

A Scaffold-Level Genome Assembly of the Pirate Bug, *Orius Laevigatus*, and a Comparative Analysis of Insecticide Resistance-Related Gene Families with Hemipteran Crop Pests

Emma Bailey (✉ emma.bailey@rothamsted.ac.uk)

Rothamsted Research

Linda Field

Rothamsted Research

Christopher Rawlings

Rothamsted Research

Rob King

Rothamsted Research

Fady Mohareb

Cranfield University

Keywan-Hassani Pak

Rothamsted Research

David Hughes

Rothamsted Research

Martin Williamson

Rothamsted Research

Christoph Zimmer

Syngenta Crop Protection

Eric Ganko

Syngenta Crop Protection

Benjamin Buer

Bayer AG, R&D

Ralf Nauen

Bayer AG, R&D

Research Article

Keywords: *Orius laevigatus*, pirate bug, PacBio, Illumina, whole genome sequencing, beneficial predator, insecticide resistance, comparative genomics, Hemiptera, crop pests

Posted Date: May 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-537204/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: The pirate bug, *Orius laevigatus*, is a highly effective beneficial predator of crop pests including aphids, spider mites and thrips in integrated pest management (IPM) programmes. No genomic information is currently available for *O. laevigatus*, as is the case for the majority of beneficial predators which feed on crop pests. In contrast, genomic information for crop pests is far more readily available. The lack of publicly available genomes for beneficial predators to date has limited our ability to perform comparative analyses of genes encoding potential insecticide resistance mechanisms between crop pests and their predators. These mechanisms include several gene/protein families including cytochrome P450s (P450s), ATP binding cassette transporters (ABCs), glutathione S-transferases (GSTs), UDP-glucosyltransferases (UGTs) and carboxyl/cholinesterases (CCEs).

Methods and findings: In this study, a high-quality scaffold level *de novo* genome assembly for *O. laevigatus* has been generated using a hybrid approach with PacBio long-read and Illumina short-read data. The final assembly achieved a scaffold N50 of 125,649bp and a total genome size of 150.98Mb. The genome assembly achieved a level of completeness of 93.6% using a set of 1,658 core insect genes present as full-length genes. The assembly was annotated with 15,102 protein-coding genes - 87% of which were assigned a putative function - and the resultant gene set achieved a completeness of 84.5%.

Comparative analyses revealed gene expansions of sigma class GSTs and CYP3 P450s. Conversely the UGT gene family showed limited expansion. Differences were seen in the distributions of resistance-associated gene families at the subfamily level between *O. laevigatus* and some of its targeted crop pests. A target site mutation in ryanodine receptors (I4790M, PxRyR) which has strong links to diamide resistance in crop pests and had previously only been identified in lepidopteran species was found to also be present in hemipteran species, including *O. laevigatus*.

Conclusion and significance: This assembly is the first published genome for the Anthocoridae family and will serve as a useful resource for further research into target-site selectivity issues and potential resistance mechanisms in beneficial predators. Furthermore, the expansion of gene families often linked to insecticide resistance may be an indicator of the capacity of this predator to detoxify selective insecticides. These findings could be exploited by targeted pesticide screens and functional studies to increase effectiveness of IPM strategies, which aim to increase crop yields by effectively controlling pests in a sustainable and environmentally-friendly manner, without impacting beneficial predator populations.

Background

Loss of crops to insect pests can account for ~ 10% of potential yield, as a result of both direct feeding damage and the transfer of viral plant diseases [1]. Thus, to maximise crop yields and sustain food production for a growing world population, pests need to be controlled. At present this control relies mainly on the use of synthetic pesticides, many of which are non-selective and are therefore toxic to both their target pest species and to beneficial predators and parasitoids. As a result there may be a reduction in the predator populations to a level where they are no longer able to contribute natural pest control. This, along with the development of insecticide resistance in pests, can lead to pest populations surging, sometimes to even higher levels than pre-pesticide application [2–4]. Beneficial predators, such as those in the genus *Orius*, have proven to be especially effective in the biological control of crop pests [5]. As generalist predators, *Orius* species target a wide variety of pest species including aphids, beet armyworm, leafhoppers, mites, thrips and whiteflies, many of which are the world's most damaging crop pests [6,7]. Some *Orius* species are commercially available as biological control agents and are widely used for this purpose as part of integrated pest management (IPM) strategies, especially in covered crops [8–10].

Whole genome sequences of insects are helping us to understand many aspects of their biology and behaviour, and this can be applied to potential insecticide resistance mechanisms in pest insects and their natural enemies. However, only a few genomes of beneficial predator species have been published to date, including a phytoseiid mite, *Galendromus occidentalis* [11]; three parasitoid wasps, *Nasonia giraulti*, *Nasonia longicornis* and *Nasonia vitripennis* [12] and two lady beetles, *Harmonia axyridis* and *Coccinella septempunctata* [13]. At the time of writing there are no published genomes for species of the *Anthocoridae* family of predators. In contrast, a growing number of genomes of crop pests are available [14–22]. This may be because the genomes of the pests themselves have appeared more useful in terms of developing targeted pesticides and investigating mechanisms of pesticide resistance. However, agriculture is now moving increasingly away from pesticide use – particularly with the Directive on Sustainable Use of Pesticides 2009/128/EC [23] - and towards IPM strategies, which includes the use of beneficial predators, and so the focus of research should also follow this trend [24–28].

The aim of the work reported here was to develop a high-quality genome assembly for *O. laevigatus*, to serve as a resource for research into this species as well as the wider *Anthocoridae* family, which consists of 400–600 mostly predaceous species - a potentially valuable source of biological control agents [29]. The *O. laevigatus* genome was then used for comparative analyses between beneficial predators and crop pests, focusing on genes encoding potential insecticide resistance mechanisms.

There are two main types of insecticide resistance mechanisms: increased expression of genes encoding protein families involved in metabolic resistance and point mutations in genes encoding insecticide target proteins [30]. Gene families involved in insecticide resistance in pest species are known to include cytochrome P450 monooxygenases (P450s), ATP binding cassette transporters (ABCs), glutathione S-transferases (GSTs), UDP-glucosyltransferases (UGTs) and carboxyl/cholinesterases (CCEs) [31–36]. Comparisons of the genes/proteins which may be involved in insecticide resistance in crop pests with the corresponding genes in beneficial insects could aid the development of insecticides which target crop pests but have limited impact on beneficial predator populations. This could prove key to developing successful IPM strategies which exploit differences in the ability of predators and crop pests to tolerate pesticides. Improving the availability of beneficial predator genomes could also help the selection of beneficial predators with genes/mutations for inherent insecticide resistance before being released in the field for biological control [37].

The results presented here provide a comprehensive foundation for further study of potential insecticide resistance mechanisms in beneficial predators and how they compare to crop pests.

