anopheles* mosquitoes

Since 2000, research has been ongoing on insecticide resistance in malaria vectors (Hemingway *et al.*, 2002). A common approach of investigating the molecular basis of insecticide resistance in *Anopheles* mosquitoes is to collect transcriptomic data from wild resistant populations and use them to compare gene expression with susceptible laboratory strains. This approach has led to the identification of several

key candidate genes that are involved in the emerging pyrethroid resistance in sub-saharan Africa. A brief summary on the current status of insecticide resistance in malaria vectors is further presented.

1.3.2.1 Metabolic resistance

Metabolic resistance is highly prevalent in *Anopheles* mosquitoes. As a result, transcriptomic studies commonly identify cytochrome P450 genes with elevated expression levels in resistant *A. gambiae* and *A. coluzzii* populations (Riveron *et al.*, 2018). Not surprisingly, these P450s mostly belong to the CYP4, CYP6 and CYP9 families, with most of those implicated in resistance belonging to the CYP6 family (Riveron *et al.*, 2018). Taking into account substrate specificity and selectivity, the four P450s which are mostly associated with insecticide resistance in malaria vectors are CYP6P3 (Edi *et al.*, 2014), CYP6P4 (Toé *et al.*, 2015), CYP6M2 (Mitchell *et al.*, 2013) and CYP9K1 (Vontas *et al.*, 2018). Furthermore, other detoxification genes including GSTs, CCEs and ABC transporters contribute to the development of metabolic resistance in *Anopheles* mosquitoes. However, with the exception of GSTE2 which has a well-characterized role in pyrethroid resistance (Adolfi *et al.*, 2019; Riveron *et al.*, 2014), they usually have a secondary role compared to that of cytochrome P450s (David *et al.*, 2013).

1.3.2.2 Cuticular Resistance

Cuticular resistance has been recently established as responsible for the multi-resistance phenotypes of the *A. coluzzii* VK7 strain (Balabanidou *et al.*, 2016). The detailed investigation of this resistance mechanism demonstrated that mosquitoes remodel their leg cuticle by enhancing the deposition of cuticular proteins and chitin, as well as by producing higher and different relative amounts of cuticular hydrocarbons (Balabanidou *et al.*, 2019). Cuticular hydrocarbon (CHC) biosynthesis is mediated by two cytochrome P450s of the CYP4G family, CYP4G16 and CYP4G17 (Kefi *et al.*, 2019). Both these cytochrome P450s function as decarbonylases but are functionally distinct and produce different CHC blends (Kefi *et al.*, 2019).

1.3.2.3 Novel mechanisms of resistance

Recent transcriptomic studies have suggested the emergence of novel resistance mechanisms mediated by protein families with no previous implications in insecticide resistance, including hexamerins (Ingham *et al.*, 2018), salivary gland proteins (Isaacs *et al.*, 2018) and sensory appendage proteins (Ingham *et al.*, 2020). More specifically,

it has been proposed that these protein families contribute to resistance by binding and sequestering insecticides (Ingham *et al.*, 2018).
In this context, a recent study demonstrated that SAP2, a sensory appendage protein, is overexpressed in resistant *A. gambiae* and *A. coluzzii* populations and confers pyrethroid resistance by binding and sequestering insecticides at the first point of mosquito contact (Ingham *et al.*, 2020). Further, genome data revealed that *SAP2* is under recent positive selection in *A. coluzzii* populations from Burkina Faso and is associated with an increase in the prevalence of pyrethroid resistance in the region (Ingham *et al.*, 2020).

1.3.2.4 Insecticide Induction

Insecticide resistance is not a static but a rather dynamic phenomenon, which is developed through the interplay of multiple intracellular pathways and insect physiological responses. Insecticide induction includes the mechanisms underlying the temporal and spatial regulation of resistance after insecticide detection and is particularly relevant for understanding the dynamics of insecticide response. However, it has not been characterized in detail.
Initial studies have demonstrated that long-term permethrin exposure leads to overexpression of cytochrome P450s and other detoxification genes in *A. gambiae* resistant strains (Vontas *et al.*, 2005). Furthemore, recent studies in *Culex* mosquitoes have provided evidence suggesting that G protein-coupled receptors (GPCRs) have an important role in the development of insecticide resistance (Li *et al.*, 2014; Li *et al.*, 2015; Zhou *et al.*, 2018). More specifically, it has been proposed that distinct GPCRs regulate pyrethroid resistance by controlling the expression of cytochrome P450s (Li *et al.*, 2014) through conserved signalling pathways (Li *et al.*, 2015; Zhou *et al.*, 2018).
Most importantly, a recent study has demonstrated that permethrin exposure induces GPCR signaling thereby leading to enhanced P450 expression and the development of resistance (Li & Liu, 2019). These results suggest that GPCRs have a particularly important role in insecticide resistance.

1.4 Insecticide resistance in *Bactrocera oleae*

Bactrocera oleae (Diptera: Tephritidae), also known as the olive fly, is the most destructive pest of olive orchards worldwide (Pavlidi *et al.*, 2017). The monophagous larva of this pest is capable of feeding on the olive mesocarp while coping with the oxidative stress imposed by the high levels of phenolic compounds found in the green, unripe olives (Pavlidi *et al.*, 2017).

B. oleae causes significant reductions in olive production and quality, thus damaging the marketing value of table olives and their most important derivative product, olive oil (Margaritopoulos *et al.*, 2008). It has been estimated that the annual economic losses due to *B. oleae* reach up to 30-35% of the total olive production in Greece (Pavlidi *et al.*, 2018). In the island of Crete, where olive trees cover up to 65% of total agricultural land, *B. oleae* causes economic losses of >20 million euros per year, while more than one million euros are spent for its control annually (Kampouraki *et al.*, 2018). Globally, the annual economic losses caused by *B. oleae* extend to USD 800 million (Daane & Johnson, 2010).

Control of the olive fly in Greece and other Meditteranean countries is based on insecticide applications, including insecticide baits and cover sprays (Kampouraki *et al.*, 2018). More specifically, olive fly control was initially achieved with the use of organophosphates, mainly dimethoate, while more recently pyrethroids, mainly alpha-cypermethrin (Pavlidi *et al.*, 2018), and the macrocyclic lactone spinosad have been introduced and are increasingly used for this purpose (Kampouraki *et al.*, 2018).

Due to the high mobility and the general tendency of spatial dispersion that characterize the tephritid flies fly populations, it is generally considered that they have a relatively limited potential for developing insecticide resistance in the field compared to other dipteran insects (Vontas *et al.*, 2011). However, the extensive use of these insecticides has reportedly led to the development of insecticide resistance in *B. oleae* field populations. Kampouraki *et al.* (2018) monitored the alpha-cypermethrin and spinosad resistance status of 55 field populations from Greece collected from 2009 to 2016 (Kampouraki *et al.*, 2018). This study demonstrated that since 2013 *B. oleae* is rapidly developing resistance to alpha-cypermethrin (Kampouraki *et al.*, 2018). As a result, extreme resistance phenotypes have been observed (Kampouraki *et al.*, 2018).

Furthermore, a recent transcriptomic study (Pavlidi *et al.*, 2018) has associated two overexpressed cytochrome P450 genes with alpha-cypermethrin resistance in *B. oleae*, in agreement with previous studies (Margaritopoulos *et al.*, 2008). Remarkably, no target-site resistance mutation was detected in this study, thus indicating the contribution of detoxification mechanisms in the observed resistance phenotype (Pavlidi *et al.*, 2018).

Until now, the olive fruit fly detoxification genes have not been systematically studied at a genome-wide level. This is probably due to the limitations imposed by the low quality of the initially sequenced *B. oleae* genome. The first effort to systematically

characterize detoxification genes in the olive fruit fly was conducted by Pavlidi *et al.* (2013) using a wide *B. oleae* transcriptome dataset (Pavlidi *et al.*, 2013).
The recently published and significantly improved version of the *B. oleae* genome (Bayega *et al.*, 2020) offers far better opportunities for annotating and characterizing the olive fly detoxification genes, including cytochrome P450s, glutathione-S-transferases (GSTs), carboxylesterases (CCEs) and ABC transporters. These new genomic resources will substantially facilitate molecular studies in *B. oleae* and will contribute to further understanding of the emerging resistance. Overall, this approach can lead to more efficient strategies of monitoring and overcoming pyrethroid resistance in *B. oleae* field-evolving populations.

2. Aim of this study

The aim of this study is to analyze insecticide resistance in mosquitoes and agricultural pests by using genomic and transcriptomic data. Mosquito leg is the first contact point of the insecticide and the most relevant tissue for insecticide uptake. To gain insights into the involvement of the mosquito leg in pyrethroid resistance, we compared the leg transcriptomes of a highly resistant *A. coluzzii* strain with two susceptible *A. coluzzii* strains. Moreover, to broaden our understanding over the mechanisms underlying deltamethrin induction we compared the leg transcriptomes of deltamethrin-exposed and unexposed highly resistant mosquitoes. Finally, we performed a detailed analysis of the olive fly detoxification genes. More specifically, we manually curated the *B. oleae* CYPome and subsequently performed phylogenetic comparisons with the CYPomes of *D. melanogaster* and *Ceratitis capitata*.

3. Materials and Methods

3.1 Data availability

The reference genome assemblies and gene annotations of *A. coluzzii* Ngousso and *A. gambiae* PEST were downloaded from VectorBase (Giraldo-Calderón *et al.*, 2015). The recently published and improved de novo assembly of the *B. oleae* genome (Bayega *et al.*, 2020) was downloaded from GenBank under the accesion assembly number GCA_001188975.4.

3.2 Reference-based RNA-seq analysis

The raw RNAseq reads from all four strains were mapped on the recently published *A. coluzzii* str. Ngousso genome (Kingan *et al.*, 2019) using hisat2 (Kim *et al.*, 2019). Next, expression was quantified at the gene level by using featureCounts (Liao *et al.*, 2014) and the differential expression analyses were performed with EdgeR (Robinson & Oshlack, 2010) using Trinity-bundled perl scripts.

3.3 De novo transcriptome assembly

As a validation step, the transcriptomic analysis was also carried out based on *de novo* transcriptome assembly (Haas *et al.*, 2013). In particular, the RNAseq reads from all samples, except for the samples which were removed by the PCA analysis, were assembled with Trinity v2.8.6 (Grabherr *et al.*, 2011), using the following parameters: "--seqType fq --SS_lib_type RF --max_memory 350G --CPU 24 --trimmomatic --no_salmon". Next, Kallisto (Bray *et al.*, 2016) was used to estimate transcript abundance and the differential expression analyses were performed as previously described (Sakka *et al.*, 2020).

3.4 Annotation of the *A. coluzzii* Ngousso genome

Since there was no available functional gene annotation, all predicted genes in the *A. coluzzii* Ngousso genome were functionally annotated. In particular, we used BLASTP v. 2.6.0+ (Camacho *et al.*, 2008) against Uniref50 (Suzek *et al.*, 2015) for identifying similar genes and InterProScan v. 5.28-67.0 (Jones *et al.*, 2014) for

finding conserved protein domains. Further, we identified the one-to-one orthologs between the *A. coluzzii* Ngousso (AcolN1.0) and the *A. gambiae* PEST (AgamP4.12) reference gene sets by implementing a best reciprocal hit (BRH) approach, using SWIPE (Rognes, 2011) with an e-value cut-off of 10^{-3}. This analysis identified 12,523 one-to-one orthologs between the two species.

3.5 GO Term Enrichment Analysis

The one-to-one orthologs between *A. coluzzii* Ngousso and *A. gambiae* were subsequently used on Gene Ontology (GO) enrichment analyses. More specifically, g:Profiler (Reimand *et al.*, 2016) was used to perform functional enrichment analyses and find significantly over-represented GO terms on the differentially expressed gene sets compared to the *A. gambiae* PEST reference genome (AgamP4), which is accessible via the Ensembl Genomes Metazoa database (Howe *et al.*, 2019).

3.6 Data editing and Visualization

Normalized expression values for all genes, namely Counts Per Million (CPM) and Transcripts Per Million (TPM), were computed using custom perl scripts. Moreover, Principal Component Analysis (PCA) was carried out using EdgeR, while the PCA plot was constructed with custom R scripts.

The Venn diagram was constructed using the VennDiagram R package (Chen & Boutros, 2011), while visualization of the GO enrichment results was implemented with custom R scripts which make use of the ggplot2 R package (Wickham, 2016).

3.7 Manual Curation

Gene prediction was performed in-house on the recently published *B. oleae* genome (Bayega *et al.*, 2020), using the BRAKER pipeline (Hoff *et al.*, 2019). Subsequently, I extracted the predicted P450 gene models based on the conserved protein domain (Pfam: PF00067) which is distinctive for cytochrome P450s. The predicted P450 gene models were then manually curated on an in-house implementation of the Web Apollo Genome Browser (Lee *et al.*, 2013).

Manual curation was conducted based on expression data from four developmental stages (egg, larvae, pupa and adult) and four *B. oleae* field-collected adult populations from Greece (Kampouraki *et al.*, 2018), as well as based on protein sequence similarity with the respective cytochrome P450s of other closely related tephritid

species, including *Bactrocera dorsalis*, *Bactrocera latifrons* and *Zeugodacus cucurbitae* (Bayega *et al.*, 2020). In total, 79/118 (67%) cytochrome P450 gene models had been correctly predicted, while 39/118 (33%) predicted gene models were edited.

3.8 Phylogeny Reconstruction

The curated *D. melanogaster* P450 gene set was downloaded by David R. Nelson's website (https://drnelson.uthsc.edu/). The *C. capitata* P450 gene set, which is curated by Pedro Hernandez-Crespo's lab, was provided by René Feyereisen.

Multiple sequence alignment was performed with Mafft v7.310 (Katoh & Standley, 2013) using the default parameters. The produced alignments were automatically trimmed using trimAl (Capella-Gutiérrez *et al.*, 2009) and a custom Bash script was used to convert the trimmed alignments to a phylip format file. Finally, the phylogenetic tree was built under the maximum likelihood optimality criterion using RaxML 8.2.11 (Stamatakis, 2014).

4. Results and Discussion

4.1 Leg-specific transcriptomic profiling reveals mechanisms of constitutive resistance and deltamethrin induction

To identify genes which are possibly implicated in the *A. coluzzii* VK7 multi-resistant phenotype and to broaden our understanding over the leg-specific deltamethrin detection, we performed an RNA-seq-based leg transcriptomic profiling of four *A. coluzzii* strains, generating a total of >968 million Illumina reads. In particular, we studied two *A. coluzzii* VK7 strains (Toé *et al.*, 2015) of varying levels of resistance, hereafter named VK7-LR (lowly resistant), VK7-HR (highly resistant), and VK7-IN (deltamethrin-induced), while Ngousso (Toé *et al.*, 2015) was used as the susceptible strain. It should be noted that the VK7-IN strain is actually the VK7-HR strain, after 1 hour of deltamethrin exposure, followed by 1 hour of recovery.

Before implementing further investigation, a Principal Component Analysis (PCA) was performed to assess the quality of the replicates, using the transcription levels of all genes (Fig. 1). The results of this analysis showed that most of the biological replicates for each strain clustered together and separately from the replicates of the other strains. However, one replicate from each of the Ngousso, VK7-HR and VK7-IN (Ngousso_3, VK7-HR_4 & VK7-IN_2) samples did not follow the expected pattern (Fig. 1). As a result, they were subsequently excluded to improve the reliability of downstream analyses.

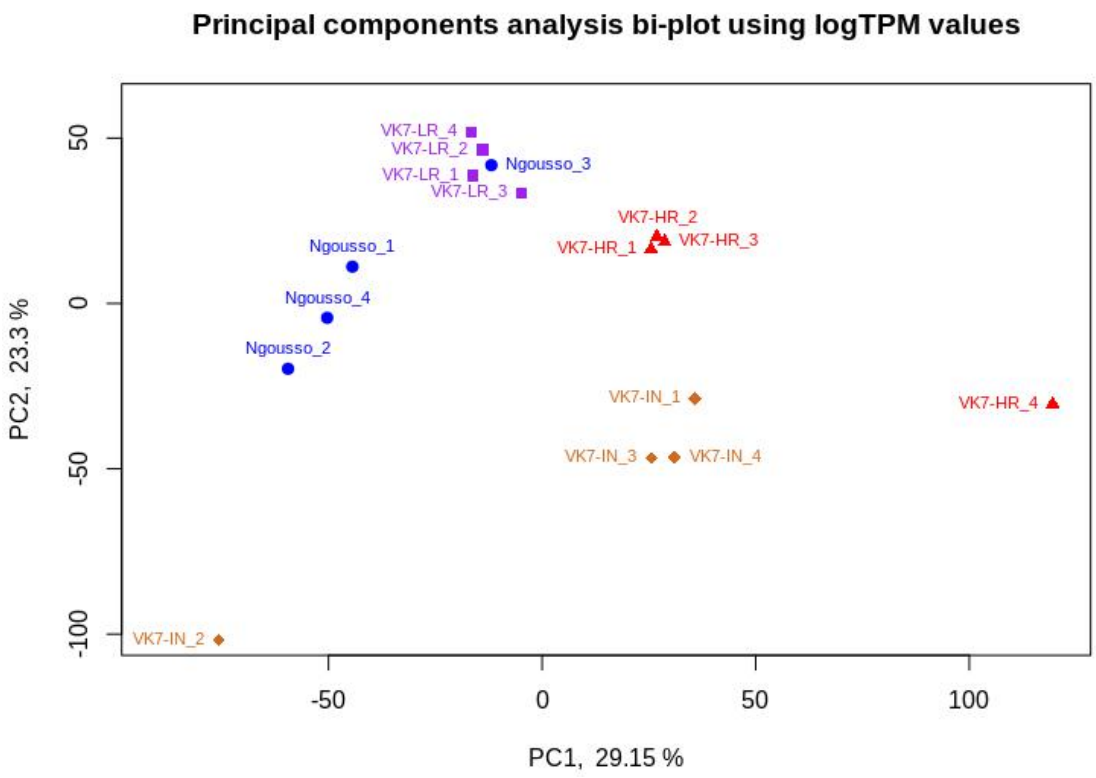

Figure 1. Principal components analysis of the gene expression levels for the four *A. coluzzii* strains. Replicates Ngousso_3, VK7-HR_4 and VK7-IN_2 that do not follow the expected pattern were excluded from all downstream analyses.

There are two comparisons that are relevant to the aims of this study; (1) comparing the resistant VK7-HR strain with the two susceptible strains, VK7-LR and Ngousso. This comparison against two different rates of susceptibility will give information on genes that are related to constitutive resistance. (2) comparing VK7-IN and VK7-HR will focus on genes related to deltamethrin induction.

4.1.1 Constitutive Resistance

Using a set of strict statistical parameters (log|FC| > 2, FDR < 0.01), we identified 555 differentially expressed genes in VK7-HR against Ngousso (111 up-regulated, 444 down-regulated) and 432 differentially expressed genes in VK7-HR against VK7-LR (120 up-regulated, 312 down-regulated) (Fig. 2). The top 20 most up-regulated and top 20 most down-regulated genes are presented on Table S1. Interestingly, 76 of the total 156 up-regulated genes (48.7%) were commonly over-expressed, whereas 163 of the 593 down-regulated genes (27.5%) were commonly under-expressed in both comparisons (Fig. 2). The total 76 commonly up-regulated genes are presented on Table 1, while the 163 commonly down-regulated genes are presented on Table S2 (Appendix).

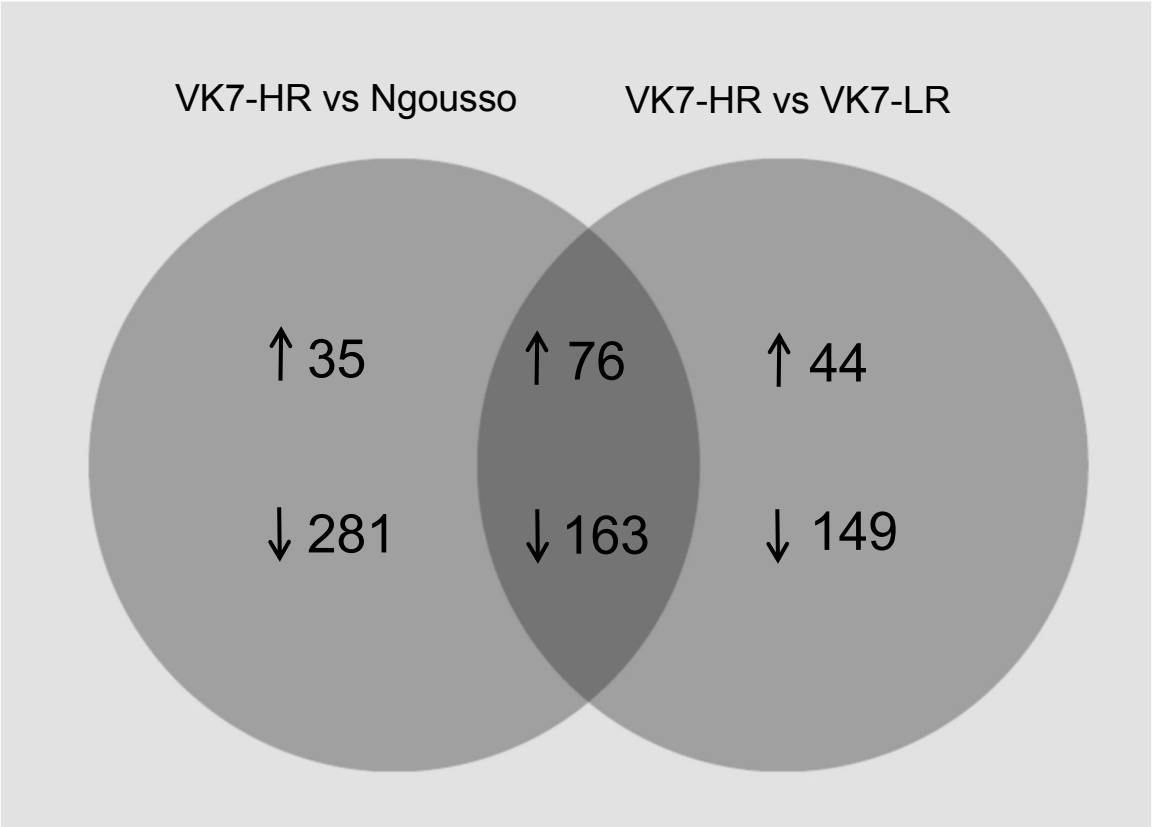

Figure 2. Venn diagram showing the number of the differentially expressed genes (log|FC| > 4, p-value < 10^{-3} & FDR < 0.01) between VK7-HR (resistant) and the two susceptible strains, Ngousso and VK7-LR. Upward arrows indicate over-expressed genes, while downward arrows represent under-expressed genes in VK7-HR compared to each of the susceptible strains respectively.

We focused on the subset of the 76 commonly up-regulated genes (log|FC| > 2, FDR < 0.05) in VK7-HR against both susceptible strains. The functional enrichment analysis of the consistently up-regulated genes revealed the overrepresentation of GO terms that are strongly related to P450 activity and insecticide detoxification (Fig. 3). These include iron ion binding (GO:0005506), monooxygenase activity (GO:0004497), heme binding (GO:0020037), tetrapyrrole binding (GO:0046906), oxidoreductase activity (GO:0016491), drug catabolic process (GO:0042737) and cofactor binding (GO:0048037) (Fig. 3). Importantly, the majority of these GO terms have been associated with a recent case of strong pyrethroid resistance in *A. funestus* (Weedall *et al.*, 2019). The results of the functional enrichment analysis among the 163 commonly down-regulated genes is presented in Fig. S1 (Appendix).

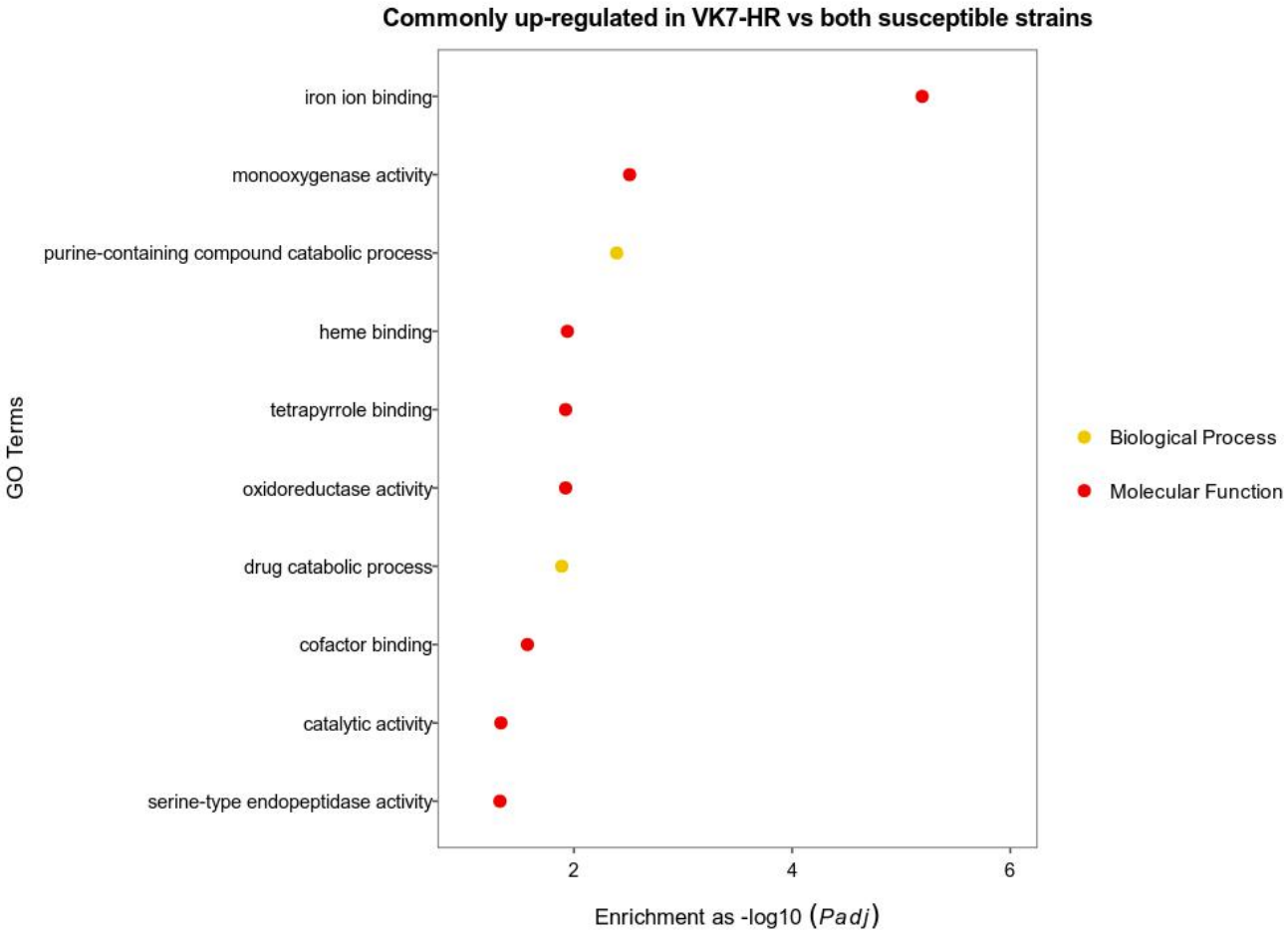

Figure 3. Over-represented GO terms on the 76 commonly up-regulated genes in the two comparisons related to constitutive resistance, VK7-HR vs VK7-LR and VK7-HR vs Ngousso.

Table 1. Overview of the 76 commonly up-regulated genes in VK7-HR (resistant) compared to both susceptible strains, VK7-LR and Ngousso.

		VK7-HR vs VK7-LR		VK7-HR vs Ngousso	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**	**Fold Change**	**Corrected p-value**
ACON012199	Histone H2B	232.01	4.62E-35	1176.53	7.42E-31
ACON029698	EGF-like domain containing protein	202.43	3.27E-15	59.34	6.68E-11
ACON006913	CUB domain containing protein	65.36	9.51E-17	56.48	1.11E-12
ACON010132	Class B Scavenger Receptor (CD36 domain) SCRBQ1	58.19	1.84E-175	64.32	2.90E-145
ACON007951	ionotropic receptor IR68a	57.96	7.38E-74	27.14	6.41E-47
ACON002210	cytochrome P450 CYP325B1	46.72	7.98E-12	142.64	2.77E-10
ACON013134	uncharacterized protein	46.37	1.05E-35	28.37	1.68E-17
ACON029319	uncharacterized protein	28.39	3.40E-122	15.55	1.68E-16
ACON006468	I-BAR domain containing protein MTSS1/MTSS2	22.85	2.22E-10	11.88	9.05E-08
ACON013465	cuticular protein (putative) CPLCP1	22.76	5.70E-07	65.71	4.80E-34
ACON005457	Ser/Thr protein phosphatase/nucleotidase	21.53	3.65E-15	33.50	4.42E-18
ACON013123	glucose dehydrogenase (acceptor)	20.19	2.61E-15	20.89	4.74E-13
ACON010831	thioester-containing protein 8 (TEP8)	19.79	4.83E-18	35.82	2.88E-17
ACON013316	uncharacterized protein	14.30	2.58E-14	50.16	1.19E-16
ACON000818	cytochrome P450 CYP9K1	13.72	3.41E-63	20.11	4.94E-15
ACON006931	cuticular protein RR-2 family 111 - CPR111	13.48	5.63E-15	5.92	1.11E-07

ACON010110	Acyl-CoA N-acyltransferase	13.16	1.86E-27	9.23	4.75E-15
ACON006486	Prss3	12.99	2.72E-13	22.79	1.00E-16
ACON006914	fibrinogen-related protein 1	12.87	2.76E-33	14.81	4.93E-15
ACON003381	cuticular protein with chitin-binding R&R consensus	12.66	2.76E-08	22.42	4.68E-08
ACON006487	trypsin-alpha	12.42	1.62E-11	16.72	7.47E-13
ACON029432	uncharacterized protein	10.54	1.43E-27	9.53	1.06E-20
ACON010830	thioester-containing protein 9 - TEP9	10.40	4.20E-10	23.87	1.21E-10
ACON003383	cuticular protein with chitin-binding R&R consensus	9.97	2.95E-10	38.00	5.48E-18
ACON029290	uncharacterized protein	9.89	7.80E-15	55.08	2.01E-58
ACON029314	Fibrinogen-like protein	9.74	1.00E-22	6.35	1.38E-12
ACON003380	cuticular protein RR-2 family 119 - CPR119	9.44	2.54E-07	16.06	3.35E-08
ACON013241	cytochrome P450 CYP4D16	9.41	1.26E-25	7.01	1.45E-18
ACON011121	pancreatic triacylglycerol lipase	9.34	6.35E-28	4.88	6.19E-18
ACON028407	uncharacterized protein	9.15	2.22E-18	9.05	1.45E-17
ACON007077	Acyltransferase 3-like	9.11	5.73E-47	12.10	9.22E-79
ACON001864	cytochrome P450 CYP4H15	8.91	4.86E-58	7.48	7.58E-32
ACON029279	Transient receptor potential cation channel subfamily A member 1	8.71	8.65E-10	6.94	1.10E-4S
ACON006191	chitinase Cht24	8.67	6.39E-19	7.22	5.51E-13
ACON007158	Alpha crystallin/Heat shock protein	8.40	2.02E-08	4.48	2.50E-06
ACON000155	alkylated DNA repair protein alkB homolog	7.95	1.37E-110	6.97	8.34E-63

ACON006225	aldehyde oxidase	7.87	5.17E-61	7.20	1.58E-38
ACON002867	cytochrome P450 CYP6P4	7.75	1.43E-08	20.26	1.52E-53
ACON001596	Putative cuticle protein	7.68	3.40E-11	6.33	2.89E-14
ACON008446	cuticular protein CPLCG family - CPLCG3	7.57	1.65E-26	4.52	6.43E-16
ACON008447	cuticular protein CPLCG family - CPLCG4	7.43	1.99E-08	7.50	4.56E-15
ACON006223	glucosyl/glucuronosyl transferase	7.38	6.68E-31	8.34	1.18E-26
ACON004449	pancreatic triacylglycerol lipase	7.20	9.62E-11	4.28	2.92E-05
ACON006209	carboxypeptidase B	7.20	1.28E-70	4.17	1.03E-21
ACON001531	sodium-independent sulfate anion transporter	6.90	5.61E-09	8.85	5.64E-08
ACON000976	uncharacterized protein	6.60	7.11E-09	17.45	7.34E-11
ACON007159	alpha-crystallin B chain	6.56	3.66E-07	4.46	1.61E-06
ACON006222	glucosyl/glucuronosyl transferase	6.54	2.58E-67	6.49	9.64E-46
ACON004665	cytochrome P450 CYP306A1	6.52	7.22E-34	6.25	1.43E-23
ACON009218	Peptidase S1A chymotrypsin family	6.33	1.22E-05	7.81	9.45E-06
ACON028147	uncharacterized protein	6.28	5.65E-16	5.26	8.03E-09
ACON001664	cuticular protein RR-2 family 1 (CPR1)	6.00	8.62E-21	5.69	3.39E-12
ACON007782	Transcription regulator GCM domain superfamily	5.88	5.45E-08	14.09	5.91E-16
ACON029479	Harbinger transposase-derived nuclease domain	5.54	4.35E-31	9.00	1.17E-48
ACON006485	trypsin-alpha	5.48	1.10E-06	20.46	1.44E-12
ACON007616	EGF-like protein	5.24	4.01E-19	9.14	3.15E-20

ACON007060	Leucine-rich repeat domain superfamily	5.13	2.28E-08	4.25	1.02E-05
ACON029782	sulfotransferase St1	5.10	4.93E-06	12.09	7.04E-44
ACON008893	cuticular protein CPLCP8	5.06	1.21E-10	7.52	1.46E-08
ACON005634	chitinase Cht2	5.05	2.20E-17	4.22	4.62E-11
ACON001503	separase	4.94	1.29E-22	5.86	9.82E-15
ACON007042	cuticular protein RR-1 family 62 - CPR62	4.93	4.21E-08	10.92	1.23E-13
ACON012850	cytochrome P450	4.93	1.16E-4	14.70	6.58E-58
ACON002391	carboxylesterase COEAE5O	4.87	4.46E-47	8.61	1.85E-35
ACON009273	Seminal fluid protein HACP027	4.81	1.30E-18	19.39	4.77E-11
ACON011794	CLIPA1 protein	4.77	1.14E-10	14.99	1.58E-19
ACON005458	Ser/Thr protein phosphatase/nucleotidase	4.77	9.84E-4	7.92	7.33E-06
ACON001774	G protein-coupled receptor rhodopsin-like	4.75	5.57E-05	4.37	9.22E-05
ACON011348	arylsulfatase b	4.75	4.61E-26	4.20	3.96E-12
ACON012988	uncharacterized protein	4.71	3.77E-21	5.18	4.78E-13
ACON002495	PH-like domain superfamily	4.65	9.52E-09	5.46	1.04E-13
ACON029108	Peptidase S1A chymotrypsin family	4.48	1.66E-37	5.81	5.32E-20
ACON002100	Dynein assembly factor 5 axonemal	4.31	1.62E-82	10.38	1.13E-108
ACON006489	Peptidase S1 PA clan	4.30	5.20E-11	5.90	1.44E-17
ACON006224	aldehyde oxidase	4.22	3.71E-45	4.12	2.43E-33
ACON029114	Peptidase S1 PA clan	4.00	8.51E-25	6.51	3.37E-67

Strikingly, 10 of the 76 (13.1%) commonly up-regulated genes encode detoxification enzymes, including cytochrome P450s (CYP), glutathione-S-transferases (GSTs), UDP-gluconotransferases (UGTs) and carboxylesterases (CCEs), while 10 of the 76 (13.1%) commonly up-regulated genes code for cuticular proteins (Fig. 4).

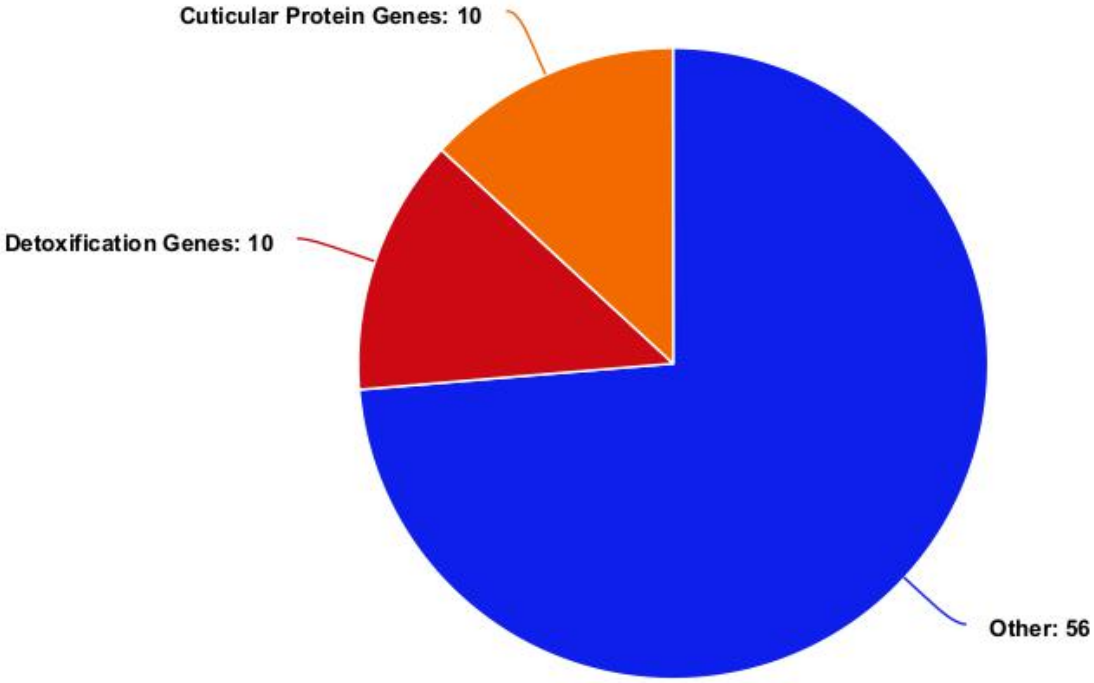

Figure 4. Distribution of the 76 commonly up-regulated genes in the two comparisons related to constitutive resistance, VK7-HR vs VK7-LR and VK7-HR vs Ngousso.

More specifically, seven CYP genes were commonly over-expressed in both comparisons (Table 2). Among them, *CYP9K1* and *CYP6P4* have been previously associated with pyrethroid resistance in *Anopheles* mosquitoes. More specifically, CYP9K1 is capable of metabolizing deltamethrin and it was responsible for the recent development of deltamethrin resistance in *A. coluzzii* populations in Bioko Island, West Africa (Vontas *et al.*, 2018). Further, functional studies have demonstrated that CYP6P4 can bind but cannot metabolize deltamethrin (Ibrahim *et al.*, 2016b). Nevertheless, due to its ability to metabolize very effectively permethrin and λ-cyhalotrin, it has been associated with pyrethroid resistance in *A. funestus* (Ibrahim *et al.*, 2016b), *A. arabiensis* (Simma *et al.*, 2019), *A. gambiae* (Edi *et al.*, 2014) and the *A. coluzzii* VK7 multi-resistant strain (Toé *et al.*, 2015). In addition, CYP4H15 has been previously associated with DDT resistance in *A. gambiae*, but it cannot directly metabolize DDT (Chiu *et al.*, 2008).

Furthermore, two other CYP genes were up-regulated only against the Ngousso susceptible strain: *CYP4H25* and *CYP6Z1* (Table 1).Moreover, CYP4H25 has been previously associated with λ-cyhalotrin resistance in *A. funestus* (Samb *et al.*, 2016), while CYP6Z1 metabolizes very effectively DDT (Chiu *et al.*, 2008), permethrin and deltamethrin, and less effectively bendiocarb, thus leading to carbamate/pyrethroid cross-resistance in *A. funestus* (Ibrahim *et al.*, 2016a).

Three GST genes were found to be up-regulated in VK7-HR legs, one (*GSTe2*) in VK7-HR against Ngousso and two (*GSTd3* and *GSTd10*) in VK7-HR against VK7-LR (Table 2). Two of them, *GSTe2* and *GSTd3*, have been previously implicated in pyrethroid resistance in *Anopheles* mosquitoes. More specifically, functional experiments have demonstrated that GSTe2 directly metabolizes both DDT and permethrin (Riveron *et al.*, 2014), thereby leading to DDT/pyrethroid cross-resistance in *A. funestus* (Riveron *et al.*, 2014), *A. gambiae* (Yahouédo *et al.*, 2017) and the *A. coluzzii* VK7 strain (Toé *et al.*, 2015). Moreover, GSTd3 has been associated with permethrin and deltamethrin resistance in *A. arabiensis* (Simma *et al.*, 2019) and *A. gambiae* (Stica *et al.*, 2019).

Further, there are several detoxification genes (encoding cytochrome P450s, UGTs and CCEs) with pronounced over-expression in VK7-HR compared to both susceptible strains, which are completely uncharacterized or poorly characterized with regard to insecticide resistance. In addition, 11 of the 163 (6.7%) commonly down-regulated genes code for detoxification enzymes. A summary of all differentially expressed detoxification genes is presented on Table 2.

Table 2. Overview of the differentially expressed detoxification genes in VK7-HR (resistant) compared to both susceptible strains, VK7-LR and Ngousso.