Methods

Sample preparation and sequencing

Orius laevigatus (commonly known as pirate bugs) were obtained from 'Bioline AgroSciences'. CO₂ was used for anaesthesia to allow the insects to be sorted from the substrate. Both adults and nymphs were then flash frozen with liquid N₂ and stored at -80°C. The whole process was done within 48 hours of arrival.

O. laevigatus's small size of ~2mm and a low yield of high molecular weight DNA per individual required ~1000 individuals to be pooled for genomic DNA/RNA extractions, which were carried out in-house at Rothamsted Research. The commercial DNAzol reagent was used for the DNA extractions, and the Bioline Isolate II RNA Mini Kit was used for the RNA extractions. The DNA and RNA were sent for library preparation and sequencing by Genewiz (New Jersey, US).

The genome assembly was developed using a hybrid assembly strategy with both Illumina short reads and Pacific Biosciences (PacBio) long reads.

Short reads were sequenced using 2mg of DNA and a library with an insert size of 200bp. Sequencing was done using Illumina HiSeq 4000 with a 2x150bp paired-end configuration. 413,143,574 reads were obtained.

To obtain long read PacBio data, 3.7mg of DNA first underwent blue pippin size selection (>=10kb) to remove low molecular weight DNA. <500ng of DNA remained after size selection, and so a low input protocol was used for library construction with an insert size of 20kb. Sequencing was done using the PacBio Sequel I platform and 537,651 reads were obtained.

Transcriptome sequencing used 10mg of RNA and a library construction with an insert size of 150bp and PolyA selection for rRNA removal. Sequencing was done using Illumina HiSeq 4000 with a 2x150bp paired-end configuration. 413,137,378 reads were obtained.

Genome metrics evaluation based on raw reads

K-mer counting of the raw Illumina DNA data was done using Jellyfish 2.2.6 [38]. Canonical (-C) 21-mers (-m 21) were counted and a histogram of k-mer frequencies outputted. GenomeScope 2.0 [39] was used to process this histogram with 'ploidy' set at 2 and 'maximum k-mer coverage cut-off' set at 10,000.

De novo genome assembly

The overall assembly pipeline is shown in figure 1.

FastQC v.0.11.8. [40] was used for quality checks on the raw Illumina HiSeq DNA and RNA sequence data. Adapters were trimmed, low-quality bases (below a score of 3) were removed from the start and end of reads and any reads with a length less than 36 bases were also removed. Trimmomatic v.0.38. [41] was used for these trimming steps.

The raw PacBio long reads were assembled into contigs with the Flye v2.5. *de novo* assembler [42,43]. Rascaf was then used to improve the Flye genome assembly with RNA-seq data [44]. Contigs were also produced with the raw PacBio long reads using Canu v1.8 [45] as well as with FALCON v1.3.0 and FALCON-Unzip, which is recommended for heterozygous/outbred organisms with diploid or higher ploidy [46,47].

QuickMerge v0.3 [48] was used to merge the assemblies, with Flye as the reference assembly. Multiple rounds of Pilon error polishing [49] were performed, using the Illumina short read data, until no further improvement in BUSCO score was seen.

Redundans [50] was used for scaffolding and redundant contig removal. Redundans is geared towards highly heterozygous genomes. Some redundant regions had to be removed manually, as Redundans does not detect redundancy when only part of the contig is duplicated. The nucmer tool from the MUMmer4 package [51] was used to detect these redundant regions through a whole genome self-alignment.

A BLAST search against the NCBI Reference Sequence (Refseq) database release 93 [52], was performed using the Tera-BLAST algorithm on a TimeLogic DeCypher system (Active Motif Inc., Carlsbad, CA). The results were processed with Megan [53] to identify any bacterial or viral sequences which were then removed manually in Geneious v10.2.6.

The mitochondrial genome sequence was identified and extracted by running a BLAST search of the *O. laevigatus* genome against the *Orius sauteri* mitochondrial genome which is publicly available at NCBI, GenBank accession No. KJ671626 [54].

Genome annotation

Gene prediction was performed using the MAKER v2.31.8 pipeline [55] through the incorporation of both transcriptome evidence and *ab initio* gene prediction as well as a custom repeat library (see below). Maker was run using Augustus v3.3.1 [56], GeneMark-ES v4.32 [57] and FGeneSH v8.0.0 [58] as well as EvidenceModeler v1.1.1 [59] with default masking options.

A *de novo* species specific repeat library was constructed using RepeatModeller v1.0.7 [60] to identify repeat models. These models were searched against the GenBank non-redundant (*nr*) protein database for Arthropoda (e value <10⁻³) using Blastx to remove any potential protein-coding genes. This was combined with transposon data to create a custom library. Transposons were identified from the transcriptome assembly by running HMMER: hmmscan [61] against the Pfam database [62] and filtering the resultant Pfam descriptions for those containing "transposon". A search for transposons was also done on transcripts produced from Maker and these transposons were then added to the custom repeat library which was used for a second round of Maker. RepeatMasker v4.0.7 [63] was used to mask repeats in the genome assembly using these repeat libraries, as well as to estimate the abundances of all predicted repeats.

RNA-seq reads were mapped to the genome with HISAT2 v2.0.5 [64] for assembly with StringTie v1.0.1 [65]. A *de novo* assembly was also done using Trinity v2.5.1 [66]. The best transcripts were selected from the Trinity and StringTie assemblies using Evigene v19.jan01 [67].

Evidence from ESTs (transcripts) was transferred to the genome assembly via Maker. The output from this was then used to produce a high confidence level gene model training set - overlapping and redundant gene models were removed. Augustus and GeneMark were trained using this training set prior to being used for *ab initio* gene predictions. FGeneSH was run based on the *Drosophila melanogaster* genome.

The best transcripts from both the *ab initio* gene prediction annotation and the transcriptome-based annotation were selected using Evigene and combined to create the final annotation.

Orius laevigatus protein sequences were aligned using Blastp against the non-redundant (*nr*) NCBI protein database for Arthropoda. InterProscan searches were run against several databases (CDD, HAMAP, HMMPanther, HMMPfam, HMMPIR, FPrintScan, BlastProDom, ProfileScan, HMMTigr) for functional annotation. BLAST2GO [68] was used to assign gene ontology (GO annotations). Infernal v1.1.2 [69] was used to predict and annotate non-coding RNAs.

The mitochondrial genome was annotated using MITOS2 [70] with reference database 'RefSeq 81 Metazoa' and genetic code '5 Invertebrate'.

Genome quality assessment

Basic metrics from the genome assembly were calculated using a script developed for the 'Assemblathon' [71].

The completeness of the genome assembly and annotation for *Orius laevigatus* was assessed using the Benchmarking Universal Single-Copy Orthologs (BUSCO) [72] of the insect gene set (insecta odb9). 'Fly' was used as the training species for Augustus gene prediction. BUSCO assessments were then run with default parameters.