		VK7-HR vs VK7-LR		**VK7-HR vs Ngousso**	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**	**Fold Change**	**Corrected p-value**
Commonly up-regulated detoxification genes					
ACON002210	cytochrome P450 CYP325B1	46.72	7.98E-12	142.64	2.77E-10
ACON000818	cytochrome P450 CYP9K1	13.72	3.41E-63	20.11	4.94E-15
ACON013241	cytochrome P450 CYP4D16	9.41	1.26E-25	7.01	1.45E-18
ACON006225	aldehyde oxidase	7.87	5.17E-61	7.20	1.58E-38
ACON002867	cytochrome P450 CYP6P4	7.75	1.43E-08	20.26	1.52E-53

ACON006223	glucosyl/glucuron osyl transferase	7.38	6.68E-31	8.34	1.18E-26
ACON004665	cytochrome P450 CYP306A1	6.54	7.22E-34	6.49	1.43E-23
ACON006222	glucosyl/glucuron osyl transferase	6.52	2.58E-67	6.25	9.64E-46
ACON012850	cytochrome P450	4.93	1.16E-04	14.70	6.58E-58
ACON002391	carboxylesterase COEAE5O	4.87	4.46E-47	8.61	1.85E-35
ACON006224	aldehyde oxidase	4.22	3.71E-45	4.12	2.43E-33
Commonly down-regulated detoxification genes					
ACON006775	glucosyl/glucuron osyl transferases	-41.80	1.49E-32	-17.38	6.23E-07
ACON005166	Short-chain dehydrogenase/re ductase SDR	-15.90	1.02E-90	-15.12	9.84E-21
ACON002429	cytochrome P450 CYP314A1	-8.15	3.81E-21	-4.62	2.03E-10
ACON005834	carboxylesterase COEJHE2E	-7.45	6.64E-10	-4.08	5.75E-34
ACON005774	cytochrome P450 CYP49A1	-6.55	1.04E-09	-7.39	3.74E-28
ACON003608	cytochrome P450 CYP4AA1	-6.47	2.28E-04	-11.41	1.28E-11
ACON010077	cytochrome P450 CYP303A1	-6.02	3.61E-06	-7.40	2.46E-06
ACON006048	cytochrome P450 CYP4J5	-5.53	4.83E-66	-8.95	2.80E-33
ACON006049	cytochrome P450 CYP4J10	-5.25	5.31E-50	-8.40	3.57E-31
Up-regulated detoxification genes in VK7-HR vs VK7-LR					
ACON001864	cytochrome P450 CYP4H15	8.91	4.86E-58		
ACON005163	glucosyl/glucuron osyl transferase	7.12	6.85E-07		
ACON004382	glutathione S-transferase delta class 3 - GSTD3	4.95	4.72E-22		
ACON004383	glutathione S-transferase delta class 10 -	4.47	4.11E-06		

	GSTD10				
ACON006221	aldehyde oxidase	4.02	1.04E-06		
Down-regulated detoxification genes in VK7-HR vs VK7-LR					
ACON007028	glucosyl/glucuronosyl transferases	-7.03	3.21E-08		
ACON005501	dehydrogenase/reductase SDR family member 11 precursor	-6.23	4.28E-05		
ACON006724	carboxylesterase COEAE3G	-5.87	3.74E-05		
ACON011824	thioredoxin peroxidase 4 - TPX4	-5.27	4.69E-05		
ACON009799	ABC transporter family C member 13 - ABCC13	-4.90	9.97E-04		
ACON011507	carboxylesterase COE13O	-4.81	3.53E-09		
ACON000088	cytochrome P450 CYP4H19	-4.72	9.58E-08		
ACON005837	carboxylesterase COEJHE5E	-4.21	3.45E-23		
ACON005658	cytochrome P450 CYP15B1	-4.14	2.79E-04		
Up-regulated detoxification genes in VK7-HR vs Ngousso					
ACON009194	glutathione S-transferase epsilon class 2 - GSTE2			11.02	7.26E-30
ACON013305	cytochrome P450 CYP4H25			5.15	1.77E-15
ACON008219	cytochrome P450 CYP6Z1			4.95	2.80E-07
ACON007589	glucosyl/glucuronosyl transferase			4.47	7.50E-42
Down-regulated detoxification genes in VK7-HR vs Ngousso					
ACON006728	carboxylesterase COEAE7G			-27.57	1.18E-20

ACON001333	protein scarlet	-16.38	3.43E-32
ACON028695	carboxylesterase	-6.75	1.45E-04
ACON006047	cytochrome P450 CYP4J9	-5.56	5.30E-08
ACON011518	ABC transporter family A member 4 - ABCA4	-4.58	1.00E-13
ACON028426	carboxylesterase COE09941	-4.10	4.51E-05

Recent studies (Balabanidou *et al.*, 2016; 2018; 2019) have demonstrated the increasingly important role of the leg cuticle in the development of multi-resistance phenotypes against many classes of insecticides in *Anopheles* mosquitoes. In total, 11 cuticular protein genes were commonly up-regulated in both comparisons (14.5%), while six were up-regulated only against VK7-LR and four were up-regulated only against the Ngousso strain (Table 3).

More than half of these genes (12/21) encoded cuticular proteins which belong to the most abundant CPR family: eleven of these are members of the RR-2 sub-family, while only one (CPR62) belongs to the RR-1 sub-family. The remaining genes encoded cuticular proteins which belong to the CPLCG (5), CPLCP (2) and the TWDL (1) families respectively. Importantly, five of the over-expressed CPRs were recently identified in the leg proteome of the multi-resistant *A. coluzzii* VK7 strain and they are considered to have an important role in cuticle thickening that leads to pyrethroid resistance (Balabanidou *et al.*, 2019). In addition, three of the over-expressed CPLG genes, *CPLCG3* (ACON008446; 4.5-7.6 fold), *CPLCG4* (ACON008447; 7.5-fold) and *CPLCG15* (ACON008459; 7.7-17.7 fold) have been associated with deltamethrin resistance in *A. gambiae* (Nkya *et al.*, 2014; Yahouédo *et al.*, 2017), while immunostaining experiments have demonstrated that CPLCG3 and CPLCG4 are localized in the leg endocuticle, thus contributing to cuticle thickening and penetration rate of the insecticide (Vannini *et al.*, 2014). Further, *CPR111* (ACON006931; 6-13.5 fold) and *TWDL12* (ACON004576; 11-fold) were recently found to be over-expressed in multiple pyrethroid-resistant *A. funestus* strains (Weedall *et al.*, 2019). At last, *CPCLP1* (ACON013465; 23-66 fold) was among the most up-regulated genes against both susceptible strains, while the ortholog of *CPLCP3* (ACON008817; 282-fold) in *D. melanogaster*, *Vajk-4*, participates in cuticle barrier formation (Cinege *et al.*, 2017). Only two cuticular protein genes were commonly down-regulated in both comparisons (Table 3). A summary of all over-expressed genes encoding cuticular proteins is presented in Table 3.

Overall, the results of the functional enrichment analyses and the subsequent manual investigation of the commonly up-regulated genes, suggest that the constitutive resistance phenotype is possibly mediated by metabolic resistance mechanisms with a

minor contribution of cuticular resistance. Consistently, several genes coding for cytochrome P450s (Table 2) and cuticular proteins (Table 3) were among the most highly over-expressed genes in VK7-HR legs in both comparisons.

Table 3. Overview of the differentially expressed cuticular protein genes in VK7-HR (resistant) compared to the two susceptible strains, VK7-LR and Ngousso.

		VK7-HR vs VK7-LR		**VK7-HR vs Ngousso**	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**	**Fold Change**	**Corrected p-value**
Commonly up-regulated cuticular protein genes					
ACON013465	cuticular protein (putative) CPLCP1	22.76	5.70E-07	65.71	4.80E-34
ACON006931	cuticular protein RR-2 family 111 - CPR111	13.48	5.63E-15	5.92	1.11E-07
ACON003381	cuticular protein with chitin-binding R&R consensus	12.66	2.76E-08	22.42	4.68E-08
ACON003383	cuticular protein with chitin-binding R&R consensus	9.97	2.95E-10	38.00	5.48E-18
ACON003380	cuticular protein RR-2 family 119 - CPR119	9.44	2.54E-07	16.06	3.35E-08
ACON001596	Putative cuticle protein	7.68	3.40E-11	6.33	2.89E-14
ACON008446	cuticular protein CPLCG family - CPLCG3	7.57	1.65E-26	4.52	6.43E-16
ACON008447	cuticular protein CPLCG family - CPLCG4	7.43	1.99E-08	7.50	4.56E-15
ACON001664	cuticular protein RR-2 family 1 (CPR1)	6.00	8.62E-21	5.69	3.39E-12
ACON008893	cuticular protein CPLCP8	5.06	1.21E-10	7.52	1.46E-08
ACON007042	cuticular protein RR-1 family 62 - CPR62	4.93	4.21E-08	10.92	1.23E-13

Commonly down-regulated cuticular protein genes					
ACON002994	cuticular protein RR-2 family 10 - CPR10	-43.75	8.75E-06	-14.91	3.48E-04
ACON006283	cuticular protein RR-2 family 70 - CPR70	-4.17	1.38E-08	-4.18	3.18E-05
Up-regulated cuticular protein genes in VK7-HR vs VK7-LR					
ACON004576	cuticular protein TWDL family - TWDL12	11.30	7.50E-24		
ACON001669	cuticular protein RR-2 family 6 - CPR6	7.21	5.60E-12		
ACON006261	cuticular protein RR-2 family 135 - CPR135	6.28	2.60E-09		
ACON001668	cuticular protein RR-2 family 5 - CPR5	5.56	8.82E-11		
ACON008445	cuticular protein CPLCG family - CPLCG2	5.36	3.46E-15		
ACON000820	cuticular protein RR-2 family 125 - CPR125	4.43	6.59E-27		
Down-regulated cuticular protein genes in VK7-HR vs VK7-LR					
ACON006007	cuticular protein RR-1 family 28 - CPR28	-5.66	3.96E-05		
ACON010122	cuticular protein RR-2 family 132 - CPR132	-3.95	5.15E-04		
ACON006497	cuticular protein RR-2 family 134 - CPR134	-2.70	2.08E-05		
ACON010369	cuticular protein RR-2 family 112 - CPR112	-2.41	3.13E-07		
Up-regulated cuticular protein genes in VK7-HR vs Ngousso					
ACON008817	cuticular protein (putative) CPLCP3			282.91	3.16E-13

ACON008459	cuticular protein CPLCG family - CPLCG15	17.65	5.74E-06
ACON003385	cuticular protein RR-2 family 123 - CPR123	9.61	4.94E-04
ACON003377	cuticular protein RR-2 family 115 - CPR115	8.56	1.38E-04
Down-regulated cuticular protein gene in VK7-HR vs Ngousso			
ACON006009	cuticular protein RR-1 family 30 - CPR30	-3.24	8.80E-07

4.1.2 Deltamethrin induction

Using the same set of strict statistical parameters (log|FC| > 2, FDR < 0.01), we identified 432 differentially expressed genes in VK7-IR against VK7-HR (376 up-regulated, 56 down-regulated). We then investigated the over-represented GO terms in the up-regulated gene set using g:Profiler (Reimand *et al.*, 2016). The top 20 up-regulated and down-regulated genes in VK7-HR vs VK7-LR are presented on Table 4.

Table 4. Accession numbers and putative functions for the top 20 most up-regulated genes and the top 20 most down-regulated genes in the VK7-IN vs VK7-HR comparison.

		VK7-IN vs VK7-HR	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**
Top 20 most up-regulated genes			
ACON000612	salivary gland protein 1 - SG1	885.27	3.33E-26
ACON002629	Mucin related 18B - Mur18B	755.29	8.94E-08
ACON006505	uncharacterized protein	620.68	2.96E-15
ACON004038	heme peroxidase 8 - HPX8	474.59	1.16E-08
ACON000610	salivary gland protein 1-like 6 - SG1f	443.80	1.36E-28
ACON000607	salivary gland protein 1-like 3 - SG1c	435.58	1.06E-20
ACON000835	uncharacterized protein	405.95	6.63E-06
ACON008216	salivary gland protein 7 - SG7	383.15	7.18E-29
ACON028181	uncharacterized protein	355.68	6.17E-21
ACON010364	chitin binding domain containing protein	328.54	3.90E-07
ACON000605	putative conserved secreted protein	235.39	1.67E-05
ACON006504	salivary gland protein 2-like - SG2b	235.12	4.19E-42
ACON008215	salivary gland protein 7-like - SG7b	224.95	3.07E-14
ACON000548	salivary gland protein 1-like 2 - SG1b	215.55	4.77E-27
ACON006507	salivary gland protein 3 - SG3	200.59	6.44E-66
ACON011216	heme peroxidase 16 - HPX16	199.79	1.40E-48
ACON000609	salivary gland protein 1-like 5 - SG1e	154.26	5.33E-19
ACON006506	salivary gland protein 2 - SG2	152.95	2.95E-83

ACON000427	vitellogenin receptor	144.95	2.24E-07
Top 20 most down-regulated genes			
ACON010695	elongation of very long chain fatty acids protein 4	-28.88	5.31E-105
ACON002210	cytochrome P450 CYP325B1	-11.18	1.10E-06
ACON005389	CRAL-TRIO lipid binding domain containing protein	-10.25	3.97E-08
ACON007909	protocadherin-16/23	-9.84	7.25E-14
ACON013017	uncharacterized protein	-9.33	5.21E-12
ACON004494	Ecdysone-induced protein 93F (Eip93F)	-9.04	7.88E-81
ACON009090	Dopa decarboxylase	-8.30	1.35E-08
ACON002205	cytochrome P450 CYP325C2	-8.08	3.21E-16
ACON011153	FGF-like proteein	-7.34	6.80E-10
ACON010520	krueppel-like factor	-7.21	4.42E-36
ACON029331	Zinc finger C2H2-type protein (jing ortholog)	-6.78	2.37E-29
ACON029329	uncharacterized protein	-6.46	1.63E-18
ACON005388	cellular retinaldehyde binding protein	-6.41	1.55E-11
ACON009868	cuticular protein RR-1 family 73 CPR73	-6.27	4.35E-19
ACON008445	cuticular protein CPLCG family - CPLCG2	-6.04	4.78E-10
ACON005600	sugar transporter ERD6-like 6	-5.95	5.03E-05
ACON005207	SCY1-like protein 2	-5.93	4.40E-21
ACON029401	uncharacterized protein	-5.78	1.48E-12
ACON002138	cytochrome P450 CYP325H1	-5.75	1.85E-41
ACON007802	Argonaute hook domain protein (gawky ortholog)	-5.66	2.03E-29

The functional enrichment analysis of the up-regulated genes in VK7-IN versus VK7-HR legs demonstrated a significant enrichment for GO terms related to G protein-coupled receptor (GPCR) activity and drug catabolism (Fig. 5). Interestingly, several genes coding for GPCRs, cytochrome P450s, ABC transporters and odorant binding proteins (OBPs) were over-expressed after deltamethrin exposure. These results are in agreement with a previous study which has monitored gene expression in resistant and susceptible *A. gambiae* strains after pyrethroid exposure (Vontas *et al.*, 2005).

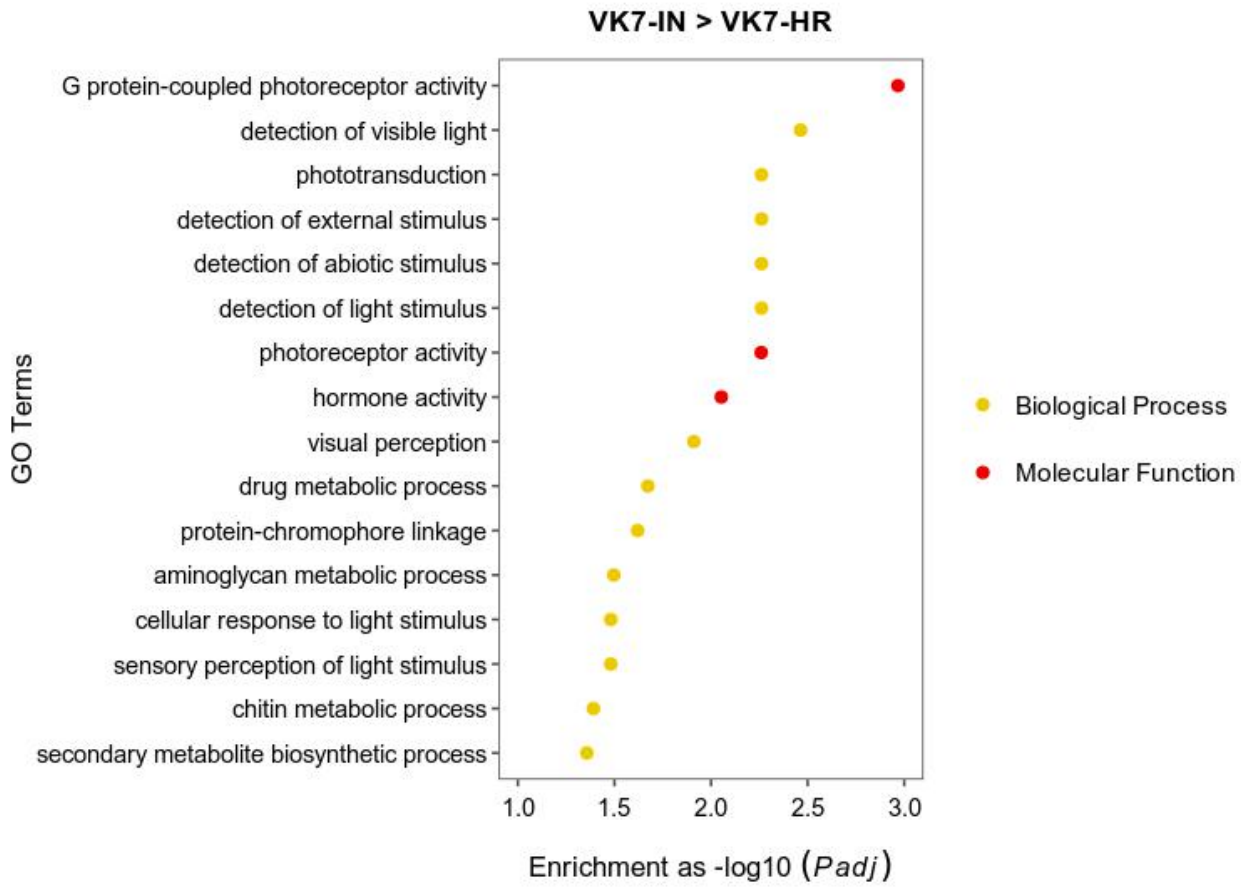

Figure 5. Functional Enrichment Analysis of the up-regulated genes in VK7-HR legs after deltamethrin exposure (VK7-IN versus VK7-HR). Deltamethrin Induction is mediated by GPCR-related signal transduction mechanisms and drug catabolism.

The genome of *A. gambiae* contains 11 genes, six of which belong to the Long Wavelength (LW)-sensitive family, one in each of the Short Wavelength (SW)-sensitive, Ultraviolet (UV)-sensitive and Rh7-like opsin families, and two are characterized as pteropsin (non-visual opsin) genes (Giraldo-Calderón *et al.*, 2017).

The transcriptomic analysis identified seven up-regulated GPCR genes, six of which code for opsins. In particular, four of the six LW-sensitive and the SW-sensitive and UV-sensitive opsin genes were found to be up-regulated after deltamethrin exposure (Table 5), thus supporting the opsin-related functional enrichment found by g:Profiler.

Table 5. Overview of the differentially expressed opsin and arrestin genes after deltamethrin exposure.

		VK7-IN vs VK7-HR	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**
Up-regulated GPCR genes			
ACON012982	GPROP3 long wavelength sensitive opsin	31.30	1.09E-04
ACON012985	GPROP4 long wavelength sensitive opsin	30.08	1.32E-04
ACON006126	GPROP8 ultraviolet wavelength sensitive opsin	29.78	1.38E-05
ACON013149	GPROP1 long wavelength sensitive opsin	29.43	1.40E-04
ACON010134	arrestin-1	28.40	7.76E-05
ACON001161	GPROP6 long wavelength sensitive opsin	25.08	1.94E-06
ACON010089	GPROP9 short wavelength sensitive opsin	20.78	3.99E-04
ACON006263	ARR2 arrestin Arr2-like	19.09	2.38E-05
ACON005002	GPRNNA18 putative GPCR class a orphan receptor 18	7.38	1.21E-07
Down-regulated GPCR genes			
ACON000658	GPRGHP3 putative growth hormone releasing hormone receptor 3 (Pyrokinin 1 receptor ortholog)	-5.20	3.09E-19
ACON008871	GPRFZ4 Putative frizzled 4-like receptor	-4.47	5.30E-28

Opsins are sensory GPCRs with a well-characterized role in sensing light and regulating downstream signaling pathways in insects (Leung & Montell, 2017). Additionally, recent studies in *Drosophila* have provided evidence suggesting several light-independent roles, thus establishing opsins as polymodal sensors with a wide array of cellular and physiological functions (Leung & Montell, 2017).

Most importantly, a recent study demonstrated that an opsin, *NYD-OP7* (Hu *et al.*, 2007; Wu *et al.*, 2004), leads to deltamethrin resistance in *Culex pipiens pallens* by regulating the expression of several cytochrome P450 genes through a phospholipase

C (PLC)-mediated signalling pathway (Zhou *et al.*, 2018). Knockdown of the *NYD-OP7* gene repressed the expression and the enzymatic activity of PLC, thus leading to reduced expression of downstream cytochrome P450 genes and increased susceptibility to deltamethrin (Zhou *et al.*, 2018).

Interestingly, among the six up-regulated opsin genes in VK7-HR legs after deltamethrin exposure, there are three paralogs that belong to the expanded LW-sensitive opsin family (Giraldo-Calderón *et al.*, 2017) and have equal protein sequence identity with *NYD-OP7* (~87%): *GPROP1* (ACON013149), *GPROP3* (ACON012982) and *GPROP4* (ACON012985) (Fig. 6). Two of these paralogs, *GRPOP1* and *GRPOP3*, were also found to be over-expressed on a previous study which characterized the transcriptomic profile of the *A. coluzzii* VK multi-resistant populations (Kwiatkowska *et al.*, 2013). Arrestin is a small protein that interacts with GPCRs and regulates their activity (Merrill *et al.*, 2003). Interestingly, two arrestin genes, *arrestin-1* (ACON006263) and *ARR2* (ACON010134), were also found to be over-expressed after deltamethrin exposure (Table 5).

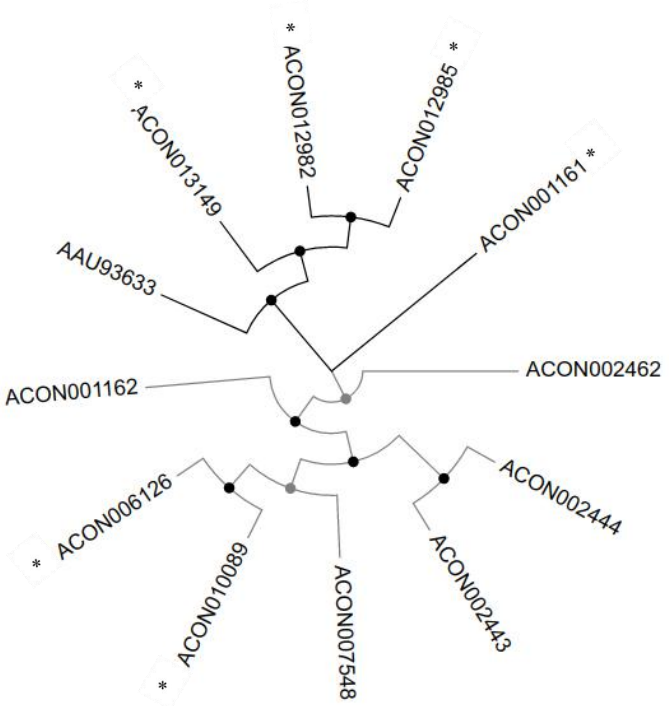

Figure 6. Phylogenetic comparison of *NYD-OP7* (AAU93633) with the 11 opsin genes found in the *A. coluzzii* Ngousso genome. Asterisk is used to denote up-regulated genes after deltamethrin exposure, while the *NYD-OP7*-like clade is noted with black branch color.

Furthermore, recent studies have provided evidence suggesting that GPCR-regulated pathways are implicated in the development of cytochrome P450-mediated resistance in *Culex quinquefasciatus* after exposure to serial concentrations of permethrin (Li *et al.*, 2014; 2015; Li & Liu, 2019). Most importantly, functional experiments in *Spodoptera frugiperda* (*Sf9*) cells demonstrated that GPCR signaling is essential for achieving elevated cytochrome P450 gene expression which subsequently leads to the development of resistance after permethrin exposure (Li & Liu, 2019).

Further, several candidate genes were found to be up-regulated in VK7-HR legs after deltamethrin exposure. These included 17 detoxification genes (4.5%), as well as genes coding for other proteins which have been recently associated with pyrethroid resistance in *Anopheles* mosquitoes, such as odorant binding proteins (OBPs), salivary gland proteins (Table 7) and hexamerin genes (Table 8).

More specifically, three cytochrome P450 genes were over-expressed after deltamethrin exposure: *CYP6P3*, *CYP6M2* and *CYP307A1* (Table 6). Both CYP6P3 and CYP6M2 are capable of metabolizing permethrin and deltamethrin (Edi *et al.*, 2014; Mitchell *et al.*, 2012; Stevenson *et al.*, 2011), while they can also metabolize bendiocarb and DDT respectively (Edi *et al.*, 2014). As a result, CYP6P3 and CYP6M2 have been previously associated with pyrethroid/DDT/carbamate cross-resistance resistance in multiple *A. gambiae* strains (Djouaka *et al.*, 2008; Mitchell *et al.*, 2012; Müller *et al.*, 2008; Edi *et al.*, 2014) as well as in the *A. coluzzii* VK multi-resistant populations (Kwiatkowska *et al.*, 2013). Furthermore, both *CYP6P3* and *CYP6M2* were found to be up-regulated after one-hour exposure to deltamethrin in resistant *A. gambiae* populations from Kenya (Bonizzoni *et al.*, 2015). Interestingly, the overexpression of both these cytochrome P450s has been reported in all aforementioned cases of pyrethroid resistance. Last, *CYP307A1* was recently found to be over-expressed in a pyrethroid-resistant *A. funestus* strain (Mugenzi *et al.*, 2019).

In total, six genes coding for ABC transporters were found to be up-regulated after deltamethrin exposure (Table 6). These genes belong to the ABCC (n = 2), ABCE (n = 1) and ABCG (n = 2) subfamilies. Among the over-expressed ABC transporters in VK7-IN legs, *ABCC2* was recently found to be over-expressed in *A. stephensi* after 12-hr deltamethrin exposure (He *et al.*, 2019), while permethrin exposure leads to expression changes of specific ABC transporter genes in *A. stephensi* (Epis *et al.*, 2014) and in larvae of *A. gambiae* (Mastrantonio *et al.*, 2019).

Table 6. Overview of the differentially expressed detoxification genes after deltamethrin exposure.

		VK7-IN vs VK7-HR	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**
Up-regulated detoxification genes			
ACON012467	ABC transporter type 1	41.10	1.53E-43
ACON002865	cytochrome P450 CYP6P3	9.17	1.30E-09
ACON011852	Short-chain dehydrogenease/reductase	6.69	2.87E-09
ACON000506	ABC transporter family G member 2 - ABCG2	6.31	2.44E-04
ACON005502	dehydrogenase/reductase SDR family member 11 precursor	5.85	2.05E-06
ACON005645	dehydrogenase/reductase SDR family member 11 precursor	5.72	4.71E-42
ACON009463	ABC transporter family G member 11 - ABCG11	5.63	6.23E-49
ACON001775	ABC transporter family C member 2 - ABCC2	5.34	1.03E-59
ACON008212	cytochrome P450 CYP6M2	5.23	1.94E-13
ACON001039	cytochrome P450 CYP307A1	4.93	2.26E-04
ACON029588	ABC transporter family E member 2 - ABCE2	4.84	4.04E-66
ACON005166	Short-chain dehydrogenase/reductase SDR	4.59	6.87E-07
ACON005834	carboxylesterase COEJHE2E	4.52	1.78E-23
ACON005501	dehydrogenase/reductase SDR family member 11 precursor	4.46	6.33E-04
ACON009799	ABC transporter family C member 13 - ABCC13	4.44	5.22E-05
ACON007990	glucosyl/glucuronosyl transferases	4.38	1.46E-05
ACON003904	AIF-like mitochondrial oxidoreductase (Nfrl)	4.07	4.25E-59
Down-regulated detoxification genes			
ACON002210	cytochrome P450 CYP325B1	-11.18	1.10E-06

ACON002205	cytochrome P450 CYP325C2	-8.08	3.21E-16
ACON002138	cytochrome P450 CYP325H1	-5.75	1.85E-41
ACON007802	Argonaute hook domain protein (gawky ortholog)	-5.66	2.03E-29
ACON012155	ABC transporter family A member 6 - ABCA6	-5.22	1.08E-05
ACON028729	ABC transporter family G member 18 - ABCG18	-5.01	4.85E-35

Furthermore, four genes coding for odorant binding proteins (OBPs) were found to be up-regulated after deltamethrin exposure, however none of them has been previously associated with insecticide resistance. In addition, six OBPs which quaintly belong to the D7 family of salivary gland proteins (SGPs), were found to be up-regulated in VK7-HR legs after deltamethrin exposure: four of these belong to the short-form D7 SGPs (D7r1-4), while the remaining two code for the long-form D7 SGPs (D7L1-2) (Arcà *et al.*, 2017). Interestingly, these genes have recently been associated with carbamate (bendiocarb) / pyrethroid (etofenprox, permethrin, deltamethrin, lambda-cyhalothrin) cross-resistance in *A. funestus*, *A. arabiensis*, *A. gambiae* and *A. coluzzii* (Isaacs *et al.*, 2018). Moreover, Isaacs *et al.* (2018) proposed that D7 SGPs act by binding and scavenging insecticides, while they also predicted a mode for bendiocarb binding based on comparative molecular docking analyses. Furthermore, it has been recently proposed that the overexpression of *D7r3* and *D7r4* is associated with pyrethroid resistance in A. funestus (Elanga-Ndille *et al.*, 2019).

In total, 33 genes coding for SGPs were found to be over-expressed and none under-expressed in VK7-HR legs after deltamethrin exposure (Table 7). Strikingly, 9 of them were among the 20 most up-regulated genes in VK7-HR legs after deltamethrin exposure (Table 4). SGPs is a group consisting of functionally and phylogenetically diverse protein families, including proteins with enzymatic activities, as well as proteins implicated in blood feeding (Arcà *et al.*, 2017).

These proteins have been studied and characterized solely based on their expression in mosquito saliva in the context of pathogen-vector-host interactions. As a result, the strong overexpression of multiple SGP genes after deltamethrin exposure cannot be directly associated with a specific function or protein group. Therefore, the main SGP candidates for further investigation in this study remained the D7 family salivary gland protein genes.

Table 7. Overview of the differentially expressed salivary gland protein (SGP) genes after deltamethrin exposure.

		VK7-IN vs VK7-HR	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**
ACON000612	salivary gland protein 1 - SG1	885.27	3.33E-26
ACON000610	salivary gland protein 1-like 6 - SG1f	443.80	1.36E-28
ACON000607	salivary gland protein 1-like 3 - SG1c	435.58	1.06E-20
ACON008216	salivary gland protein 7 - SG7	383.15	7.18E-29
ACON006504	salivary gland protein 2-like - SG2b	235.12	4.19E-42
ACON008215	salivary gland protein 7-like - SG7b	224.95	3.07E-14
ACON000548	salivary gland protein 1-like 2 - SG1b	215.55	4.77E-27
ACON006507	salivary gland protein 3 - SG3	200.59	6.44E-66
ACON000609	salivary gland protein 1-like 5 - SG1e	154.26	5.33E-19
ACON006506	salivary gland protein 2 - SG2	152.95	2.95E-83
ACON005822	putative 56 kDa salivary secreted protein	111.40	9.65E-93
ACON001374	TRIO salivary gland protein trio	105.38	4.21E-21
ACON000611	salivary gland protein 1-like 4 - SG1d	96.70	1.00E-15
ACON013423	salivary gland protein 9 - SG9	85.49	2.83E-45
ACON008307	Putative salivary protein hyp12	62.00	9.71E-19
ACON008306	Putative salivary protein hyp12	49.04	1.46E-13
ACON008283	D7 short form salivary protein - D7r3	35.28	1.90E-05
ACON003841	salivary gland protein 10 - SG10	34.64	1.36E-14
ACON008281	D7 short form salivary protein - D7r4	32.77	1.91E-07
ACON008282	D7 short form salivary protein - D7r2	28.16	5.42E-05

ACON008278	D7 long form salivary protein - D7L1	21.73	1.61E-05
ACON008279	D7 long form salivary protein - D7L2	19.90	5.13E-06
ACON000150	salivary gland protein 6 - SG6	17.92	5.35E-06
ACON008284	D7 short form salivary protein - D7r1	17.33	4.91E-05
ACON006495	Putative secreted salivary basic peptide hyp6.2	15.26	8.44E-08
ACON004334	salivary gland protein 5 - SG5	14.94	5.75E-05
ACON007906	Spink6	14.90	6.63E-08
ACON007907	salivary Kazal 1	11.02	1.26E-04
ACON011026	5' nucleotidase	10.89	2.30E-06
ACON009527	Kazal domain-containing peptide	9.03	1.06E-19
ACON010647	salivary gland protein 8 - SG8	8.50	2.66E-04
ACON006899	basic tail-containing salivary secreted peptide	7.65	4.46E-11
ACON028406	Hyp5.6 salivary secreted peptide	7.15	2.25E-60

Further, five hexamerin genes were found up-regulated after deltamethrin exposure (Table 8). Hexamerins are haemolyph-localized proteins with a well-characterized amino-acid storage role, which are also implicated in the binding and transport of small organic compounds (Willis *et al.*, 2012). The *A. coluzzii* genome contains 8 putative hexamerin genes, five of which were found up-regulated after deltamethrin exposure. Moreover, a recent meta-analysis study has found five of these genes over-expressed in several pyrethroid-resistant *A. arabiensis*, *A. gambiae* and *A. coluzzii* populations (Ingham *et al.*, 2018). Further, knockdown of one of these genes (AGAP001659) led to increased susceptibility in a pyrethroid resistant *A. coluzzii* strain, while no significant changes were observed in terms of mortality when other hexamerin genes were silenced (Ingham *et al.*, 2018).

Based on previous studies (Haunerland & Bowers, 1986; Ingham *et al.*, 2018), the leading hypothesis is that hexamerins are implicated in resistance by binding and sequestering insecticides. The up-regulated hexamerin genes are presented in Table 8.

Table 8. Overview of the differentially expressed hexamerin genes after deltamethrin exposure.

		VK7-IN vs VK7-HR	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**
ACON001659	hexamerin	16.2	2.37E-05
ACON001345	hexamerin	13.52	1.79E-05
ACON001657	hexamerin	13.17	2.98E-04
ACON005766	putative hexamerin	12.1	7.09E-06
ACON005767	putative hexamerin	12	3.49E-04

Finally, three heme peroxidases were found to be highly up-regulated after deltamethrin exposure. A recent study has suggested that the overexpression of heme peroxidases is associated with oxidative stress defense in pyrethroid-resistant *A. arabiensis* and *A. funestus* strains (Oliver & Brooke, 2016), as a homeostatic control mechanism to reduce the heavily elevated oxidative burden due to the activity of insecticide detoxification enzymes.

Overall, the up-regulated genes in VK7-IN legs, suggest that short-term (1-hr) deltamethrin exposure induces an orchestrated, multilayered response which involves enhanced drug detoxification and transport as well as binding of the insecticide (deltamethrin induction). This response could be mediated by over-expressed GPCRs which are responsible for detecting insecticides and coordinating the development of resistance by regulating the expression of many downstream target genes, including genes encoding detoxification enzymes and chemosensory proteins with a pyrethroid-protecting role. More specifically, insecticide binding could have a crucial contribution as a direct and complementary resistance mechanism, by inactivating or by reducing the availability of the insecticide. In particular, insecticide binding may contribute by reducing the insecticide burden, thus increasing the available time for metabolically-based mechanisms to act and inactivate the insecticide before it reaches its target.

4.2 Molecular Evolution of *B. oleae* cytochrome P450 genes

To broaden our understanding over the evolution of cytochrome P450s in the olive fruit fly (*B. oleae*, Diptera: Tephritidae), we manually curated all predicted cytochrome P450 genes in the *B. oleae* genome and subsequently used them to perform comparative genomic analyses.

More specifically, the curated gene models were used to study the phylogenetic relationships of the *B. oleae* cytochrome P450 genes with those from *D. melanogaster* (Diptera: Drosophilidae) and a well-studied tephritid species, the Mediterranean fruit fly (medfly), *C. capitata* (Diptera: Tephritidae).

4.2.1 Manual Curation of cytochrome P450 genes in the olive fly genome

At first, I extracted the predicted P450 gene models based on the conserved protein domain (Pfam: PF00067) which is distinctive for cytochrome P450s. The extracted gene models were then manually curated based on expression data from four developmental stages (egg, larvae, pupa and adult), as well as based on protein sequence similarity with the respective proteins of other closely related tephritid species. In total, 120 cytochrome P450 predicted gene models were manually curated using this process. A detailed report of the manually curated *B. oleae* P450 gene set is presented on Table 9.

Table 10. Detailed summary of the curated cytochrome P450 gene set of *B. oleae*.

Gene ID	Scaffold	Strand	Size (aa)	Best BLAST hit (SwissProt)
g358.t1	LGAM02018580.1	-	87	CYP28D1 probable pseudogene
g716.t1	LGAM02018580.1	-	542	CYP318A1
g774.t1	LGAM02018580.1	-	503	CYP309A2
g775.t1	LGAM02018580.1	-	502	CYP309A1
g776.t1	LGAM02018580.1	-	503	CYP309A1
g777.t1	LGAM02018580.1	+	513	CYP28D1

g910.t1	LGAM02018580.1	+	545	CYP306A1
g911.t1	LGAM02018580.1	-	538	CYP18A1
g1303.t1	LGAM02004382.1	+	515	CYP302A1
g1962.t1	LGAM02006010.1	-	496	CYP6U1
g2121.t1	LGAM02036686.1	-	566	CYP4G15
g3190.t1	LGAM02013110.1	-	499	CYP305A1
g3540.t1	LGAM02013110.1	+	548	CYP314A1 - Ecdysone 20-monooxygenase
g3842.t1	LGAM02036228.1	+	67	CYP49A1 gene fragment
g4062.t1	LGAM02019890.1	+	552	CYP4G1
g4063.t1	LGAM02019890.1	+	552	CYP4G1
g5451.t1	LGAM02038715.1	+	513	CYP6D4
g5488.t1	LGAM02038715.1	+	512	CYP6D5
g5488.t1a	LGAM02038715.1	+	513	CYP6D5
g5584.t1c	LGAM02038715.1	+	508	CYP313A4
g5611.t1	LGAM02038715.1	+	587	CYP313A4
g5612.t1	LGAM02038715.1	+	494	CYP313A4
g5614.t1	LGAM02038715.1	+	498	CYP313A4
g5615.t1	LGAM02038715.1	-	499	CYP313A4
g6103.t1	LGAM02036878.1	-	269	CYP6A13
g6105.t1	LGAM02036878.1	-	500	CYP6A13
g6636.t1	LGAM02008275.1	+	141	CYP12E1 gene fragment
g7078.t1	LGAM02028583.1	+	489	CYP4S3
g7240.t1	LGAM02014826.1	-	280	CYP9H1 probable pseudogene
g7242.t1	LGAM02014826.1	-	503	CYP9b2
g7243.t1	LGAM02014826.1	-	504	CYP9b2
g7490.t1	LGAM02005276.1	-	511	CYP9F2
g8074.t1	LGAM02038611.1	+	502	CYP311A1
g8253.t1	LGAM02038804.1	+	517	CYP4P1
g8254.t1	LGAM02038804.1	+	514	CYP4P1
g8685.t1	LGAM02016086.1	+	553	CYP49A1
g8993.t1	LGAM02003224.1	+	94	CYP12E1 probable pseudogene

g10466.t1	LGAM02011584.1	-	502	CYP4D2
g10467.t1	LGAM02011584.1	-	510	CYP4AE1
g10468.t1	LGAM02011584.1	-	519	CYP4D2
g10472.t1	LGAM02011584.1	-	515	CYP4D1
g10473.t1	LGAM02011584.1	-	524	CYP4D1
g10684.t1	LGAM02025437.1	-	396	CYP12E1-like probable pseudogene
g11326.t1	LGAM02021086.1	+	515	CYP4AC1
g11327.t1	LGAM02021086.1	+	153	CYP4AC1 gene fragment
g12118.t1	LGAM02003756.1	-	502	CYP6A21
g13863.t1	LGAM02035380.1	+	518	CYP4E1
g13864.t1	LGAM02035380.1	-	45	CYP4AE1 probable pseudogene
g13867.t1	LGAM02035380.1	+	527	CYP4e2
g13885.t1	LGAM02035380.1	+	495	CYP6A13
g13886.t1	LGAM02035380.1	+	496	CYP6A23
g13887.t1a	LGAM02035380.1	+	513	CYP6A14
g13887.t1aa	LGAM02035380.1	+	504	CYP6A13
g13888.t1	LGAM02035380.1	+	498	CYP6A23
g13889.t1	LGAM02035380.1	+	514	CYP6A23
g13889.t1a	LGAM02035380.1	+	500	CYP6A20
g13890.t1	LGAM02035380.1	-	504	CYP6A9
g13891.t1	LGAM02035380.1	-	504	CYP6A9
g13892.t1	LGAM02035380.1	-	507	CYP6A21
g13893.t1	LGAM02035380.1	-	504	CYP6A8
g13894.t1	LGAM02035380.1	-	322	CYP6A2 probable pseudogene
g13896.t1	LGAM02035380.1	-	504	CYP6A9
g13898.t1	LGAM02035380.1	+	526	CYP317A1
g13899.t1	LGAM02035380.1	+	531	CYP6A13
g14391.t1	LGAM02035380.1	+	519	CYP6G1
g14392.t1	LGAM02035380.1	-	520	CYP6G2
g14393.t1	LGAM02035380.1	+	519	CYP6G1
g14394.t1	LGAM02035380.1	+	519	CYP6G1
g14395.t1	LGAM02035380.1	-	95	CYP6G2 gene fragment

g14396.t1	LGAM02035380.1	+	527	CYP6T3
g14397.t1	LGAM02035380.1	-	535	CYP6G2-like
g14398.t1	LGAM02035380.1	-	535	CYP6G2-like
g15072.t1	LGAM02023717.1	-	134	CYP313a4 gene fragment
g16378.t1	LGAM02016867.1	+	490	CYP4d8
g16379.t1	LGAM02016867.1	+	515	CYP316A1
g16413.t1	LGAM02016867.1	+	508	CYP28D1
g16559.t1	LGAM02016867.1	-	123	CYP4D14 probable pseudogene
g17537.t1	LGAM02021800.1	+	261	CYP12E1 probable pseudogene
g17678.t1	LGAM02025750.1	+	519	CYP6V1
g18109.t1	LGAM02016039.1	-	213	CYP12E1 gene fragment
g18207.t1	LGAM02009281.1	-	514	CYP309a2
g18801.t1	LGAM02033874.1	-	548	CYP313B1
g18868.t1	LGAM02033874.1	+	559	CYP315A1
g19257.t1	LGAM02033874.1	+	501	CYP313A4
g19258.t1	LGAM02033874.1	-	518	CYP313A4
g19260.t1	LGAM02033874.1	+	490	CYP313A4
g19268.t1	LGAM02033874.1	+	518	CYP313A4
g19289.t1	LGAM02033874.1	+	525	CYP307A1-like
g19511.t1	LGAM02025756.1	-	506	CYP6A2
g20401.t1	LGAM02007670.1	-	534	CYP4C3
g20466.t1	LGAM02007670.1	+	532	CYP12E1
g20467.t1	LGAM02007670.1	+	533	CYP12E1
g20468.t1	LGAM02007670.1	+	532	CYP12E1
g20469.t1	LGAM02007670.1	+	534	CYP12E1
g20470.t1	LGAM02007670.1	+	533	CYP12E1
g20499.t1	LGAM02007670.1	+	514	CYP304A1
g20500.t1	LGAM02007670.1	+	514	CYP304A1
g21105.t1	LGAM02012403.1	-	515	CYP6T1
g21753.t1	LGAM02005801.1	+	492	CYP310A1
g22246.t1	LGAM02026273.1	+	543	CYP12C1
g22260.t1	LGAM02026273.1	+	551	CYP12B1

g22261.t1	LGAM02026273.1	+	542	CYP12B4
g23048.t1	LGAM02029766.1	-	506	CYP308A1
g23659.t1	LGAM02025757.1	+	501	CYP303A1
g24264.t1	LGAM02021877.1	+	581	CYP301A1
g24537.t1	LGAM02013164.1	+	541	CYP12A4
g24537.t1a	LGAM02013164.1	+	556	CYP12A4
g24538.t1	LGAM02013164.1	+	556	CYP12A5
g24538.t1a	LGAM02013164.1	+	557	CYP12A4

4.2.2 Phylogenetic analysis of the *B. oleae* cytochrome P450s

The process of manual curation identified 94 intact cytochrome P450 genes, 10 probable pseudogenes and six gene fragments. Identified gene fragments were subsequently excluded from the phylogenetic comparions. In addition, eight gene clusters consisting of three or more consecutive cytochrome P450 genes were identified by this analysis.