Comparative Genomics and Phylogenetic Analysis

To produce the species tree, orthogroup gene trees were produced using Orthofinder [73] and the tree was inferred from these using the STAG method [74].

In order to identify genes potentially involved in insecticide resistance, the PFAM domains assigned to gene models during annotation (as described in the 'Genome Annotation' methods section) were used as follows: CCEs (PF00135/IPR002018), GSTs (IPR004045/PF02798), (IPR004046/PF00043), P450s (IPR001128/PF00067), ABCs (IPR003439/PF00005) and UGTs (IPR002213/PF00201). Proteins from UniProt for the classes of interest, from hemipteran species, were used for BLAST queries against *O. laevigatus* to exclude false positives and to assist with subfamily assignment within these classes. These proteins were also used in phylogenetic trees which were produced using MAFFT alignments and RaxML v8.2.11 [75] to further assist with subfamily assignment.

Manual checks and curation were performed for genes potentially involved in insecticide resistance. Increased copy numbers of these genes often led to adjacent tandem duplications being incorrectly annotated as one gene model, therefore curation was important to prevent incorrect gene numbers being reported in later analyses. The exon/intron boundaries and start/stop codons of the genes were confirmed through visualization in IGV [76] of RNAseq data mapped to the genome using HISAT2 v2.0.5 [64] and the gene models were edited in Geneious where necessary.

The P450s were classified and named by Dr David Nelson [77]. The UGTs were classified and named by Dr Michael Court [78].

Results And Discussion

Sequencing

~1,000 individuals of *O. laevigatus* were needed to produce enough DNA and RNA for sequencing. Since they were obtained commercially, the level of inbreeding of the culture was not known. However, all individuals were obtained from a single colony within the rearing facility. A high heterozygosity level was therefore a possibility and this was kept in mind when making decisions during the assembly process.

Raw data

For the DNA sequencing, 537,651 PacBio reads were generated with a total length of 6Gb (44x) and an N50 of 11,287bp.

A total of 123 Gb (820x) was produced in the Illumina HiSeq platform for whole genome sequencing, as well as 123Gb (820x) for transcriptome sequencing.

Quality trimming of reads using Trimmomatic to remove adapters and any reads <36bp resulted in a 0.2% loss of reads for whole genome sequencing and a 5% loss of reads for transcriptome sequencing (table 1).

Table 1. Number of paired-end Illumina HiSeq DNA sequences present before and after trimming

	Illumina DNA Reads	Illumina RNA Reads
Total sequences before trimming	413,143,574	413,137,378
Total sequences after trimming	412,474,208	389,150,727
Sequences lost	669,366	23,986,651

Genome metrics evaluation based on raw reads

The raw read k-mer analysis with GenomeScope 2.0 estimated a haploid genome size of ~141Mb (table 2), in line with the final assembly size of 151Mb. Total genome repeat size was 20Mb, 16.5% of the total estimated genome size.

Table 2. Genome characteristics obtained from GenomeScope v2.0

Using a k-mer length of 21 and a maximum k-mer coverage of 10,000

Genome characteristics	Minimum	Maximum
Heterozygosity, %	1.197	1.297
Genome Haploid Length (Mb)	140.7	142.2
Genome Repeat Length (Mb)	20.2	20.5
Genome Unique Length (Mb)	120.4	121.8
Read Error Rate, %	0.86	0.86

The heterozygosity rate ranged from 1.20% to 1.30%. This alongside the small 'shoulder' to the left of the main 'full-model' peak (figure 2), indicates a fairly high level of heterozygosity, which was taken into consideration in the assembly strategy.

Assembly

Flye, Falcon and Canu were used to produce 3 separate genome assemblies. The statistics for these assemblies, as well as for subsequent versions of the assembly outlined in this section are shown in Additional file 2. Rascaf improved the contiguity of the Flye assembly through alignment of the RNA-seq data to the genome, likely because it is less affected by the use of multiple individuals versus genome assembly tools which include non-conserved sequences from a population of individuals. Falcon Unzip improved the Falcon assembly contiguity with a 4.5-fold decrease in the total number of scaffolds (although this coincided with a ~9% loss of complete gene models found using BUSCO).

Flye (both with and without Rascaf) had the best assembly statistics in terms of scaffold N50 and BUSCO score. However, Falcon Unzip achieved the largest 'longest scaffold' of the three assemblers.

Quickmerge was used to merge the Falcon-Unzip assembly, Rascaf improved Flye assembly and the Canu assembly. The resultant merged assembly had better continuity than any of the stand-alone assemblies, however, the BUSCO completeness was slightly worse (and worsened with the second round of Quickmerge). This was likely due to mis-assemblies in the component assemblies causing alignment issues, which resulted in sections of the misassembled contigs being discarded.

Pilon was used for error polishing and improved the BUSCO completeness score. Redundans (redundancy removal and scaffolding/gap-closing) improved the scaffold N50 and removed redundant scaffolds.

A comparison of the gene models (core insect genes from the insecta odb9 BUSCO gene set) found in the original Flye / Falcon / Canu assemblies versus the merged assembly showed that some of the gene models were found in at least one of the original assemblies, but were missing in the merged assembly. Of the 154 missing or fragmented genes in the merged assembly (out of a total 1,658 core insect genes), 5 were found in the Falcon assembly, 5 in the Flye assembly and 46 in the Canu assembly. Manual editing to bring the full-length contigs containing these missing genes into the merged assembly took the BUSCO completeness score up by 5%. A final round of Pilon improved this score by an additional 0.5% (further rounds of Pilon did not improve the score).

This brought the final assembly statistics to 93.6% BUSCO (insecta) complete, scaffold N50: 125,649bp, the longest scaffold: 2,051,674bp and 89.4% of scaffolds >10k in length (table 3).

Table 3. Final assembly statistics for the *O. laevigatus* genome

Number of scaffolds	2,050
Total size of scaffolds	150,957,203 bp
Longest scaffold	2,051,674 bp
Shortest scaffold	1,007 bp
Number of scaffolds > 1K nt	2,050 (100.0%)
Number of scaffolds > 10K nt	1,832 (89.4%)
Number of scaffolds > 100K nt	386 (18.8%)
Number of scaffolds > 1M nt	4 (0.2%)
Number of scaffolds > 10M nt	0 (0.0%)
N50 scaffold length	125,649 bp
Number of N's	21,965 *
Number of gaps	187 *

*(1 gap was 17,239 N's, and another gap was 1,243 N's. All other gaps were <100 N's.)

Annotation

Gene prediction with MAKER identified 15,102 protein-coding genes with the encoded proteins having a mean length of 464 amino acids. Of these, 12,949 (86%) had a match to NCBI's non-redundant (*nr*) database and 11,616 (77%) contained InterPro motifs, domains or signatures. In total, 13,112 (87%) were annotated with either blastp or InterPro and 10,192 were annotated with a GO ID. More information on the InterPro member database annotations is given in additional file 1. The longest protein found was an 'egf-like protein' at 14,628 amino acids.

From the Infernal tool inference of RNA alignments, a total of 791 non-coding RNA elements and 269 *cis*-regulatory elements were found in the genome (table 4).

Table 4. Number of ncRNAs predicted in the *Orius laevigatus* genome

ncRNA element	Number of elements
tRNA	503
rRNA	182
snRNA	53
miRNA	41
srpRNA	6
snoRNA	3
lncRNA	3

Repeat Annotation

Transposable and repetitive elements made up 27.07% of the assembled *O. laevigatus* genome (table 5) and the majority of these (20.4%) were unclassified repeats. This is close to the reported repeat content of other Hemipteran species, for example: *Cimex lectularius* - 31.63% [79] and *Acyrtosiphon pisum* - 38% [15]. One exception is *Rhodnius prolixus* which has an unusually low repeat content of 5.6% [80].

Table 5. Summary of transposable and repetitive elements in the *Orius laevigatus* genome

	Number of Elements	Length Occupied	Percentage of Sequence
SINES	705	59,683 bp	0.04%
LINES	3,309	1,556,653 bp	1.03%
LINE1	0	0 bp	0.00%
LINE2	496	257,681 bp	0.17%
L3/CR1	2,310	890,133 bp	0.59%
LTR elements	959	501,171 bp	0.33%
DNA elements	5,490	1,715,984 bp	1.14%
hAT-Charlie	784	222,164 bp	0.15%
TcMar-Tigger	99	41,650 bp	0.03%
Unclassified	105,531	30,830,578 bp	20.42%
Total interspersed repeats	NA	34,664,069 bp	22.96%
Small RNA	127	35,035 bp	0.02%
Satellites	4,867	3,456,707 bp	2.29%
Simple repeats	30,022	2,273,603 bp	1.51%
Low complexity	7,742	444,236 bp	0.29%
Total:	NA	42,285,278 bp	27.07%

Mitochondrial Genome

A circularized mitochondrial genome of 16,246bp, assembled and annotated using MITOS2, consisted of 13 protein coding genes, 19 tRNA genes, 2 rRNA genes and an A+T rich region with a length of 1,460bp and an A+T content of 72.7% (figure 3). This closely matches the *Orius sauteri* mitochondrial genome, which is also 16,246bp and has 13 protein-coding genes, 22 tRNA genes, 2rRNA genes and an A+T rich region of 1,758 bp and an A+T content of 73.5% [54].

Phylogeny

OrthoFinder assigned 318,985 genes (88.8% of total) to 27,481 orthogroups. There were 1,621 orthogroups with all species present and 45 of these consisted entirely of single-copy genes.

Phylogenetic analysis correctly clustered *O. laevigatus* within the Hemipteran clade (figure 4) and identified *Cimex lectularius* as its closest relative.

Comparative Genomics

ABC Transporters

ATP-binding cassette transporters (ABCs), the largest known group of active transporters, can eliminate xenobiotic compounds - such as secondary metabolites produced by plants or insecticides - through translocation [32]. These transporters are subdivided into eight subfamilies: ABCA-H. ABCB, ABCC and ABCG are the subfamilies most associated with resistance to a variety of insecticides including pyrethroids, carbamates, organophosphates and neonicotinoids [81]. 41 of the 64 transporters in *O. laevigatus* belong to these 3 class-specific expansions (table 6) which could confer resistance to insecticides (a phylogenetic tree showing relationships of ABC transporters in *O. laevigatus* is included in Additional file 3, full sequences are included in Additional file 4).

Table 6 shows a comparison of numbers of ABC transporter genes found in the current study with those reported for some pest species. The gene family expansions were generally seen in the ABCC and ABCG classes for all hemipteran species and slightly larger expansions were seen in some crop pests compared to *O. laevigatus* for the ABCC class, however, the expansions were of very similar size for both crop pests and *O. laevigatus* in the ABCG class. Overall, the total numbers of ABC transporter genes were similar across all the hemipteran species compared.

Table 6. Numbers of ABC transporter genes annotated in *O. laevigatus* (this study), *C. lectularius* [82], *L. hesperus* [83], *F. occidentalis* [84], *T. palmi* [85], *A. gossypii* [86], *T. vaporariorum* [87] *D. noxia* and *B. tabaci* [88]

	<i>O. laevigatus</i> + close relatives			Crop pests					
	<i>Orius laevigatus</i>	<i>Cimex lectularius</i>	<i>Lygus hesperus</i>	<i>Frankliniella occidentalis</i>	<i>Thrips palmi</i>	<i>Diuraphis noxia</i>	<i>Aphis gossypii</i>	<i>Trialeurodes vaporariorum</i>	<i>Bemisia tabaci</i>
ABCA	11	6	11	3	3	3	4	3	8
ABCB	9	7	6	5	4	6	5	9	3
ABCC	9	6	12	19	12	24	25	7	6
ABCD	1	2	2	2	2	3	2	4	2
ABCE	1	1	1	1	2	1	1	1	1
ABCF	5	4	3	3	3	3	4	3	3
ABCG	23	23	19	22	16	26	30	9	23
ABCH	2	2	11	13	7	11	0	9	9
Total	64	51	65	70	49	77	71	45	55

Glutathione S-Transferases

The glutathione S-transferases (GSTs) protein family is large and functionally diverse, and is known to confer resistance to all main insecticide classes. GST-mediated detoxification of insecticides takes place via several different mechanisms, including protecting against oxidative stress, binding and sequestration of the insecticide and by catalysing the conjugation of glutathione to the insecticide to reduce their toxicity [33].

The number of GST genes in *O. laevigatus* was fairly similar to other hemipteran close relatives, with the exception of the sigma class, which was notably larger (table 7, full sequences included in Additional file 4). Of the 16 genes in the sigma class, 9 genes (mRNA13082 and mRNA13086-13093) were adjacent on the same scaffold, indicating a lineage specific expansion (fig 5). Expansions in this class have been reported in several hemipteran species including *Triatoma infestans* and *Myzus persicae* [89,90]. The sigma class has been found to play an important role in detoxification of organophosphorus insecticides in hemipteran species [91], therefore this expansion could potentially confer some tolerance to organophosphates in *O. laevigatus*. The delta and epsilon classes of GSTs are linked to insecticide resistance to pyrethroids [92,93] and the delta class is much larger in the crop pest species than in *O. laevigatus* and its close relatives which suggests these crop pests could exhibit a higher level of GST-mediated pyrethroid resistance.

Table 7. Numbers of GST genes annotated in *Orius laevigatus* (this study), *Cimex lectularius* [94], *Rhodnius prolixus*, *Triatoma infestans* [90], *Thrips palmi* [85], *Myzus persicae*, *Acyrtosiphon pisum*, *Trialeurodes vaporariorum* and *Bemisia tabaci* [95].