These gene clusters result from the expansion of particular cytochrome P450 gene families and will be further discussed in this study. Next, we aligned the amino-acid sequences of the curated cytochrome P450 gene sets from *B. oleae* (n = 104), *D. melanogaster* (n = 89) and *C. capitata* (n = 115) and based on this *all-vs-all* alignment we built a comprehensive phylogenetic tree. The tree was rooted using the human *CYP51A1* gene as an outgroup (Figure 7).

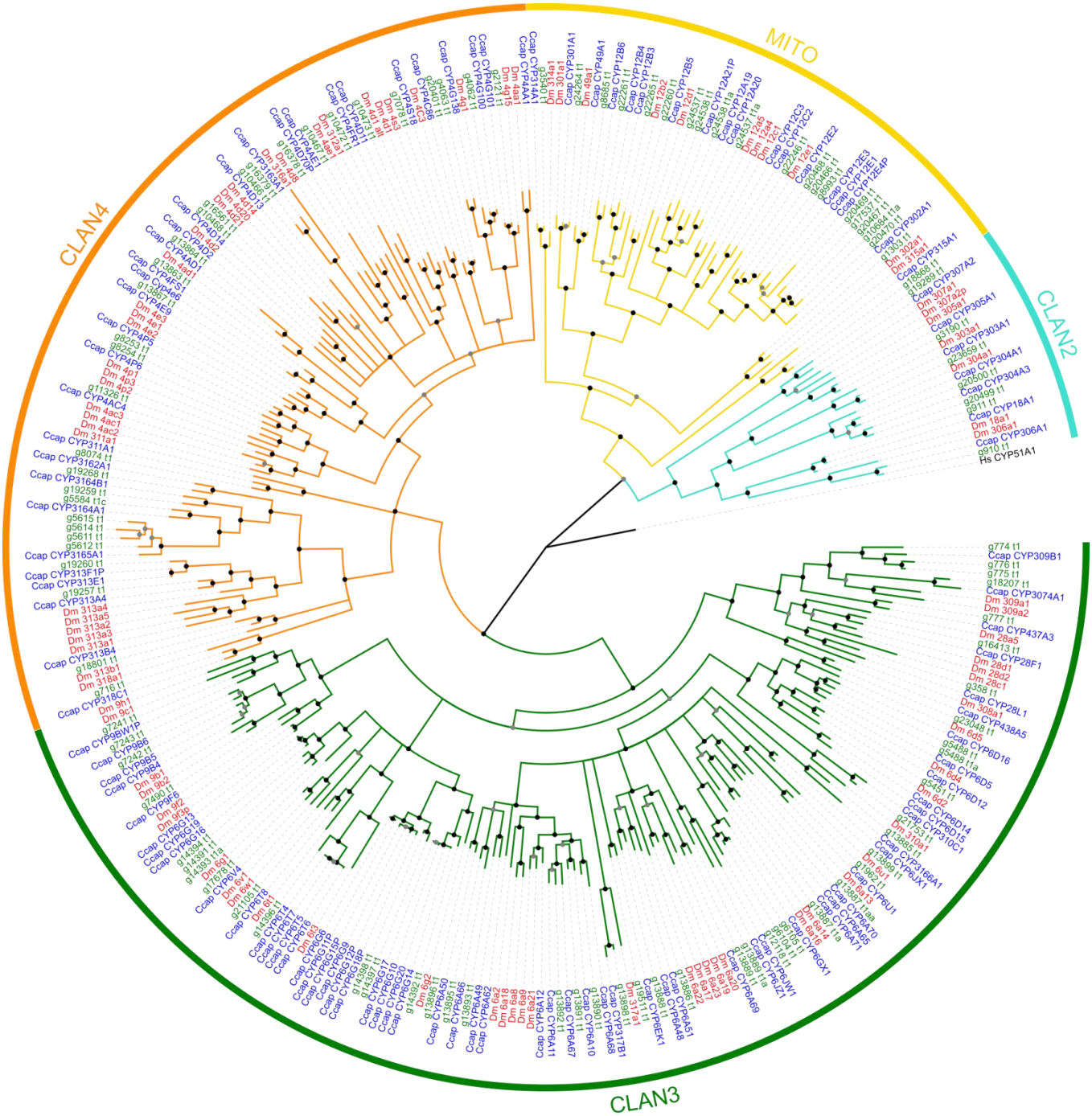

Figure 7. Maximum likelihood phylogenetic tree of cytochrome P450 amino acid sequences from the olive fruit fly *Bactrocera oleae* (names in green), *Drosophila melanogaster* (names in red) and the medfly *Ceratitis capitata* (names in blue). The tree was rooted using the human *CYP51A1* gene as an outgroup. Nodes with less than 50% bootstrap support were collapsed into multifurcating nodes, while nodes with bootstrap support between 50% and 75% are indicated with gray circles and nodes with bootstrap support greater than 75% are indicated with black circles.

4.2.3 Comparing P450 copy number variation between *B. oleae*, *D. melanogaster* and *C. capitata*

The *B. oleae* cytochrome P450 gene set is composed of 94 genes, 10 probable pseudogenes and six gene fragments. These numbers are greater compared to *D. melanogaster* (88 genes and one pseudogene), but less than those of *C. capitata*, (103 genes, nine pseudogenes and three gene fragments).

Furthermore, the *B. oleae* cytochrome P450 genes belong to the four clans that insects typically possess (Feyereisen, 2012); mito clan (21 genes), clan 2 (7 genes), clan 3 (45 genes) and clan 4 (31 genes) (Table 1). *B. oleae* and *D. melanogaster* have an almost equal gene number of clan 2 and clan 4 cytochrome P450s (Table 10). Moreover, CYP3 is the most abundant P450 clan in the *B. oleae* genome with 45 genes and it is larger than clan 3 of *D. melanogaster*, where 37 genes are found (Table 10). However, it is considerably less expanded in comparison to its corresponding clan in *C. capitata*, which is composed of 59 genes (Table 10). Notably, *B. oleae* possesses 21 mito clan genes, twice as many compared to *D. melanogaster* (n = 11) (Table 10).

Table 10. Number of the *B. oleae*, *D. melanogaster* and *C. capitata* cytochrome P450 genes in each of the four insect P450 clans.

Number of genes					
Species	**Mito Clan**	**Clan 2**	**Clan 3**	**Clan 4**	**Total**
B. oleae	21	7	45	31	104
D. melanogaster	11	7	37	32	87
C. capitata	18	7	59	31	115

In further detail, the increased number of cytochrome P450 genes in the olive fruit fly genome compared to *D. melanogaster* is largely due to expansions in the CYP12 (mito clan) and CYP6 (clan 3) families. The expansion of the same gene families has been also suggested to account for the elevated number of cytochrome P450 genes found in the *C. capitata* genome (Papanicolaou *et al.*, 2016). Subsequently, we focused on gene copy number variation within each of the cytochrome P450 clans.

4.2.3.1 Mito clan

Mitochondrial P450 gene families have a relatively stable gene copy number across insect genomes (Feyereisen, 2012). As an exception, CYP12 is a rapidly evolving gene family that shows considerable copy number fluctuations between insect species, thus resembling the more variable P450 clans 3 and 4 (Feyereisen, 2012). *B. oleae* shows an expansion in the CYP12A and CYP12E subfamilies compared to both *D. melanogaster* & *C. capitata* (Table 11).

The olive fruit fly CYP12A subfamily is composed of four genes, while three and two CYP12A genes are found in the genomes of *C. capitata* and *D. melanogaster*, respectively (Table 1). Interestingly, three of the four CYP12A genes in the *B. oleae* genome have one-to-one orthologs with *C. capitata* (Fig. 7) whereas the fourth (g24537.t1) is sister to all other CYP12A genes, thereby suggesting that the CYP12A gene subfamily expansion has occurred before the divergence of the two species. Furthermore, all four CYP12A genes are located in tandem and comprise a 14.2-kb gene cluster in the *B. oleae* genome, thus indicating that they have emerged by gene duplication. A similar pattern of genetic organization has been observed for two CYP12A genes in the *C. capitata* genome (Papanicolaou *et al.*, 2016).

Members of the CYP12A family have been previously implicated in insecticide resistance in Diptera. More specifically, *CYP12A4* has been shown to confer resistance to lufenuron, a chitin biosynthesis inhibitor (benzoylurea pesticide), in a natural population of *D. melanogaster* (Bogwitz *et al.*, 2005). In addition, an interesting case of two *CYP12A4*/*CYP12A5* chimeric genes occuring via interparalog exchange in the genome of *D. melanogaster* lines from the DGRP (Mackay *et al.*, 2012) has been recently associated with highly elevated levels of luneferon resistance (Good *et al.*, 2014).

The olive fly CYP12E subfamily is also expanded in both the olive fly and the medfly, compared to *D. melanogaster*. More specifically, only one *CYP12E1* gene is found in the *D. melanogaster* genome, while *B. oleae* and *C. capitata* possess several paralogs of *CYP12E1* and *CYP12E2* (Table 11). *CYP12E2* is a *CYP12E1* duplicate which has originated before the split of *B. oleae* and *C. capitata* (Fig. 7).

Additionally, *B. oleae* and *C. capitata* share two CYP12E ortholog groups probably emerging from several *CYP12E1* and *CYP12E2* duplication events that have occurred independently after the split of the two species (Fig. 7). Interestingly, all five CYP12E intact genes (three *CYP12E1* and two *CYP12E2* paralogs), are tandemly located at a 26.8-kb genomic cluster, while the three CYP12E pseudogenes and the *CYP12E1* gene fragment are dispersed in the *B. oleae* genome. This fact could possibly indicate that *cis* co-regulation of expression is a decisive factor for the preservation of duplicated genes in the CYP12E subfamily.

The cytochrome P450s belonging to the CYP12E subfamily are not functionally characterized. Nevertheless, the expansion of this gene subfamily in the olive fruit fly

genome and the adjacent genome topology of this expansion could indicate its involvement in environmental response mechanisms related to the *B. oleae*-specific lifestyle and the environmental stresses associated with its particular ecological niche.

Table 11. Number of the mitochondrial (mito clan) cytochrome P450 genes found in the *D. melanogaster*, *C. capitata* and *B. oleae* genomes.

Species	Genes	Pseudogenes
B. oleae	4 *CYP12A*, 3 *CYP12E1*, 2 *CYP12E1*	2 *CYP12E1*, 1 *CYP12E2*
D. melanogaster	*CYP12A4, CYP12A5, CYP12E1*	-
C. capitata	3 *CYP12A*, 2 *CYP12E1*, 1 *CYP12E2*	1 *CYP12E1*

4.2.3.2 Clan 2

B. oleae possesses an equal number of clan 2 genes (n =7) with *D. melanogaster* and *C. capitata* (Table 1). Furthermore, all CYP2 clan P450 genes, except for *CYP304A* and *CYP307A*, have a 1:1:1 ortholog relationship between *B. oleae*, *C. capitata* and *D. melanogaster* (Fig. 7). This pattern can be attributed to the conserved role of several of these genes in insect development and physiology, such as *CYP306A1* and *CYP307A2* which are involved in the biosynthesis of 20-hydroxyecdysone (20H) (Chung *et al.*, 2009). The essentiallity of clan 2 P450s in core physiological processes has been previously proposed to account for the relatively conserved number of CYP2 gene copies across insects (Feyereisen, 2012).

4.2.3.3 Clan 3

Most clan 3 gene copy number variations between *B. oleae* and *D. melanogaster* are found in members of the CYP6 family. CYP6 is a rapidly proliferating gene family that is frequently occurring in gene clusters and it exhibits high evolvability (Feyereisen, 2012). In addition, members of several CYP6 subfamilies, including CYP6A, CYP6G and CYP6D, have been heavily associated with insecticide resistance and xenobiotic metabolism in higher Diptera species (Feyereisen, 2012).

The olive fly CYP6 family (30 genes, two pseudogenes) has a greater level of

expansion compared to *D. melanogaster* (23 genes), largely as a result of the notable expansion of the CYP6A and CYP6G subfamilies (Fig. 7). However, the *B. oleae* CYP6 family is noticeably less expanded compared to *C. capitata* (40 genes, 4 pseudogenes and 1 gene fragment).

More specifically, the olive fly CYP6A gene family (16 genes, two pseudogenes) is slightly larger than in *C. capitata* (15 genes) and *D. melanogaster* (13 genes) (Fig. 7). In addition, most CYP6A genes (12 genes and one pseudogene) are located in a 54.3-kb gene cluster in the *B. oleae* genome, which is composed of 15 consecutive cytochrome P450 genes and is flanked by orthologs of the *D. melanogaster mtt* (upstream) and *kank* (downstream) genes. Furthermore, most *B. oleae* CYP6A genes have 1:1 orthologs with *C. capitata*, thereby demonstrating that this expansion has mostly occurred before the divergence of the two species (Fig. 7). This finding is further supported by the existence of a related gene cluster in the genome of *C. capitata*. This cluster consists of 18 genes, most of which belong to the CYP6A family (n=13), and is also flanked by orthologs of the *kank* and *mtt* genes (Papanicolaou *et al.*, 2016).

Cytochrome P450s belonging to the CYP6A subfamily are well-known to be implicated in insecticide resistance of several pests of medical and agricultural importance (Battlay *et al.*, 2018; Feyereisen, 2012; Mitchell *et al.*, 2012; Riveron *et al.*, 2013; Zhu *et al.*, 2010). Interestingly, a cytochrome P450 gene which was previously found to be over-expressed in resistant *B. oleae* populations from Greece (unpublished data) is the 1:1 ortholog of *CYP6A51* from *C. capitata*. *CYP6A51* was initially found to be over-expressed in resistant field populations from Spain and was therefore associated with λ-cyhalotrin resistance in *C. capitata* (Arouri *et al.*, 2015). Furthermore, a recent functional study has demonstrated that CYP6A51 is capable of metabolizing both λ-cyhalotrin and deltamethrin, thus confering pyrethroid resistance in transgenic *D. melanogaster* strains over-expressing this gene (Tsakireli *et al.*, 2019).

The olive fly CYP6G subfamily (six genes, one gene fragment) is expanded compared to *D. melanogaster* (two genes), even though it has contracted compared to *C. capitata* (10 genes, four pseudogenes). More specifically, each of the two *D. melanogaster* CYP6G genes, *CYP6G1* and *CYP6G2*, has expanded in the olive fly and the medfly genomes (Fig. 7). Interestingly, *B. oleae* and *C. capitata* share an additional CYP6G ortholog group that is absent from *D. melanogaster* (Fig. 7). This ortholog group derives from a *CYP6G2* duplicate that emerged in a common ancestor of the two species after the split with the *D. melanogaster* clade (Fig. 7). Remarkably, the CYP6G subfamily has been independently expanded in *B. oleae* and *C. capitata* after the split of the two species, thus leading to the formation of the three distinct CYP6G ortholog groups (Fig. 7). The olive fly CYP6G genes are located in a 32-kb genomic cluster that contains all six *CYP6G* and one *CYP6T* genes. The two genes belonging to the *CYP6G2*-like ortholog group, which is absent from *D. melanogaster*, are located at the end of this cluster. Three related, tandem located *CYP6G*/*CYP6T* gene clusters are found in the *C. capitata* genome (Papanicolaou *et al.*, 2016).

The CYP6G subfamily has been implicated in multiple cases of insecticide resistance in dipteran species. More specifically, previous studies in *D. melanogaster* have demonstrated that CYP6G1 is capable of conferring cross-resistance to DDT (Daborn *et al.*, 2002), neonicotinoids, carbamates, organophosphates and insect growth regulators (Le Goff & Hilliou, 2017). In addition, an ortholog of *CYP6G1* has been recently shown to confer resistance to multiple insecticides in the Australian sheep blowfly, *Lucilia cuprina* (Traylor *et al.*, 2017). Further, CYP6G2 has been associated with neocotinoid and organophosphate cross-resistance in *D. melanogaster* (Daborn *et al.*, 2007). In this context, the expansion of the CYP6G subfamily could provide *B. oleae* with a great potential for developing insecticide resistance. However, it should be noted that there is no clear and direct correlation between the number of detoxification genes and the capacity for metabolizing xenobiotics and thus developing insecticide resistance (Ioannidis *et al.*, 2017). At last, the olive fruit fly CYP6D and CYP6T subfamilies have an equal gene number with *D. melanogaster*, consisting of three and two genes respectively (Fig. 7). Two of the three CYP6T genes are located in tandem at the *B. oleae* genome.

4.2.3.4 Clan 4

B. oleae possesses a comparable number of clan 4 cytochrome P450 genes with *D. melanogaster* and *C. capitata* (Table 1). However, it has a smaller number of genes belonging to the CYP4AC, CYP4AD, CYP4D, CYP4E and CYP4P subfamilies compared to their respective subfamilies in *D. melanogaster* and *C. capitata* (Fig. 7).

Furthermore, the CYP313A subfamily has a greater level of expansion in *B. oleae* (eight genes, one pseudogene and one gene fragment) compared to both *C. capitata* (six genes, one pseudogene) and *D. melanogaster* (five genes) (Fig. 7). The expansion of this subfamily in the *B. oleae* genome is based on two independent expansion events that occurred after the split with the *D. melanogaster* clade. More specifically, the first expansion event occurred in the common ancestor of *B. oleae* and *C. capitata*, as there are four *CYP313A* genes in the olive fruit fly genome with closely related orthologs in *C. capitata* (Fig. 7). Three of these genes are located in a 22.2-kb gene cluster, while the fourth *CYP313A* gene is located approximately 125 kb away. The second expansion event has occurred after the split of *B. oleae* and *C. capitata* (Fig. 7) and the four genes responsible are located in tandem at a distinct 70-kb gene cluster in the olive fruit fly genome. Moreover, the CYP313A subfamily has also been independently expanded by gene duplication events in the lineage leading to *D. melanogaster* (Fig. 7).

The CYP313A subfamily is not functionally characterized, but a recent study has found *CYP313A4* to be up-regulated in a DDT-resistant *D. melanogaster* strain (Seong *et al.*, 2018). The two-step expansion of this subfamily could be related to the environmental stress associated with a tephritid-specific lifestyle.

Conclusions

Insecticide resistance imposes great challenges in public health and food security on a global scale. Understanding how insects develop resistance to insecticides provides the basis for developing effective and sustainable resistance management strategies. This study investigated molecular mechanisms of insecticide resistance by using genomic and transcriptomic data. More specifically, we focused on insecticide resistance in the major malaria vector *Anopheles coluzzii* and the olive fly *Bactrocera oleae*.

At the first part of the study, we performed a leg-specific transcriptomic study of the multi-resistant *A. coluzzii* VK7 strain. In brief, the transcriptomic data suggest that the constitutive resistance phenotype is primarily mediated by detoxification enzymes, mainly cytochrome P450s, with a minor contribution of cuticular resistance. Moreover, we compared leg-specific transcriptomic data from deltamethrin-exposed and -unexposed highly resistant *A. coluzzii* VK7 mosquitoes. Since the mosquito leg is the first contact point of insecticides, this analysis provided valuable insights regarding the development of resistance after exposure to the insecticide. The results of this comparison suggest that deltamethrin exposure induces a complex response that is mediated by detoxification enzymes and chemosensory proteins with a pyrethroid-protecting role. Particular opsins (GPCRs) are possibly responsible for regulating the development of resistance after detection of the insecticide. Studying the phylogeny of these proteins across multiple mosquito species and inferring their genome localization will be particularly useful in the context of delineating novel regulatory mechanisms that regulate the development of resistance after insecticide exposure.

At the second part of this study, we studied the evolution of the olive fly cytochrome P450s. More specifically, we manually curated the total *B. oleae* P450ome and subsequently performed phylogenetic comparisons with those of *Drosophila melanogaster* and *Ceratitis capitata*. In summary, *B. oleae* shows expansions in the mitochondrial (mito) clan and the CYP3 clan compared to *D. melanogaster*. More specifically, *B. oleae* demonstrates a lineage-specific expansion of the CYP12A and CYP12E subfamilies. The enlargement of the CYP3 clan is caused by expansions of the CYP6A and CYP6G families, which have mainly occured before and after the split with *C. capitata* respectively.

This study provides a good framework for the implementation of future transcriptomic and comparative genomic studies in tephritid flies, with a particular interest in olive fruit fly biology and evolution. More specifically, by making use of the available transcriptomic data from field-caught resistant olive fly populations, as well as from olive flies fed on ripe and unripe olives we can identify novel detoxification candidate genes which are implicated in the metabolism of xenobiotics and the enriched phenolic compounds found in the olive mesocarp.

Further, by comparing gene copy number variation of the total detoxification enzyme families (including cytochrome P450s, GSTs, CCEs, UGTs and ABC transporters)

between *B. oleae*, *B. dorsalis*, *C. capitata* and *D. melanogaster* (possibly include additional Diptera species), we can address important questions regarding the ecology and evolution of tephritid flies on two levels of biological order. First, we will identify genetic signatures associated with generic features of the tephritid flies ecology and their common evolutionary origin. Second, by comparing three tephritids with constrasting trophic niches we will uncover detoxification gene family expansions and contractions which are relevant to the particular life-style and the plant-feeding capacity of each species. These comparative genomic analyses will also provide crucial insights on the resistance potential of these agricultural pests.

References

Adolfi, A., Poulton, B., Anthousi, A., Macilwee, S., Ranson, H., & Lycett, G. J. (2019). Functional genetic validation of key genes conferring insecticide resistance in the major African malaria vector, *Anopheles* gambiae. *Proceedings of the National Academy of Sciences*, *116*(51), 25764-25772.

Antwi, F. B., & Reddy, G. V. P. (2015). Toxicological effects of pyrethroids on non-target aquatic insects. *Environmental Toxicology and Pharmacology*, *40*(3), 915-923.

Arcà, B., Lombardo, F., Struchiner, C. J., & Ribeiro, J. M. (2017). Anopheline salivary protein genes and gene families: an evolutionary overview after the whole genome sequence of sixteen *Anopheles* species. *BMC Genomics*, *18*(1), 153.

Arouri, R., Le Goff, G., Hemden, H., Navarro-Llopis, V., M'saad, M., Castañera, P., Feyereisen, R., Hernández-Crespo, P., & Ortego, F. (2015). Resistance to lambda-cyhalothrin in Spanish field populations of *Ceratitis capitata* and metabolic resistance mediated by P450 in a resistant strain. *Pest Management Science*, *71*(9), 1281-1291.

Balabanidou, V., Grigoraki, L., & Vontas, J. (2018). Insect cuticle: a critical determinant of insecticide resistance. *Current Opinion in Insect Science*, *27*, 68-74.

Balabanidou, V., Kampouraki, A., MacLean, M., Blomquist, G. J., Tittiger, C., Juárez, M. P., Mijailovsky, S. J., Chalepakis, G., Anthousi, A., Lynd, A., Antoine, S., Hemingway, J., Ranson, H., Lycett, G. J., & Vontas, J. (2016). Cytochrome P450 associated with insecticide resistance catalyzes cuticular hydrocarbon production in *Anopheles gambiae*. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(33), 9268-9273.

Balabanidou, V., Kefi, M., Aivaliotis, M., Koidou, V., Girotti, J. R., Mijailovsky, S. J., Juárez, M. P., Papadogiorgaki, E., Chalepakis, G., Kampouraki, A., Nikolaou, C., Ranson, H., & Vontas, J. (2019). Mosquitoes cloak their legs to resist insecticides. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1907), 20191091.

Battlay, P., Leblanc, P. B., Green, L., Garud, N. R., Schmidt, J. M., Fournier-Level, A., & Robin, C. (2018). Structural Variants and Selective Sweep Foci Contribute to Insecticide Resistance in the Drosophila Genetic Reference Panel. *G3*, *8*(11), 3489-3497.

Bayega, A., Djambazian, H., Tsoumani, K. T., Gregoriou, M.-E., Sagri, E., Drosopoulou, E., Mavragani-Tsipidou, P., Giorda, K., Tsiamis, G., Bourtzis, K.,

Oikonomopoulos, S., Dewar, K., Church, D. M., Papanicolaou, A., Mathiopoulos, K. D., & Ragoussis, J. (2020). De novo assembly of the olive fruit fly (*Bactrocera oleae*) genome with linked-reads and long-read technologies minimizes gaps and provides exceptional Y chromosome assembly. *BMC Genomics*, *21*(1), 259.

Bhatt, S., Weiss, D. J., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K., Moyes, C. L., Henry, A., Eckhoff, P. A., Wenger, E. A., Briët, O., Penny, M. A., Smith, T. A., Bennett, A., Yukich, J., Eisele, T. P., Griffin, J. T., Fergus, C. A., Lynch, M., Lindgren, F., Cohen, J. M., Murray, C. L. J., Smith, D. L., Hay, S. I., Cibulskis, R. E., & Gething, P. W. (2015). The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature*, *526*(7572), 207-211.

Bogwitz, M. R., Chung, H., Magoc, L., Rigby, S., Wong, W., O'Keefe, M., McKenzie, J. A., Batterham, P., & Daborn, P. J. (2005). CYP12A4 confers lufenuron resistance in a natural population of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(36), 12807-12812.

Bonizzoni, M., Ochomo, E., Dunn, W. A., Britton, M., Afrane, Y., Zhou, G., Hartsel, J., Lee, M. C., Xu, J., Githeko, A., Fass, J., & Yan, G. (2015). RNA-seq analyses of changes in the *Anopheles* gambiae transcriptome associated with resistance to pyrethroids in Kenya: identification of candidate-resistance genes and candidate-resistance SNPs. *Parasites & Vectors*, *8*, 474.

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525-527.

Calla, B., Noble, K., Johnson, R. M., Walden, K., Schuler, M. A., Robertson, H. M., & Berenbaum, M. R. (2017). Cytochrome P450 diversification and hostplant utilization patterns in specialist and generalist moths: Birth, death and adaptation. *Molecular Ecology*, *26*(21), 6021-6035.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, *10*, 421.

Cao, X., Gulati, M., & Jiang, H. (2017). Serine protease-related proteins in the malaria mosquito, *Anopheles* gambiae. *Insect Biochemistry and Molecular Biology*, *88*, 48-62.

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics (Oxford, England)*, *25*(15), 1972-1973.

Chen, H., & Boutros, P. C. (2011). VennDiagram: a package for the generation of

highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*, *12*(1), 35.

Chiu, T. L., Wen, Z., Rupasinghe, S. G., & Schuler, M. A. (2008). Comparative molecular modeling of *Anopheles* gambiae CYP6Z1, a mosquito P450 capable of metabolizing DDT. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(26), 8855-8860.

Chung, H., Sztal, T., Pasricha, S., Sridhar, M., Batterham, P., & Daborn, P. J. (2009). Characterization of *Drosophila melanogaster* cytochrome P450 genes. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(14), 5731-5736.

Cinege, G., Zsámboki, J., Vidal-Quadras, M., Uv, A., Csordás, G., Honti, V., Gábor, E., Hegedűs, Z., Varga, G., Kovács, A. L., Juhász, G., Williams, M. J., Andó, I., & Kurucz, É. (2017). Genes encoding cuticular proteins are components of the Nimrod gene cluster in Drosophila. *Insect Biochemistry and Molecular Biology*, *87*, 45-54.

Daane, K. M., & Johnson, M. W. (2010). Olive Fruit Fly: Managing an Ancient Pest in Modern Times. *Annual Review of Entomology*, *55*(1), 151-169.

Daborn, P. J., Lumb, C., Boey, A., Wong, W., ffrench-Constant, R. H., & Batterham, P. (2007). Evaluating the insecticide resistance potential of eight *Drosophila melanogaster* cytochrome P450 genes by transgenic over-expression. *Insect Biochemistry and Molecular Biology*, *37*(5), 512–519.

Daborn, P. J., Yen, J. L., Bogwitz, M. R., Le Goff, G., Feil, E., Jeffers, S., Tijet, N., Perry, T., Heckel, D., Batterham, P., Feyereisen, R., Wilson, T. G., & ffrench-Constant, R. H. (2002). A single P450 allele associated with insecticide resistance in *Drosophila*. *Science* (New York, N.Y.), *297*(5590), 2253-2256.

David, J.-P., Ismail, H. M., Chandor-Proust, A., & Paine, M. J. I. (2013). Role of cytochrome P450s in insecticide resistance: impact on the control of mosquito-borne diseases and use of insecticides on Earth. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1612), 20120429.

Dermauw, W., & Van Leeuwen, T. (2014). The ABC gene family in arthropods: Comparative genomics and role in insecticide transport and resistance. *Insect Biochemistry and Molecular Biology*, *45*, 89-110.

Djouaka, R. F., Bakare, A. A., Coulibaly, O. N., Akogbeto, M. C., Ranson, H., Hemingway, J., & Strode, C. (2008). Expression of the cytochrome P450s, CYP6P3 and CYP6M2 are significantly elevated in multiple pyrethroid resistant populations of *Anopheles* gambiae s.s. from Southern Benin and Nigeria. *BMC Genomics*, *9*, 538.

Dong, K. (2007). Insect sodium channels and insecticide resistance. *Invertebrate Neuroscience*, *7*(1), 17-30.

Douris, V., Steinbach, D., Panteleri, R., Livadaras, I., Pickett, J. A., Van Leeuwen, T., Nauen, R., & Vontas, J. (2016). Resistance mutation conserved between insects and mites unravels the benzoylurea insecticide mode of action on chitin biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(51), 14692-14697.

Edi, C. V., Djogbénou, L., Jenkins, A. M., Regna, K., Muskavitch, M. A., Poupardin, R., Jones, C. M., Essandoh, J., Kétoh, G. K., Paine, M. J., Koudou, B. G., Donnelly, M. J., Ranson, H., & Weetman, D. (2014). CYP6 P450 enzymes and ACE-1 duplication produce extreme and multiple insecticide resistance in the malaria mosquito *Anopheles gambiae*. *PLoS Genetics*, *10*(3), e1004236.

Elanga-Ndille, E., Nouage, L., Binyang, A., Assatse, T., Tene-Fossog, B., Tchouakui, M., Nguete Nguiffo, D., Irving, H., Ndo, C., Awono-Ambene, P., & Wondji, C. S. (2019). Overexpression of Two Members of D7 Salivary Genes Family is Associated with Pyrethroid Resistance in the Malaria Vector *Anopheles Funestus* s.s. but Not in *Anopheles Gambiae* in Cameroon. *Genes*, *10*(3), 211.

Epis, S., Porretta, D., Mastrantonio, V., Comandatore, F., Sassera, D., Rossi, P., Cafarchia, C., Otranto, D., Favia, G., Genchi, C., Bandi, C., & Urbanelli, S. (2014). ABC transporters are involved in defense against permethrin insecticide in the malaria vector *Anopheles* stephensi. *Parasites & Vectors*, *7*, 349.

Feyereisen, R. (2005). Insect cytochrome P450. In L. I. Gilbert, K. Iatrou, & S. S. Gill (eds.). *Comprehensive Molecular Insect Science*, pp. 1-77. Oxford, UK: Elsevier.

Feyereisen, R. (2012). Insect CYP Genes and P450 Enzymes. In: Gilbert, L. I. (*ed.*). *Insect Molecular Biology and Biochemistry*, pp. 237-316. Oxford: Elsevier.

Feyereisen, R., Dermauw, W., & Van Leeuwen, T. (2015). Genotype to phenotype, the molecular and physiological dimensions of resistance in arthropods. *Pesticide Biochemistry and Physiology*, *121*, 61-77.

Feyereisen, R. (2020). Origin and evolution of the CYP4G subfamily in insects, cytochrome P450 enzymes involved in cuticular hydrocarbon synthesis. *Molecular Phylogenetics and Evolution*, 106695.

Fotakis, E. A., Mastrantonio, V., Grigoraki, L., Porretta, D., Puggioli, A., Chaskopoulou, A., Osório, H., Weill, M., Bellini, R., Urbanelli, S., & Vontas, J. (2020). Identification and detection of a novel point mutation in the Chitin Synthase gene of *Culex* pipiens associated with diflubenzuron resistance. *PLOS Neglected*

Tropical Diseases, *14*(5), e0008284.

ffrench-Constant, R. H. (2014). Insecticide resistance comes of age. *Genome Biology*, *15*(2), 106.

Giraldo-Calderón, G. I., Emrich, S. J., MacCallum, R. M., Maslen, G., Dialynas, E., Topalis, P., Ho, N., Gesing, S., VectorBase Consortium, Madey, G., Collins, F. H., & Lawson, D. (2015). VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Research*, *43*(Database issue), D707–D713.

Giraldo-Calderón, G. I., Zanis, M. J., & Hill, C. A. (2017). Retention of duplicated long-wavelength opsins in mosquito lineages by positive selection and differential expression. *BMC Evolutionary Biology*, *17*(1), 84.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., Henschel, R., LeDuc, R. D., Friedman, N., & Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494-1512.

Haunerland, N. H., & Bowers, W. S. (1986). Binding of insecticides to lipophorin and arylphorin, two hemolymph proteins of *Heliothis zea*. *Archives of Insect Biochemistry and Physiology*, *3*, 87-96.

Hawkins, N. J., Bass, C., Dixon, A., & Neve, P. (2018). The evolutionary origins of pesticide resistance. *Biological reviews of the Cambridge Philosophical Society*, *94*(1), 135-155.

He, Q., Yan, Z., Si, F., Zhou, Y., Fu, W., & Chen, B. (2019). ATP-Binding Cassette (ABC) Transporter Genes Involved in Pyrethroid Resistance in the Malaria Vector *Anopheles* sinensis: Genome-Wide Identification, Characteristics, Phylogenetics, and Expression Profile. *International Journal of Molecular Sciences*, *20*(6), 1409.

Hemingway, J., Field, L., & Vontas, J. (2002). An Overview of Insecticide Resistance. *Science*, *298*(5591), 96-97.

Hemingway, J., Ranson, H., Magill, A., Kolaczinski, J., Fornadel, C., Gimnig, J., Coetzee, M., Simard, F., Roch, D. K., Hinzoumbe, C. K., Pickett, J., Schellenberg, D., Gething, P., Hoppé, M., & Hamon, N. (2016). Averting a malaria disaster: will insecticide resistance derail malaria control? *The Lancet*, *387*(10029), 1785-1788.

Howe, K. L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., Alvarez-Jarreta, J., Barba, M., Bolser, D. M., Cambell, L., Carbajo, M., Chakiachvili, M., Christensen, M., Cummins, C., Cuzick, A., Davis, P., Fexova, S., Gall, A., George, N., Gil, L., Gupta, P., Hammond-Kosack, K. E., Haskell, E., Hunt, S. E., Jaiswal, P., Janacek, S. H., Kersey, P. J., Langridge, N., Maheswari, U., Maurel, T., McDowall, M. D., Moore, B., Muffato, M., Naamati, G., Naithani, S., Olson, A., Papatheodorou, I., Patricio, M., Paulini, M., Pedro, H., Perry, E., Preece, J., Rosello, M., Russell, M., Sitnik, V., Staines, D. M., Stein, J., Tello-Ruiz, M. K., Trevanion, S. J., Urban, M., Wei, S., Ware, D., Williams, G., Yates, A. D., & Flicek, P. (2019). Ensembl Genomes 2020-enabling non-vertebrate genomic research. *Nucleic Acids Research*, *48*(D1), D689-D695.

Hu, X., Sun, Y., Wang, W., Yang, M., Sun, L., Tan, W., Sun, J., Qian, J., Ma, L., Zhang, D., Yan, G., & Zhu, C. (2007). Cloning and characterization of *NYD-OP7*, a novel deltamethrin resistance associated gene from *Culex* pipiens pallens. *Pesticide Biochemistry and Physiology*, *88*(1), 82-91.

Ibrahim, S. S., Ndula, M., Riveron, J. M., Irving, H., & Wondji, C. S. (2016a). The P450 CYP6Z1 confers carbamate/pyrethroid cross-resistance in a major African malaria vector beside a novel carbamate-insensitive N485I acetylcholinesterase-1 mutation. *Molecular Ecology*, *25*(14), 3436-3452.

Ibrahim, S. S., Riveron, J. M., Stott, R., Irving, H., & Wondji, C. S. (2016b). The cytochrome P450 CYP6P4 is responsible for the high pyrethroid resistance in knockdown resistance-free *Anopheles* arabiensis. *Insect Biochemistry and Molecular Biology*, *68*, 23-32.

Ingham, V. A., Wagstaff, S. & Ranson, H. (2018). Transcriptomic meta-signatures identified in *Anopheles* gambiae populations reveal previously undetected insecticide resistance mechanisms. *Nature Communications*, *9*, 5282.

Ioannidis, P., Simao, F. A., Waterhouse, R. M., Manni, M., Seppey, M., Robertson, H. M., Misof, B., Niehuis, O., & Zdobnov, E. M. (2017). Genomic Features of the Damselfly *Calopteryx splendens* Representing a Sister Clade to Most Insect Orders. *Genome Biology and Evolution*, *9*(2), 415-430.

Isaacs, A. T., Mawejje, H. D., Tomlinson, S., Rigden, D. J., & Donnelly, M. J. (2018). Genome-wide transcriptional analyses in *Anopheles* mosquitoes reveal an unexpected association between salivary gland gene expression and insecticide resistance. *BMC Genomics*, *19*(1), 225.

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S. Y., Lopez, R., & Hunter, S. (2014). InterProScan 5:

genome-scale protein function classification. *Bioinformatics*, *30*(9), 1236-1240.

Kampouraki, A., Stavrakaki, M., Karataraki, A., Katsikogiannis, G., Pitika, E., Varikou, K., Vlachaki, A., Chrysargyris, A., Malandraki, A., Sidiropoulos, N., Paraskevopoulos, A., Gkilpathi, D., Roditakis, E., & Vontas, J. (2018). Recent evolution and operational impact of insecticide resistance in olive fruit fly *Bactrocera oleae* populations from Greece. *Journal of Pest Science*, *91*(4), 1429-1439.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772-780.

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907-915.

Kingan, S. B., Heaton, H., Cudini, J., Lambert, C. C., Baybayan, P., Galvin, B. D., Durbin, R., Korlach, J., & Lawniczak, M. (2019). A High-Quality De novo Genome Assembly from a Single Mosquito Using PacBio Sequencing. *Genes*, *10*(1), 62.

Kwiatkowska, R. M., Platt, N., Poupardin, R., Irving, H., Dabire, R. K., Mitchell, S., Jones, C. M., Diabaté, A., Ranson, H., & Wondji, C. S. (2013). Dissecting the mechanisms responsible for the multiple insecticide resistance phenotype in *Anopheles* gambiae s.s., M form, from Vallée du Kou, Burkina Faso. *Gene*, *519*(1), 98-106.

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930.

Le Goff, G., & Hilliou, F. (2017). Resistance evolution in *Drosophila*: the case of CYP6G1. *Pest Management Science*, *73*(3), 493-499.

Lee, E., Helt, G. A., Reese, J. T., Munoz-Torres, M. C., Childers, C. P., Buels, R. M., Stein, L., Holmes, I. H., Elsik, C. G., & Lewis, S. E. (2013). Web Apollo: a web-based genomic annotation editing platform. *Genome Biology*, *14*(8), R93.

Leung, N. Y., & Montell, C. (2017). Unconventional Roles of Opsins. *Annual Review of Cell and Developmental Biology*, *33*, 241-264.

Li, T., Cao, C., Yang, T., Zhang, L., He, L., Xi, Z., Bian, G., & Liu, N. (2015). A G-protein-coupled receptor regulation pathway in cytochrome P450-mediated permethrin-resistance in mosquitoes, *Culex* quinquefasciatus. *Scientific Reports*, *5*, 17772.