	<i>O. laevigatus</i> + close relatives				Crop pests				
	<i>Orius laevigatus</i>	* <i>Cimex lectularius</i>	<i>Rhodnius prolixus</i>	<i>Triatoma infestans</i>	<i>Thrips palmi</i>	<i>Myzus persicae</i>	<i>Acyrtosiphon pisum</i>	<i>Trialeurodes vaporariorum</i>	<i>Bemisia tabaci</i>
Delta	1	1	1	1	14	3	16	9	14
Epsilon	0	0	0	0	0	0	1	1	0
Omega	2	1	1	0	1	1	2	0	1
Sigma	16	5	7	9	6	12	6	3	6
Theta	1	2	3	2	1	1	2	0	0
Zeta	1	1	1	0	2	0	0	2	2
Microsomal	3	0	1	2	1	2	2	3	2
Total	24	10	14	14	25	19	29	18	25

* *C. lectularius* numbers may be an underestimate as sequencing coverage was low for this study.

Carboxyl/cholinesterases

Many esterases are linked to detoxification of organophosphorus, carbamate and pyrethroid insecticides and acetylcholinesterase (AChE) is the target for organophosphate and carbamate insecticides, with amino acid substitutions being linked to resistance [35]. 32 members of the CCE superfamily, including 1 AChE gene, were found in the *O. laevigatus* genome (table 8, full sequences included in Additional file 4) which is a similar number to that reported for *Cimex lectularius*, which had 30 CCE genes and 2 AChE genes [82].

Table 8. Numbers of CCEs annotated in *Orius laevigatus* (this study), *Cimex lectularius* [94], *Rhodnius prolixus* [96], *Triatoma infestans* [90], *Frankliniella occidentalis* [84], *Myzus persicae* [89], *Acyrtosiphon pisum*, *Bemisia tabaci* [97] and *Trialeurodes vaporariorum* [98] and their distribution across classes and clades.

Class/Clade	<i>O. laevigatus</i> + close relatives				Crop pests					
	<i>Orius laevigatus</i>	<i>Cimex lectularius</i>	<i>Rhodnius prolixus</i>	<i>Triatoma infestans</i>	<i>Frankliniella occidentalis</i>	<i>Myzus persicae</i>	<i>Acyrtosiphon pisum</i>	<i>Trialeurodes vaporariorum</i>	<i>Bemisia tabaci</i>	
Dietary class	0	0*	22 (0)***	0	28	5	5	12	6	
Hormone/semiochemical processing class	16	20*	9 (31)***	18	7	12	16	6	19	
Neuro-developmental class	Glutactins	1	0*	2	0	2	0	0	1	1
	AChE	2	1*	2	1	2	3	2	2	4
	uncharacterised	1	1*	2	0	2	1	1	1	1
	gliotactin	3	0*	1	0	1	1	1	1	1
	neuroligin	8	0*	4	0	7	0	3	3	10
	neurotactin	1	0*	1	0	1	0	0	1	0
	Subtotal	16	2*	12	1	15	5	7	9	17
Total	32	22*	43	19	50	22	28	27	42	
		(30)**								

* *C. lectularius* numbers may be an underestimate as sequencing coverage was low for this study, clade assignment was also uncertain as a result.

** A more recent study [82] found 30 CCE genes in *C. lectularius*, and is more likely to be a true representation, but they had not been assigned into classes/clades.

*** Numbers in brackets represent the possible true numbers of *R. prolixus* CCEs, based on a potential misassignment of 22 genes to the dietary class instead of the hormone/semiochemical processing class

The dietary class of CCEs is involved in insecticide and xenobiotic detoxification [99]. *O. laevigatus* has no genes within this class, in line with *T. infestans* and *C. lectularius*, whereas the crop pest species (i.e. thrips, aphids and whiteflies) all have at least 5 members in this class (table 8). *R. prolixus* has 22 genes which have been classed as dietary, however this assignment was based heavily on a species-specific expansion which is characteristic of the dietary class. The real number of genes in the dietary class for *R. prolixus* may be 0, since this clade of 22 genes clusters with the hormone/semiochemical class in both the *R. prolixus* study [96] and this study (Additional file 5). A lack of dietary esterases in *R. prolixus* would make sense, as *R. prolixus*, *C. lectularius* and *T. infestans* are all blood-sucking insects and do not require dietary esterases to process the secondary metabolites found in plants. This could also explain why *O. laevigatus*, a beneficial predator of crop pests in both nymph and adult life stages, does not require dietary esterases.

The dietary class is involved in pyrethroid resistance [100], however, *T. infestans* exhibits pyrethroid esterase activity despite having no dietary esterases [101]. *O. laevigatus* has also shown the ability to develop pyrethroid resistance - although the exact mechanism of this resistance is not yet known [102]. The hormone and semiochemical processing class is also involved in insecticide metabolism, due to the presence of β -esterases [103,104]. There may be some redundancy in genes potentially involved in insecticide detoxification from the dietary and hormone/semiochemical processing classes. This might explain why only one of these classes shows an increased number of genes for each of these hemipteran species (table 8), as having increased numbers of both classes would be redundant, whilst very low numbers of both classes would be detrimental. The lack of the dietary class may therefore not impact the xenobiotic resistance abilities of *O. laevigatus*, as it has 16 genes within the hormone/semiochemical processing class.

The remaining CCEs in *O. laevigatus* belong to the neurodevelopmental class and include the neuroligins, gliotactins, glutactins and neurotactins, which are non-catalytic due to the lack of a critical serine residue. Acetylcholinesterase is the only protein in this class which has been linked to organophosphate resistance [105,106].

UGTs

UDP-glucosyltransferases (UGTs) are detoxification enzymes which are speculated to be involved in insecticide metabolism. Although the exact mechanisms of UGT-mediated resistance have not yet been identified, their upregulation has been shown in resistant strains of *P. xylostella* [31] and they have been linked to diamide resistance in *Chilo suppressalis* [107] and neonicotinoid resistance in *Diaphorina citri* [108] and they also contribute to insecticide detoxification via the elimination of oxidative stress in *Apis cerana* [109].

The number of UGT genes in *O. laevigatus* was much lower than for other hemipteran species (table 9). The UGTs were submitted to Dr Michael Court for naming and are included in Additional file 4. Numbers of UGTs have been reported to be lower in non-phytophagous insects [85], which could explain the low numbers seen in *O. laevigatus* and *R. prolixus* compared to crop pests. This suggests that UGT-mediated detoxification may be lower in *O. laevigatus* than in crop pests.

Table 9. Numbers of UDP glucuronosyltransferase genes found in *O. laevigatus* (this study), *Rhodnius prolixus*, *Tetranychus urticae*, *Nilaparvata lugens*, *Acyrtosiphon pisum*, *Bemisia tabaci* [19], *Myzus persicae* [110] and *Trialeurodes vaporariorum* [111].

	<i>O. laevigatus</i> + close relatives		Crop pests					
	<i>Orius laevigatus</i>	<i>Rhodnius prolixus</i>	<i>Tetranychus urticae</i>	<i>Nilaparvata lugens</i>	<i>Myzus persicae</i>	<i>Acyrtosiphon pisum</i>	<i>Trialeurodes vaporariorum</i>	<i>Bemisia tabaci</i>
Total	10	16	81	20	101	72	55	76

Cytochrome P450s

Cytochrome P450s are a diverse superfamily capable of metabolizing a huge variety of endogenous and exogenous substrates. In insects they are associated with growth and development, metabolism of pesticides and plant toxins as well as the production and metabolism of insect hormones and pheromones. P450s are associated with resistance to insecticides from a variety of classes, including pyrethroids, carbamates and neonicotinoids. They are also linked to the activation of organophosphates and other pro-insecticides [34]. Upregulation of P450s in insects has been shown to confer insecticide resistance [112–115], and conversely downregulation occurs in response to pro-insecticides [116,117].

A total of 58 full-length P450 genes were identified in the *O. laevigatus* genome, 11 P450 fragment genes were also found as well as 1 pseudogene. (Sequences are included in Additional file 4). These sequences were named by Dr David Nelson using his in-house pipeline [77]. The majority of these genes (34) belonged to the diverse CYP3 class, which was a similar size to other hemipteran species (table 10).

Table 10. Total numbers of P450s genes annotated in *O. laevigatus* (this study), *C. lectularius* [82], *R. prolixus*, *T. infestans* [90], *F. occidentalis*, *T. palmi* [84], *M. persicae*, *A. pisum* [89] *T. vaporariorum* [87] and *B. tabaci* [118].

	<i>O. laevigatus</i> + close relatives				Crop pests					
	<i>Orius laevigatus</i>	<i>Cimex lectularius</i>	<i>Rhodnius prolixus</i>	<i>Triatoma infestans</i>	<i>Frankliniella occidentalis</i>	<i>Thrips palmi</i>	<i>Myzus persicae</i>	<i>Acyrtosiphon pisum</i>	<i>Trialeurodes vaporariorum</i>	<i>Bemisia tabaci</i>
CYP2	6	6	7	1	12	12	3	10	7	18
CYP3	34(41)*	36	55	65	22	26	63	33	41	76
CYP6	11	10**	8	15	18	-	-	29	34	47
CYP9	0	0**	0	0	0	-	-	0	0	0
Other	23	26**	47	50	4	-	-	4	7	-
CYP4	13(17)*	11	49	22	37	42	48	32	25	73
Mitochondrial	5	6	8	6	10	11	1	8	7	4
Total	58	59	119	94	81	91	115	83	80	171

*Values in brackets represent total gene numbers including partial and fragment genes. For other species partial and fragment p450 genes were excluded in cases where they were listed as such - some may remain in the counts if official naming and curation had not taken place.

**Values used are those from [82], but values differed by study - [119] identified 5 CYP9s, 35 CYP6s and 5 others, [120] identified 0 CYP9s, 8 CYP6s and 15 others (these were also officially named by David Nelson)

The CYP3 clade is currently the P450 clade most associated with insecticide resistance - notably the CYP6 and CYP9 families [121]. Interestingly the CYP9 family was not present in *O. laevigatus*, as found for *T. infestans* and *R. prolixus* [90]. Further investigation into the assignment of classes within the CYP3 clade suggests the lack of the CYP9 class could be a common feature within hemiptera (table 10).

Expansion of the CYP397 gene family was seen in *O. laevigatus*, (fig 6) with 7 full-length CYP397 genes and 1 fragment. CYP397B1, CYP397B2, CYP397B6 and CYP397C1 were directly adjacent on the same scaffold, indicating tandem duplications. Sequence similarity of the CYP397 genes to CYP397B1 ranged from 52% to 86%, which suggests a variation in ages of these tandem duplications. *Cimex lectularius* also showed an increased copy number of CYP397 with 6 copies (A1-A6) [82]. CYP397A1 is significantly upregulated (>36 fold) in pyrethroid-resistant strains of *C. lectularius* [120], therefore the expansion of this gene family could potentially confer some tolerance to pyrethroids in *O. laevigatus*

A previous study [122] looked at the effect of insecticide synergists on *Orius tristicolor* (another pirate bug of the Anthocoridae family), and found that PBO (piperonyl butoxide, an inhibitor of P450s and esterases) significantly increased the mortality rate when combined with indoxacarb (an oxadiazine insecticide). Whereas inhibition of solely GSTs or esterases did not reduce mortality. Upregulation of P450s, esterases and GSTs have all been seen in response to oxadiazines [123], therefore the fact that only P450 inhibition had an impact on mortality rate suggests P450s may be the primary detoxification mechanism of *O. laevigatus*.

Target site mutations

Point mutations resulting in amino acid substitutions in the target proteins of insecticides have been characterised in many insecticide resistant insect species, including in the sodium channel gene *para* which confers resistance to pyrethroids [124], the acetylcholinesterase-1 (*ace-1*) enzyme associated with

organophosphate resistance [125] and the acetyl-coenzyme A carboxylase (ACC) enzyme linked to keto-enol (e.g. spiromesifen) resistance [126]. Despite these mutations having been observed in a variety of hemipteran crop pests, none were observed in *O. laevigatus*. This is likely because, unlike crop pests, there has been no selection pressure for these mutations in populations of the predator.

The ryanodine receptor (RyR) is the target of diamide insecticides, and two target site resistance mutations conferring amino acid substitutions (I4790M and G4946E - numbering according to *Plutella xylostella*, PxRyR) have been identified in lepidopteran pests [21,127]. Interestingly, *O. laevigatus* wildtype RyR has the I4790M substitution (full sequence for RyR in Additional file 4) which has been shown to confer varying levels of resistance to diamides. M4790 was also present in other hemipteran species as shown in figure 7 (except for *Lygus hesperus* which had L4790). I4790M has been detected in lepidopteran populations across the globe and is considered to be a 'selectivity switch' for diamides [128]. *Orius tristicolor* showed high levels of resistance to chlorantraniliprole (a diamide insecticide) with <5% mortality [122] with the I4790M substitution being the main cause [129]. It is therefore possible that I4790M may confer some tolerance to diamides in *O. laevigatus*, and indeed, diamide resistance has been reported in *O. laevigatus* [130]. However, I4790M could potentially also confer diamide tolerance in crop pests - diamide resistance has already been shown in *F. occidentalis* [130]. Therefore this would likely not be an exploitable difference for IPM strategies.