Li, T., Liu, L., Zhang, L., & Liu, N. (2014). Role of G-protein-coupled Receptor-related Genes in Insecticide Resistance of the Mosquito, *Culex quinquefasciatus*. *Scientific Reports*, *4*(1), 6474.

Li, T., & Liu, N. (2019). Role of the G-Protein-Coupled Receptor Signaling Pathway in Insecticide Resistance. *International Journal of Molecular Sciences*, *20*(17), 4300.

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923-930.

Margaritopoulos, J. T., Skavdis, G., Kalogiannis, N., Nikou, D., Morou, E., Skouras, P. J., Tsitsipis, J. A., & Vontas, J. (2008). Efficacy of the pyrethroid alpha-cypermethrin against *Bactrocera oleae*populations from Greece, and improved diagnostic for an iAChE mutation. *Pest Management Science*, *64*(9), 900-908.

Mastrantonio, V., Ferrari, M., Negri, A., Sturmo, T., Favia, G., Porretta, D., Epis, S., & Urbanelli, S. (2019). Insecticide Exposure Triggers a Modulated Expression of ABC Transporter Genes in Larvae of *Anopheles* gambiae s.s. *Insects*, *10*(3), 66.

McCarthy, D. J., Chen, Y., & Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, *40*(10), 4288-4297.

Merrill, C. E., Pitts, R. J., & Zwiebel, L. J. (2003). Molecular characterization of arrestin family members in the malaria vector mosquito, *Anopheles gambiae*. *Insect Molecular Biology*, *12*(6), 641-650.

Mitchell, S. N., Stevenson, B. J., Müller, P., Wilding, C. S., Egyir-Yawson, A., Field, S. G., Hemingway, J., Paine, M. J., Ranson, H., & Donnelly, M. J. (2012). Identification and validation of a gene causing cross-resistance between insecticide classes in *Anopheles* gambiae from Ghana. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(16), 6147-6152.

Morand, S., & Lajaunie, C. (2017). *Biodiversity and Health: Linking Life, Ecosystems and Societies 1st Edition*. Elsevier. ISBN: 9781785481154.

Mugenzi, L., Menze, B. D., Tchouakui, M., Wondji, M. J., Irving, H., Tchoupo, M., Hearn, J., Weedall, G. D., Riveron, J. M., & Wondji, C. S. (2019). Cis-regulatory CYP6P9b P450 variants associated with loss of insecticide-treated bed net efficacy against *Anopheles* funestus. *Nature Communications*, *10*(1), 4652.

Müller, P., Warr, E., Stevenson, B. J., Pignatelli, P. M., Morgan, J. C., Steven, A., Yawson, A. E., Mitchell, S. N., Ranson, H., Hemingway, J., Paine, M. J., & Donnelly, M. J. (2008). Field-caught permethrin-resistant *Anopheles* gambiae overexpress CYP6P3, a P450 that metabolises pyrethroids. *PLoS Genetics*, *4*(11), e1000286.

Mutero, A., Pralavorio, M., Bride, J. M., & Fournier, D. (1994). Resistance-associated point mutations in insecticide-insensitive acetylcholinesterase. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(13), 5922–5926.

Nkya, T. E., Akhouayri, I., Poupardin, R., Batengana, B., Mosha, F., Magesa, S., Kisinza, W., & David, J. P. (2014). Insecticide resistance mechanisms associated with different environments in the malaria vector *Anopheles* gambiae: a case study in Tanzania. *Malaria Journal*, *13*, 28.

Oliver, S. V., & Brooke, B. D. (2016). The Role of Oxidative Stress in the Longevity and Insecticide Resistance Phenotype of the Major Malaria Vectors *Anopheles* arabiensis and *Anopheles* funestus. *PloS ONE*, *11*(3), e0151049.

Papanicolaou, A., Schetelig, M. F., Arensburger, P., Atkinson, P. W., Benoit, J. B., Bourtzis, K., Castañera, P., Cavanaugh, J. P., Chao, H., Childers, C., Curril, I., Dinh, H., Doddapaneni, H., Dolan, A., Dugan, S., Friedrich, M., Gasperi, G., Geib, S., Georgakilas, G., Gibbs, R. A., Giers, S. D., Gomulski, L. M., González-Guzmán, G., Guillem-Amat, A., Han, Y., Hatzigeorgiou, A. G., Hernández-Crespo, P., Hughes, D. S. T., Jones, J. W., Karagkouni, D., Koskinioti, P., Lee, S. L., Malacrida, A. R., Manni, M., Mathiopoulos, K., Meccariello, A., Munoz-Torres, M., Murali, S. C., Murphy, T. D., Muzny, D. M., Oberhofer, G., Ortego, F., Paraskevopoulou, M. D., Poelchau, M., Qu, J., Reczko, M., Robertson, H. M., Rosendale, A. J., Rosselot, A. E., Saccone, G., Salvemini, M., Savini, G., Schreiner, P., Scolari, F., Siciliano, P., Sim, S. B., Tsiamis, G., Ureña, E., Vlachos, I. S., Werren, J. H., Wimmer, E. A., Worley, K. C., Zacharopoulou, A., Richards, S., & Handler, A. M. (2016). The whole genome sequence of the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann), reveals insights into the biology and adaptive evolution of a highly invasive pest species. *Genome Biology*, *17*(1), 192.

Pavlidi, N., Gioti, A., Wybouw, N., Dermauw, W., Ben-Yosef, M., Yuval, B., Jurkevich, E., Kampouraki, A., Van Leeuwen, T., & Vontas, J. (2017). Transcriptomic responses of the olive fruit fly *Bactrocera oleae* and its symbiont Candidatus *Erwinia dacicola* to olive feeding. *Scientific Reports*, *7*(1), 42633.

Pavlidi, N., Kampouraki, A., Tseliou, V., Wybouw, N., Dermauw, W., Roditakis, E., Nauen, R.., Van Leeuwen, T., & Vontas, J. (2018). Molecular characterization of pyrethroid resistance in the olive fruit fly *Bactrocera oleae*. *Pesticide Biochemistry and Physiology*, *148*, 1-7.

Pelosi, P., Iovinella, I., Zhu, J., Wang, G., & Dani, F. R. (2018). Beyond chemoreception: diverse tasks of soluble olfactory proteins in insects. *Biological Reviews*, *93*(1), 184-200.

Perry, T., Batterham, P., & Daborn, P. J. (2011). The biology of insecticidal activity and resistance. *Insect Biochemistry and Molecular Biology*, *41*(7), 411-422.

Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M. V., Unger, M. F., Collins, F. H., & Feyereisen, R. (2002). Evolution of supergene families associated with insecticide resistance. *Science* (New York, N.Y.), *298*(5591), 179-181.

Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H., & Vilo, J. (2016). g:Profiler—a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Research*, *44*(W1), W83–W89.

Riveron, J. M., Irving, H., Ndula, M., Barnes, K. G., Ibrahim, S. S., Paine, M. J. I., & Wondji, C. S. (2013). Directionally selected cytochrome P450 alleles are driving the spread of pyrethroid resistance in the major malaria vector *Anopheles funestus*. *Proceedings of the National Academy of Sciences*, *110*(1), 252-257.

Riveron, J. M., Yunta, C., Ibrahim, S. S., Djouaka, R., Irving, H., Menze, B. D., Ismail, H. M., Hemingway, J., Ranson, H., Albert, A., & Wondji, C. S. (2014). A single mutation in the *GSTe2* gene allows tracking of metabolically based insecticide resistance in a major malaria vector. *Genome Biology*, *15*(2), R27.

Riveron, J. M., Tchouakui, M., Mugenzi, L., Menze, B. D., Chiang, M.-C., & Wondji, C. S. (2018). Insecticide Resistance in Malaria Vectors: An Update at a Global Scale. In: Manguin, S., & Dev, V. (*eds.*). *Towards Malaria Elimination - A Leap Forward.* ISBN: 978-1-78923-551-7.

Robinson, G. E., Hackett, K. J., Purcell-Miramontes, M., Brown, S. J., Evans, J. D., Goldsmith, M. R., Lawson, D., Okamuro, J., Robertson, H. M., & Schneider, D. J. (2011). Creating a Buzz About Insect Genomes. *Science*, *331*(6023), 1386-1386.

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, 11, R25.

Rognes, T. (2011). Faster Smith-Waterman database searches with inter-sequence SIMD parallelisation. *BMC Bioinformatics*, *12*, 221.

Sakka, M., Riga, M., Ioannidis, P., Mpaliota, G., Jagadeesan, R., Nayak, M. K., Vontas, J., & Athanassiou, C. (2020). Transcriptomic analysis of s-methoprene resistance in the lesser grain borer, *Rhyzopertha dominica*, and evaluation of

piperonyl butoxide as a resistance breaker. Research Square, DOI: 10.21203/rs.3.rs-18153/v1.

Samb, B., Konate, L., Irving, H., Riveron, J. M., Dia, I., Faye, O., & Wondji, C. S. (2016). Investigating molecular basis of lambda-cyhalothrin resistance in an *Anopheles funestus* population from Senegal. *Parasites & Vectors*, *9*(1), 449.

Seong, K. M., Coates, B. S., Kim, D., Hansen, A. K., & Pittendrigh, B. R. (2018). Differentially expressed microRNAs associated with changes of transcript levels in detoxification pathways and DDT-resistance in the *Drosophila melanogaster* strain 91-R. *PLoS ONE*, *13*(4), e0196518.

Simma, E. A., Dermauw, W., Balabanidou, V., Snoeck, S., Bryon, A., Clark, R. M., Yewhalaw, D., Vontas, J., Duchateau, L., & Van Leeuwen, T. (2019). Genome-wide gene expression profiling reveals that cuticle alterations and P450 detoxification are associated with deltamethrin and DDT resistance in *Anopheles arabiensis* populations from Ethiopia. *Pest Management Science*, *75*(7), 1808-1818.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312-1313.

Stevenson, B. J., Bibby, J., Pignatelli, P., Muangnoicharoen, S., O'Neill, P. M., Lian, L. Y., Müller, P., Nikou, D., Steven, A., Hemingway, J., Sutcliffe, M. J., & Paine, M. J. (2011). Cytochrome P450 6M2 from the malaria vector *Anopheles* gambiae metabolizes pyrethroids: Sequential metabolism of deltamethrin revealed. *Insect Biochemistry and Molecular Biology*, *41*(7), 492-502.

Stica, C., Jeffries, C. L., Irish, S. R., Barry, Y., Camara, D., Yansane, I., Kristan, M., Walker, T., & Messenger, L. A. (2019). Characterizing the molecular and metabolic mechanisms of insecticide resistance in *Anopheles gambiae* in Faranah, Guinea. *Malaria Journal*, *18*(1), 244.

Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B., Wu, C. H., & UniProt Consortium (2015). UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics*, *31*(6), 926-932.

The *Anopheles* gambiae 1000 Genomes Consortium (2017). Genetic diversity of the African malaria vector *Anopheles* gambiae. *Nature*, *552*, 96-100.

Toé, K. H., N'Falé, S., Dabiré, R. K., Ranson, H., & Jones, C. M. (2015). The recent escalation in strength of pyrethroid resistance in *Anopheles* coluzzi in West Africa is linked to increased expression of multiple gene families. *BMC Genomics*, *16*(1), 146.

Traylor, M. J., Baek, J.-M., Richards, K. E., Fusetto, R., Huang, W., Josh, P., Chen, Z.,

Bollapragada, P., O'Hair, R. A. J., Batterham, P., & Gillam, E. M. J. (2017). Recombinant expression and characterization of Lucilia cuprina CYP6G3: Activity and binding properties toward multiple pesticides. *Insect Biochemistry and Molecular Biology*, 90, 14-22.

Tsakireli, D., Riga, M., Kounadi, S., Douris, V., & Vontas, J. (2019). Functional characterization of CYP6A51, a cytochrome P450 associated with pyrethroid resistance in the Mediterranean fruit fly *Ceratitis capitata*. *Pesticide Biochemistry and Physiology*, *157*, 196-203.

Tse, E. G., Korsik, M. & Todd, M. H. (2019). The past, present and future of anti-malarial medicines. *Malaria Journal*, *18*, 93.

Vannini, L., Reed, T. W. & Willis, J. H. (2014). Temporal and spatial expression of cuticular proteins of *Anopheles* gambiae implicated in insecticide resistance or differentiation of M/S incipient species. *Parasites & Vectors*, *7*, 24.

Vontas, J., Blass, C., Koutsos, A. C., David, J. P., Kafatos, F. C., Louis, C., Hemingway, J., Christophides, G. K., & Ranson, H. (2005). Gene expression in insecticide resistant and susceptible *Anopheles* gambiae strains constitutively or after insecticide exposure. *Insect Molecular Biology*, *14*(5), 509-521.

Vontas, J., Grigoraki, L., Morgan, J., Tsakireli, D., Fuseini, G., Segura, L., Niemczura de Carvalho, J., Nguema, R., Weetman, D., Slotman, M. A., & Hemingway, J. (2018). Rapid selection of a pyrethroid metabolic enzyme CYP9K1 by operational malaria control activities. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(18), 4619-4624.

Vontas, J., Hernández-Crespo, P., Margaritopoulos, J. T., Ortego, F., Feng, H.-T., Mathiopoulos, K. D., & Hsu, J.-C. (2011). Insecticide resistance in Tephritid flies. *Pesticide Biochemistry and Physiology*, *100*(3), 199-205.

Weedall, G. D., Mugenzi, L., Menze, B. D., Tchouakui, M., Ibrahim, S. S., Amvongo-Adjia, N., Irving, H., Wondji, M. J., Tchoupo, M., Djouaka, R., Riveron, J. M., & Wondji, C. S. (2019). A cytochrome P450 allele confers pyrethroid resistance on a major African malaria vector, reducing insecticide-treated bednet efficacy. *Science Translational Medicine*, *11*(484), eaat7386.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Willis, J. H., Papandreou, N. C., Iconomidou, V. A., Hamodrakas, S. J. (2012). Cuticular Proteins. In: Gilbert, L. I. (*ed.*) *Insect Molecular Biology and Biochemistry*, pp. 134-167. Amsterdam: Elsevier/Academic Press.

Wu, H.-W., Tian, H.-S., Wu, G.-L., Langdon, G., Kurtis, J., Shen, B., Ma, L., Li, X. L., Gu, Y., Hu, X. B., & Zhu, C. L. (2004). *Culex pipiens pallens*: identification of genes differentially expressed in deltamethrin-resistant and -susceptible strains. *Pesticide Biochemistry and Physiology*, *79(*3), 75-83.

Yahouédo, G. A., Chandre, F., Rossignol, M., Ginibre, C., Balabanidou, V., Mendez, N., Pigeon, O., Vontas, J., & Cornelie, S. (2017). Contributions of cuticle permeability and enzyme detoxification to pyrethroid resistance in the major malaria vector *Anopheles* gambiae. *Scientific Report*s, *7*(1), 11091.

Zhou, D., Duan, B., Xu, Y., Ma, L., Shen, B., Sun, Y., & Zhu, C. (2018). NYD-OP7/PLC regulatory signaling pathway regulates deltamethrin resistance in *Culex* pipiens pallens (Diptera: Culicidae). *Parasites & Vectors*, *11*(1), 419.

Zhu, F., Moural, T. W., Nelson, D. R., & Palli, S. R. (2016). A specialist herbivore pest adaptation to xenobiotics through up-regulation of multiple Cytochrome P450s. *Scientific Reports*, *6*, 20421.

Zhu, F., Parthasarathy, R., Bai, H., Woithe, K., Kaussmann, M., Nauen, R., Harrison, D. A., Palli, S. R. (2010). A brain specific cytochrome P450 responsible for the majority of deltamethrin resistance in the QTC279 strain of *Tribolium castaneum*. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 8557-8562.

Wang, H., Shi, Y., Wang, L., Liu, S., Wu, S., Yang, Y., Feyereisen, R., & Wu, Y. (2018). CYP6AE gene cluster knockout in Helicoverpa armigera reveals role in detoxification of phytochemicals and insecticides. *Nature Communications*, *9*(1), 4820.

WHO (2019). *World Malaria Report 2019*. ISBN: 978-92-4-156572-1.

Wilson, A. L., Courtenay, O., Kelly-Hope, L. A., Scott, T. W., Takken, W., Torr, S. J., & Lindsay, S. W. (2020). The importance of vector control for the control and elimination of vector-borne diseases. *PLoS Neglected Tropical Diseases*, *14*(1), e0007831.

Wykes M. N. (2013). Why haven't we made an efficacious vaccine for malaria?. *EMBO Reports*, *14*(8), 661.

Appendix

Table S1. Accession numbers and putative functions for the top 20 most up-regulated genes and the top 20 most down-regulated genes in VK7-HR vs VK7-LR and their fold change in expression in the VK7-HR vs Ngousso comparison.

		VK7-HR vs VK7-LR		VK7-HR vs Ngousso	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**	**Fold Change**	**Corrected p-value**
Top 20 most up-regulated genes in VK7-HR vs VK7-LR					
ACON012199	Histone H2B	232.01	4.62E-35	76.52	7.42E-31
ACON029698	EGF-like domain containing protein	202.43	3.27E-15	59.34	6.68E-11
ACON006913	CUB domain containing protein	65.35	9.51E-17	56.47	1.11E-12
ACON010132	SCRBQ1-Class B Scavenger Receptor (CD36 domain)	58.19	1.84E-175	64.31	2.90E-145
ACON007951	ionotropic receptor IR68a	57.96	7.38E-74	27.13	6.41E-47
ACON002210	cytochrome P450 CYP325B1	46.72	7.98E-12	42.64	2.77E-10
ACON013134	uncharacterized protein	46.37	1.05E-35	28.37	1.68E-17
ACON009373	uncharacterized protein	39.23	9.47E-26		
ACON029319	uncharacterized protein	28.38	3.40E-122	15.54	1.68E-16
ACON006468	I-BAR domain containing protein MTSS1/MTSS2	22.85	2.22E-10	11.88	9.05E-08
ACON013465	cuticular protein CPLCP1	22.75	5.70E-07	65.71	4.80E-34
ACON005457	Ser/Thr protein phosphatase/nucleotidase	21.52	3.65E-15	33.49	4.42E-18
ACON013123	glucose dehydrogenase (acceptor)	20.19	2.61E-15	20.89	4.74E-13
ACON010831	thioester-containing protein 8 (TEP8)	19.78	4.83E-18	35.81	2.88E-17
ACON001203	chitin binding domain containing protein	16.51	1.46E-11		
ACON013180	spindle B	14.65	7.36E-15		
ACON006343	peptidoglycan recognition protein	14.80	8.86E-07		

	PGRPS2				
ACON013316	uncharacterized protein	14.30	2.58E-14	50.16	1.19E-16
ACON000818	cytochrome P450 CYP9K1	13.71	3.41E-63	20.11	4.94E-15
ACON009984	cell cycle control protein 50A	13.68	2.75E-11		
Top 20 most down-regulated genes in VK7-HR vs VK7-LR					
ACON028181	uncharacterized protein	-1139.99	9.04E-06	-530.51	9.13E-26
ACON013064	uncharacterized protein	-1009.79	2.02E-56	-165.43	2.10E-47
ACON029433	uncharacterized protein	-905.78	1.32E-40	-624.74	5.42E-31
ACON000612	salivary gland protein 1 (SG1)	-874.35	1.51E-05	-392.95	1.26E-06
ACON006079	odorant-binding protein 53 (Obp53)	-847.63	4.86E-90	-686.23	5.04E-141
ACON006433	chitin binding domain containing protein	-828.18	8.23E-04		
ACON011981	uncharacterized protein	-726.13	6.36E-45	-314.18	2.70E-14
ACON000610	salivary gland protein 1-like 6 (SG1f)	-595.58	7.20E-08	-437.54	2.52E-16
ACON000607	salivary gland protein 1-like 3 (SG1c)	-561.31	4.59E-08	-514.76	1.08E-15
ACON006504	salivary gland protein 2-like (SG2b)	-527.65	1.78E-05	-285.51	3.05E-09
ACON010813	thioester-containing protein 18 (TEP18)	-504.10	1.97E-17	-356.04	1.08E-12
ACON006400	alkaline phosphatase 2	-383.49	2.17E-88	-410.08	4.71E-106
ACON006507	salivary gland protein 3 (SG3)	-301.69	9.04E-06	-300.25	2.72E-08
ACON012124	zinc finger C2H2 protein	-281.69	1.59E-23	-193.21	7.73E-14
ACON029344	uncharacterized protein	-243.98	2.32E-11	-317.86	3.03E-06
ACON006506	salivary gland protein 2 (SG2)	-240.33	5.81E-05	-174.51	3.19E-07
ACON008956	uncharacterized protein	-233.41	5.10E-07	-157.36	6.48E-24
ACON013423	salivary gland protein 9 (SG9)	-229.09	2.09E-04	-152.92	1.21E-06
ACON007015	amine oxidase	-189.01	2.60E-155	-119.06	1.48E-37
ACON008216	salivary gland protein 7 (SG7)	-186.53	4.46E-08	-166.41	1.26E-15

Table S2. Overview of the 163 commonly down-regulated genes in the two comparisons comprising constitutive resistance, hence VK7-HR vs VK7-LR and VK7-HR vs Ngousso.