Overall, tolerance of insecticides by *O. laevigatus* resulting from target site differences seems unlikely compared to what is seen in crop pests, where there has been intensive selection pressure. Although, it is important to note that the *O. laevigatus* assembly was a consensus of ~1000 individuals, therefore differences in target sites would likely only be apparent if they were present in the majority of the population.

Conclusions

PacBio long-read technology combined with low error-rate short-read Illumina sequencing enabled the production of a high-quality genome and mitochondrial assembly for *O. laevigatus*. Despite the need to pool many individuals for sequencing, the genome completeness is still of a sufficient quality to aid with comparative and functional genomics analyses and provides a useful first reference genome for the *Anthocoridae* family. Hi-C based scaffolding would likely be the next best step to significantly improve this genome in the future.

Comparative analyses of *O. laevigatus* with hemipteran crop pests showed evidence of possible differences in xenobiotic tolerance, including a potential increase in GST-mediated tolerance of organophosphates in *O. laevigatus*, whilst GST-mediated pyrethroid tolerance may be more prevalent in crop pests. There may also be less UGT-mediated tolerance to diamides and neonicotinoids in *O. laevigatus* compared to crop pests - although, the I4790M target site mutation may confer some degree of diamide insensitivity to *O. laevigatus*.

A recent study shows that there is significant variation in the susceptibility of *O. laevigatus* to pyrethroids when a variety of wild and commercial populations are assessed [102]. This suggests that beneficial predators such as *O. laevigatus* are certainly capable of developing insecticide resistance, but a combination of factors result in resistance developing slower than in pest species. This could be due to beneficial predators having smaller population sizes, longer life cycles, less exposure to pesticides and a lack of continuous selection pressure - beneficial predators often need to be re-released each season as populations migrate to new areas in search of food sources. These differences will have resulted in a lesser degree of selection for resistance mechanisms in *O. laevigatus* and therefore any observed differences in potential sensitivity would only be at low levels. Further comparisons looking at differences in gene expansions, expression levels and key target site mutations between resistant and susceptible strains of *O. laevigatus* would provide more concrete evidence for the findings in this study.

In conclusion, this study indicates differences in potential mechanisms of resistance between crop pests and *O. laevigatus* which could be exploited when selecting targeted insecticides. An increase in the number of pesticides which are safe for beneficial predators such as *O. laevigatus* would be of significant impact to pest management, especially at a time when the list of pesticides effective against crop pests is growing ever shorter. The findings also suggest that *O. laevigatus* has the ability to develop resistance to a variety of insecticides which could be used to our advantage through the selective breeding and selection of resistant strains of *O. laevigatus* for use in IPM strategies.

Declarations

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The genome assembly generated in this study is available in the GenBank repository, under BioProject: PRJNA721944. The *Orius laevigatus* genes used for comparative analyses are included in the additional files.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the Pest Genomics Initiative, a collaborative project between Rothamsted Research, Bayer Crop Science and Syngenta AG. Rothamsted Research gratefully acknowledges strategic funding from the UK Biotechnology and Biological and Sciences Research Council (UKRI-BBSRC). The content is solely the responsibility of the authors and does not necessarily reflect the official views of the funding bodies.

Authors' contributions

The Pest Genomics Initiative (BB, CR, CZ, EG, KH-P, LF, RK and RN) devised the original conceptual ideas. EB performed the DNA and RNA extractions with assistance from MW. EB assembled and annotated the genome with guidance from RK and DH. EB performed the comparative analyses with input from RN. FM, RK and KH-P supervised the project. EB wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Acknowledgements go to Dr David R. Nelson of the International Committee on the Nomenclature for Cytochrome P450 Enzymes for classification of *Orius laevigatus* P450s and to Dr Michael H. Court of the UGT Nomenclature Committee for classification of *Orius laevigatus* UGTs.

References

1. Oerke E-C. Crop losses to pests. *J Agric Sci* 2006,144:31–43. <https://doi.org/10.1017/s0021859605005708>.
2. Geiger F, Bengtsson J, Berendse F, Weisser WW, Emmerson M, Morales MB, et al. Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. *Basic Appl Ecol* 2010,11:97–105. <https://doi.org/10.1016/j.baae.2009.12.001>.
3. Bottrell DG, Schoenly KG. Resurrecting the ghost of green revolutions past: The brown planthopper as a recurring threat to high-yielding rice production in tropical Asia. *J Asia Pac Entomol* 2012,15:122–40. <https://doi.org/10.1016/j.jaspen.2011.09.004>.
4. Debach P, Rosen D. Biological control by natural enemies (second edition). *J Trop Ecol* 1992,8:216–216. <https://doi.org/10.1017/s0266467400006374>.
5. Hernandez LM. A review of the economically important species of the genus *Orius* (Heteroptera: Anthocoridae) in East Africa. *Journal of Natural History* 1999,33:543–68. <https://doi.org/10.1080/002229399300245>.
6. Kelton LA. The Anthocoridae of Canada and Alaska: Heteroptera, Anthocoridae. Agriculture Canada, 1978.
7. Malais M. Knowing and Recognizing: The Biology of Glasshouse Pests and Their Natural Enemies. 1992.
8. Chambers RJ, Long S, Helyer NL. Effectiveness of *Orius laevigatus*(Hem.: Anthocoridae) for the control of *Frankliniella occidentalis* on cucumber and pepper in the UK. *Biocontrol Sci Technol* 1993,3:295–307. <https://doi.org/10.1080/09583159309355284>.
9. Sanchez JA, Sanchez JA, Lacasa A. Modelling population dynamics of *Orius laevigatus* and *O. albidipennis* (Hemiptera: Anthocoridae) to optimize their use as biological control agents of *Frankliniella occidentalis* (Thysanoptera: Thripidae). *Bull Entomol Res* 2002,92:77–88. <https://doi.org/10.1079/BER2001136>.
10. Tommasini MG, Maini S, Nicoli G. Advances in the integrated pest management in protected-eggplant crops by seasonal inoculative releases of *Orius laevigatus*. *Adv Hortic Sci* 1997,11:182–8.
11. Hoy MA, Waterhouse RM, Wu K, Estep AS, Ioannidis P, Palmer WJ, et al. Genome Sequencing of the Phytoseiid Predatory Mite *Metaseiulus occidentalis* Reveals Completely Atomized Hox Genes and Superdynamic Intron Evolution. *Genome Biol Evol* 2016,8:1762–75. <https://doi.org/10.1093/gbe/evw048>.
12. Werren JH, Richards S, Desjardins CA, Niehuis O, Gadau J, Colbourne JK, et al. Functional and Evolutionary Insights from the Genomes of Three Parasitoid *Nasonia* Species. *Science* 2010,327:343–8. <https://doi.org/10.1126/science.1178028>.
13. Ando T, Matsuda T, Goto K, Hara K, Ito A, Hirata J, et al. Repeated inversions within a pannier intron drive diversification of intraspecific colour patterns of ladybird beetles. *Nat Commun* 2018,9:1–13. <https://doi.org/10.1038/s41467-018-06116-1>.
14. Tribolium Genome Sequencing Consortium, Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, et al. The genome of the model beetle and pest *Tribolium castaneum*. *Nature* 2008,452:949–55. <https://doi.org/10.1038/nature06784>.
15. International Aphid Genomics Consortium. Genome sequence of the pea aphid *Acyrtosiphon pisum*. *PLoS Biol* 2010,8:e1000313. <https://doi.org/10.1371/journal.pbio.1000313>.
16. You M, Yue Z, He W, Yang X, Yang G, Xie M, et al. A heterozygous moth genome provides insights into herbivory and detoxification. *Nat Genet* 2013,45:220–5. <https://doi.org/10.1038/ng.2524>.
17. Xue J, Zhou X, Zhang C-X, Yu L-L, Fan H-W, Wang Z, et al. Genomes of the rice pest brown planthopper and its endosymbionts reveal complex complementary contributions for host adaptation. *Genome Biol* 2014,15:521. <https://doi.org/10.1186/s13059-014-0521-0>.
18. Nicholson SJ, Nickerson ML, Dean M, Song Y, Hoyt PR, Rhee H, et al. The genome of *Diuraphis noxia*, a global aphid pest of small grains. *BMC Genomics* 2015,16:1–16. <https://doi.org/10.1186/s12864-015-1525-1>.
19. Chen W, Hasegawa DK, Kaur N, Kliot A, Pinheiro PV, Luan J, et al. The draft genome of whitefly *Bemisia tabaci* MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. *BMC Biol* 2016,14:1–15. <https://doi.org/10.1186/s12915-016-0321-y>.
20. Wenger JA, Cassone BJ, Legeai F, Johnston JS, Bansal R, Yates AD, et al. Whole genome sequence of the soybean aphid, *Aphis glycines*. *Insect Biochem Mol Biol* 2020,123:102917. <https://doi.org/10.1016/j.ibmb.2017.01.005>.
21. Wang X, Cao X, Jiang D, Yang Y, Wu Y. CRISPR/Cas9 mediated ryanodine receptor I4790M knockin confers unequal resistance to diamides in *Plutella xylostella*. *Insect Biochem Mol Biol* 2020,125:103453. <https://doi.org/10.1016/j.ibmb.2020.103453>.
22. Schoville SD, Chen YH, Andersson MN, Benoit JB, Bhandari A, Bowsher JH, et al. A model species for agricultural pest genomics: the genome of the Colorado potato beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). *Sci Rep* 2018,8:1–18. <https://doi.org/10.1038/s41598-018-20154-1>.