		VK7-HR vs VK7-LR		VK7-HR vs Ngousso	
Gene ID	**Description**	**Fold Change**	**Corrected p-value**	**Fold Change**	**Corrected p-value**
ACON028181	uncharacterized protein	-1139.99	9.04E-04	-530.51	9.13E-26
ACON013064	uncharacterized protein	-1009.80	2.02E-56	-1165.44	2.10E-47
ACON029433	uncharacterized protein	-905.78	1.32E-40	-624.75	5.42E-31
ACON000612	SG1 - salivary gland protein 1	-874.36	1.51E-05	-392.96	1.26E-06
ACON006079	Obp53 - odorant-binding protein 53	-847.64	4.86E-90	-686.24	5.04E-141
ACON011981	uncharacterized protein	-726.14	6.36E-45	-314.18	2.70E-14
ACON000610	SG1f - salivary gland protein 1-like 6	-595.59	7.20E-08	-437.55	2.52E-16
ACON000607	SG1c - salivary gland protein 1-like 3	-561.31	4.59E-08	-514.77	1.08E-15
ACON006504	SG2b - salivary gland protein 2-like	-527.66	1.78E-05	-285.52	3.05E-09
ACON010813	TEP18 - thioester-containing protein 18	-504.10	1.97E-17	-356.04	1.08E-12
ACON006400	alkaline phosphatase 2	-383.50	2.17E-88	-410.09	4.71E-106
ACON006507	SG3 - salivary gland protein 3	-301.70	9.04E-06	-300.26	2.72E-08
ACON012124	uncharacterized protein	-281.69	1.59E-23	-193.22	7.73E-14
ACON029344	uncharacterized protein	-243.98	2.32E-11	-317.87	3.03E-06
ACON006506	SG2 - salivary gland protein 2	-240.34	5.81E-05	-174.51	3.19E-07
ACON008956	uncharacterized protein	-233.41	5.10E-07	-157.37	6.48E-24
ACON013423	SG9 - salivary gland protein 9	-229.09	2.09E-04	-152.92	1.21E-06

ACON007015	uncharacterized protein	-189.01	2.60E-155	-51.12	1.48E-37
ACON008216	SG7 - salivary gland protein 7	-186.54	4.46E-08	-166.41	1.26E-15
ACON000609	SG1e - salivary gland protein 1-like 5	-178.92	2.86E-06	-145.85	1.13E-13
ACON000548	SG1b - salivary gland protein 1-like 2	-163.28	1.27E-06	-125.42	1.04E-10
ACON005822	uncharacterized protein	-157.25	1.47E-04	-63.17	5.67E-06
ACON001717	uncharacterized protein	-109.20	3.40E-31	-34.67	2.48E-07
ACON008307	uncharacterized protein	-100.66	3.46E-04	-51.14	2.25E-05
ACON000473	proton-coupled amino acid transporter	-95.54	1.58E-10	-99.38	2.19E-45
ACON008306	uncharacterized protein	-95.48	9.03E-05	-50.52	2.34E-05
ACON001374	trio - TRIO salivary gland protein	-92.99	1.65E-05	-79.34	1.02E-13
ACON000611	SG1d - salivary gland protein 1-like 4	-84.36	4.21E-05	-70.62	2.72E-09
ACON006371	amyrel	-70.67	1.45E-04	-36.29	1.42E-04
ACON001781	uncharacterized protein	-64.32	2.45E-10	-137.51	2.23E-23
ACON001352	uncharacterized protein	-61.71	1.83E-05	-15.57	1.31E-06
ACON000946	C-4 methylsterol oxidase	-60.45	1.35E-15	-81.58	2.81E-27
ACON010601	ILP4 - insulin-like peptide 4	-60.23	1.02E-05	-30.99	1.06E-04
ACON013246	uncharacterized protein	-51.37	9.24E-10	-140.25	6.60E-24
ACON029324	uncharacterized protein	-49.52	4.16E-18	-27.96	2.91E-13
ACON002994	CPR10 - cuticular protein RR-2 family 10	-43.75	8.75E-06	-14.91	3.48E-04
ACON009918	SGS5	-42.07	3.17E-08	-18.71	4.88E-05
ACON008282	D7r2 - D7 short form	-42.04	8.10E-05	-30.35	5.37E-04

	salivary protein				
ACON006775	glucosyl/glucuronosyl transferases	-41.80	1.49E-32	-17.38	6.23E-07
ACON010604	ILP6 - Insulin-like peptide 6	-38.81	1.44E-04	-26.67	7.26E-05
ACON002629	uncharacterized protein	-35.44	9.59E-15	-1986.97	2.06E-13
ACON002292	PKP - pyrokinin	-34.66	1.59E-06	-20.68	6.38E-04
ACON011098	uncharacterized protein	-32.94	6.09E-05	-36.62	8.25E-07
ACON006194	uncharacterized protein	-31.70	1.94E-04	-35.35	1.81E-16
ACON008281	D7r4 - D7 short form salivary protein	-29.22	7.77E-06	-11.22	4.48E-04
ACON006195	uncharacterized protein	-28.61	2.95E-04	-20.96	2.31E-29
ACON000605	uncharacterized protein	-28.23	3.72E-10	-679.80	1.60E-07
ACON028406	uncharacterized protein	-27.15	1.30E-11	-7.12	4.56E-43
ACON003433	uncharacterized protein	-25.76	7.82E-08	-41.24	4.32E-14
ACON010409	OBP22 - odorant-binding protein 22	-25.34	4.35E-05	-17.81	5.17E-05
ACON028579	uncharacterized protein	-25.11	1.49E-95	-16.61	1.05E-39
ACON002769	C-4 methylsterol oxidase	-23.83	1.51E-19	-33.90	1.83E-13
ACON011785	CLIPE6 - CLIP-domain serine protease	-23.51	4.75E-06	-12.19	3.03E-04
ACON012543	uncharacterized protein	-23.20	2.20E-09	-11.32	6.19E-05
ACON013470	uncharacterized protein	-20.84	5.13E-11	-10.61	3.94E-08
ACON004883	uncharacterized protein	-20.82	4.69E-06	-17.45	9.72E-79
ACON011467	Or5 - odorant receptor 5	-20.51	7.03E-16	-7.04	5.11E-08
ACON009502	uncharacterized protein	-19.66	3.76E-07	-18.89	1.60E-05

ACON028082	uncharacterized protein	-18.80	6.29E-05	-13.34	5.35E-07
ACON008968	uncharacterized protein	-18.76	1.78E-27	-5.50	8.94E-16
ACON002556	Obp66 - odorant-binding protein 66	-18.27	9.86E-04	-11.80	2.11E-04
ACON012308	ornithine decarboxylase	-17.90	2.89E-05	-8.90	3.23E-04
ACON006709	CHYM1 - Chymotrypsin-1	-17.64	1.17E-35	-17.85	1.06E-24
ACON006026	uncharacterized protein	-17.56	9.49E-12	-10.99	2.35E-05
ACON029415	uncharacterized protein	-17.31	1.72E-17	-10.86	8.58E-09
ACON013239	juvenile hormone-inducible protein	-17.25	1.73E-14	-7.09	1.03E-05
ACON028150	uncharacterized protein	-16.40	6.26E-05	-29.11	1.87E-07
ACON005518	FMRF - FMRFamide	-16.04	4.09E-09	-10.10	5.54E-04
ACON005166	uncharacterized protein	-15.90	1.02E-90	-15.12	9.84E-21
ACON009708	uncharacterized protein	-15.74	1.19E-04	-11.90	1.72E-04
ACON003938	uncharacterized protein	-15.28	1.34E-16	-8.09	3.43E-17
ACON006711	chym3 - Chymotrypsin-2	-15.19	9.60E-30	-27.54	7.13E-10
ACON012979	glucose dehydrogenase (acceptor)	-15.66	4.13E-16	-15.12	4.03E-11
ACON029196	Fatty acyl-CoA reductase	-15.02	1.46E-05	-24.02	4.59E-34
ACON012201	uncharacterized protein	-14.78	1.14E-09	-14.07	8.58E-07
ACON011443	carboxypeptidase	-14.32	4.26E-05	-11.59	2.70E-12
ACON002451	SCRB1 - Class B Scavenger Receptor (CD36 domain).	-13.62	4.76E-04	-15.95	2.00E-04
ACON006710	CHYM2 - chymotrypsin	-11.93	2.70E-14	-34.79	2.57E-05
ACON013016	glucose dehydrogenase	-11.52	8.51E-10	-13.48	1.99E-09

	(acceptor)				
ACON004694	uncharacterized protein	-11.33	2.49E-07	-10.36	5.07E-11
ACON008065	uncharacterized protein	-11.31	5.11E-05	-23.78	3.42E-05
ACON012216	uncharacterized protein	-11.17	4.98E-36	-23.25	6.49E-39
ACON007752	uncharacterized protein	-11.13	6.57E-17	-17.49	5.77E-13
ACON012217	sterol O-acyltransferase	-11.10	2.51E-20	-38.73	8.94E-48
ACON029062	OBP1 - odorant-binding protein 1	-11.79	2.40E-05	-13.10	2.32E-05
ACON002328	uncharacterized protein	-10.71	1.41E-15	-8.34	2.24E-12
ACON008467	solute carrier family 6 (neurotransmitter transporter%2C amino acid) member 5/7/	-10.63	1.29E-07	-6.71	1.40E-04
ACON013762	uncharacterized protein	-10.57	7.09E-06	-12.64	7.42E-14
ACON012497	uncharacterized protein	-10.54	1.57E-25	-5.92	8.86E-15
ACON003547	MORN repeat-containing protein 4	-10.42	3.46E-05	-12.92	6.26E-05
ACON009635	uncharacterized protein	-10.25	3.19E-10	-4.77	3.23E-18
ACON003939	uncharacterized protein	-9.97	1.95E-42	-6.46	3.78E-16
ACON007359	uncharacterized protein	-9.80	1.40E-07	-10.27	2.76E-13
ACON008688	uncharacterized protein	-9.12	2.14E-07	-8.92	1.22E-20
ACON001240	uncharacterized protein	-9.79	6.50E-04	-51.38	1.60E-04
ACON003775	uncharacterized protein	-8.89	8.16E-20	-5.13	6.15E-11
ACON003778	uncharacterized protein	-8.70	3.12E-30	-5.19	2.24E-11
ACON005238	facilitated glucose transporter (solute	-8.57	2.04E-05	-12.75	1.12E-08

	carrier family 2)				
ACON002429	CYP314A1 - cytochrome P450	-8.15	3.81E-21	-4.62	2.03E-10
ACON005810	uncharacterized protein	-8.93	1.11E-20	-8.89	4.18E-15
ACON006081	OBP55 - odorant-binding protein 55	-8.08	7.69E-10	-4.94	1.41E-11
ACON003677	uncharacterized protein	-8.59	2.55E-16	-13.15	1.07E-27
ACON002771	uncharacterized protein	-8.29	7.58E-04	-46.75	1.25E-04
ACON001382	DH31 - diuretic hormone 31	-7.97	1.15E-04	-5.68	2.88E-04
ACON009756	ANCE6 - angiotensin-converting enzyme 6	-7.93	2.95E-07	-8.33	6.68E-04
ACON007679	uncharacterized protein	-7.60	8.08E-06	-7.38	7.83E-06
ACON010294	uncharacterized protein	-7.59	2.69E-08	-17.81	6.80E-15
ACON001890	uncharacterized protein	-7.46	2.65E-04	-10.42	1.16E-04
ACON005834	COEJHE2E - carboxylesterase	-7.45	6.64E-10	-4.08	5.75E-34
ACON010885	(S)-2-hydroxy-acid oxidase	-7.30	3.74E-39	-7.23	2.53E-49
ACON003274	solute carrier family 12 (sodium/potassium/chloride transporter) 2C member 2	-7.30	1.70E-07	-9.82	5.67E-04
ACON005099	uncharacterized protein	-7.18	3.23E-06	-8.79	3.28E-06
ACON003515	uncharacterized protein	-7.12	2.19E-21	-5.68	4.57E-05
ACON001591	cadherin-87A	-7.90	6.91E-10	-7.67	9.57E-13
ACON000140	IR100a - ionotropic receptor IR100a	-6.98	7.65E-06	-7.66	1.34E-05
ACON012182	uncharacterized protein	-6.95	6.54E-14	-6.10	1.04E-05
ACON029259	uncharacterized protein	-6.94	9.22E-16	-20.04	5.51E-26

ACON000347	CAPA - pyrokinin capa-like	-6.78	2.00E-04	-5.45	1.58E-07
ACON003608	CYP4AA1 - cytochrome P450	-6.47	2.28E-04	-11.41	1.28E-11
ACON004118	SCRAL1 - Class A Scavenger Receptor (SRCR domain) with Lysyl Oxidase domain.	-6.40	1.27E-13	-5.66	5.81E-04
ACON028554	uncharacterized protein	-6.32	3.12E-05	-8.48	6.94E-04
ACON002168	uncharacterized protein	-6.26	9.83E-05	-13.95	3.43E-13
ACON003777	uncharacterized protein	-6.16	3.68E-21	-4.69	1.19E-10
ACON005774	CYP49A1 - cytochrome P450	-6.55	1.04E-09	-7.39	3.74E-28
ACON007062	ETH - ecdysis-triggering hormone	-6.47	4.66E-12	-4.27	2.04E-05
ACON007048	uncharacterized protein	-6.04	2.50E-04	-4.34	4.49E-04
ACON010077	CYP303A1 - cytochrome P450	-6.02	3.61E-06	-7.40	2.46E-06
ACON006187	Protein G12	-6.03	6.41E-08	-4.56	8.76E-07
ACON006078	Obp52 - odorant-binding protein 52	-5.99	6.10E-04	-6.77	1.33E-14
ACON012653	uncharacterized protein	-5.90	3.73E-14	-4.01	1.82E-07
ACON028167	uncharacterized protein	-5.74	1.18E-15	-4.21	2.51E-06
ACON006048	CYP4J5 - cytochrome P450	-5.53	4.83E-66	-8.95	2.80E-33
ACON006442	uncharacterized protein	-5.44	6.47E-05	-7.34	3.28E-05
ACON007498	IR75k - ionotropic receptor IR75k	-5.41	4.45E-40	-4.91	6.99E-13
ACON028015	uncharacterized protein	-5.39	1.53E-15	-4.07	6.59E-04
ACON012321	OBP26 - odorant-binding protein 26	-5.34	5.63E-33	-9.29	2.77E-16
ACON003776	uncharacterized protein	-5.30	2.31E-13	-4.38	1.30E-08

ACON006049	CYP4J10 - cytochrome P450	-5.25	5.31E-50	-8.40	3.57E-31
ACON003091	Sterol carrier protein 2 variant 2	-5.20	3.72E-05	-5.09	3.21E-07
ACON012401	uncharacterized protein	-5.94	3.87E-05	-4.46	2.73E-33
ACON029292	uncharacterized protein	-5.60	2.22E-10	-4.60	5.35E-39
ACON000296	uncharacterized protein	-5.28	4.71E-32	-4.87	7.90E-43
ACON006024	uncharacterized protein	-4.73	4.71E-07	-5.42	9.11E-07
ACON004760	uncharacterized protein	-4.72	2.83E-04	-6.64	1.59E-07
ACON004433	Obp19 - odorant-binding protein 19	-4.67	3.77E-04	-4.96	1.47E-04
ACON012577	uncharacterized protein	-4.61	1.10E-20	-9.77	4.36E-34
ACON029795	uncharacterized protein	-4.55	1.00E-07	-13.56	8.60E-18
ACON005093	uncharacterized protein	-4.49	4.55E-04	-7.84	4.42E-09
ACON001889	uncharacterized protein	-4.45	1.48E-04	-5.01	8.88E-06
ACON001047	inositol polyphosphate 1-phosphatase	-4.44	2.57E-37	-5.38	3.33E-42
ACON007365	uncharacterized protein	-4.44	2.27E-09	-150.92	2.64E-53
ACON012400	alpha-glucosidase	-4.41	8.70E-12	-6.35	9.34E-19
ACON006437	uncharacterized protein	-4.30	1.11E-43	-4.58	5.18E-34
ACON007270	uncharacterized protein	-4.30	3.39E-04	-4.81	2.55E-05
ACON006932	citron Rho-interacting kinase	-4.29	6.27E-17	-4.77	2.69E-23
ACON000785	Synaptic vesicle protein	-4.18	5.21E-13	-5.33	1.88E-18
ACON006283	CPR70 - cuticular protein RR-2 family 70	-4.17	1.38E-08	-4.18	3.18E-05

ACON006430	CTLGA2 - C-type lectin (CTL) - galactose binding	-4.08	1.11E-10	-4.42	2.14E-07
ACON006893	uncharacterized protein	-4.60	4.24E-06	-5.59	3.54E-05
ACON006638	Oatp58Dc	-4.52	5.91E-05	-6.20	2.53E-05
ACON005610	uncharacterized protein	-4.05	1.65E-05	-5.99	1.22E-09
ACON007456	LRIM8B - leucine-rich immune protein (Short)	-4.27	1.09E-39	-4.53	7.96E-28
ACON010291	uncharacterized protein	-4.03	1.02E-08	-4.78	3.78E-05

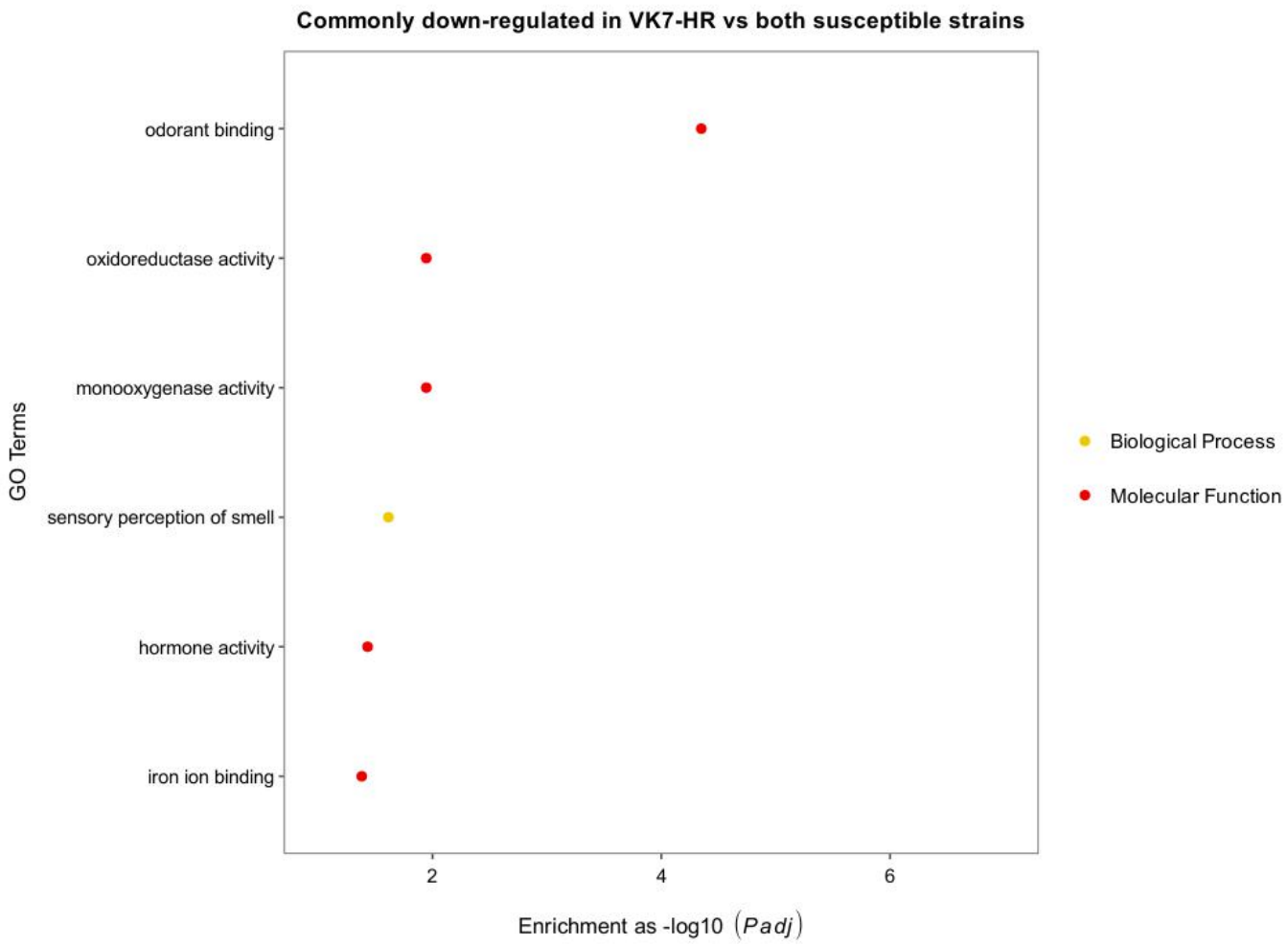

Figure S1. Functional Enrichment Analysis of the down-regulated genes in VK7-HR legs compared to both susceptible strains, VK7-LR and Ngousso.