23. European Commission. Directive 2009/128/EC on the sustainable use of pesticides. Official Journal of the European Union, 2009. <https://doi.org/10.2861/78>.
24. Cameron PJ, Walker GP, Hodson AJ, Kale AJ, Herman TJB. Trends in IPM and insecticide use in processing tomatoes in New Zealand. *Crop Prot* 2009,28:421–7. <https://doi.org/10.1016/j.cropro.2009.01.002>.
25. Kranthi KR, Russell DA. Changing Trends in Cotton Pest Management. In: Peshin R, Dhawan AK, editors. *Integrated Pest Management: Innovation-Development Process: Volume 1*, Dordrecht: Springer Netherlands, 2009, p. 499–541. https://doi.org/10.1007/978-1-4020-8992-3_17.
26. Meissle M, Mouron P, Musa T, Bigler F, Pons X, Vasileiadis VP, et al. Pests, pesticide use and alternative options in European maize production: current status and future prospects. *J Appl Entomol* 2009,134:357–75. <https://doi.org/10.1111/j.1439-0418.2009.01491.x>.
27. Hillocks RJ. Farming with fewer pesticides: EU pesticide review and resulting challenges for UK agriculture. *Crop Prot* 2012,31:85–93. <https://doi.org/10.1016/j.cropro.2011.08.008>.
28. Lechenet M, Dessaint F, Py G, Makowski D, Munier-Jolain N. Reducing pesticide use while preserving crop productivity and profitability on arable farms. *Nature Plants* 2017,3:1–6. <https://doi.org/10.1038/nplants.2017.8>.
29. Lattin JD. BIONOMICS OF THE ANTHOCORIDAE. *Annu Rev Entomol* 1999,44:207–31. <https://doi.org/10.1146/annurev.ento.44.1.207>.
30. Heckel DG. Insecticide Resistance After Silent Spring. *Science* 2012,337:1612–4. <https://doi.org/10.1126/science.1226994>.
31. Li X, Shi H, Gao X, Liang P. Characterization of UDP-glucuronosyltransferase genes and their possible roles in multi-insecticide resistance in *Plutella xylostella* (L.). *Pest Manag Sci* 2018,74:695–704. <https://doi.org/10.1002/ps.4765>.
32. Merzendorfer H. Chapter One - ABC Transporters and Their Role in Protecting Insects from Pesticides and Their Metabolites. In: Cohen E, editor. *Advances in Insect Physiology*, vol. 46, Academic Press, 2014, p. 1–72. <https://doi.org/10.1016/B978-0-12-417010-0.00001-X>.
33. Pavlidis N, Vontas J, Van Leeuwen T. The role of glutathione S-transferases (GSTs) in insecticide resistance in crop pests and disease vectors. *Curr Opin Insect Sci* 2018,27:97–102. <https://doi.org/10.1016/j.cois.2018.04.007>.
34. Scott JG. Cytochromes P450 and insecticide resistance. *Insect Biochem Mol Biol* 1999,29:757–77. [https://doi.org/10.1016/S0965-1748\(99\)00038-7](https://doi.org/10.1016/S0965-1748(99)00038-7).
35. Sogorb MA, Vilanova E. Enzymes involved in the detoxification of organophosphorus, carbamate and pyrethroid insecticides through hydrolysis. *Toxicol Lett* 2002,128:215–28. [https://doi.org/10.1016/S0378-4274\(01\)00543-4](https://doi.org/10.1016/S0378-4274(01)00543-4).
36. Rane RV, Ghodke AB, Hoffmann AA, Edwards OR, Walsh TK, Oakeshott JG. Detoxifying enzyme complements and host use phenotypes in 160 insect species. *Current Opinion in Insect Science* 2019,31:131–8. <https://doi.org/10.1016/j.cois.2018.12.008>.
37. Roderick GK, Navajas M. Genes in new environments: genetics and evolution in biological control. *Nat Rev Genet* 2003,4:889–99. <https://doi.org/10.1038/nrg1201>.
38. Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 2011,27:764–70. <https://doi.org/10.1093/bioinformatics/btr011>.
39. Rhyker Ranallo-Benavidez T, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun* 2020,11:1–10. <https://doi.org/10.1038/s41467-020-14998-3>.
40. Andrews S. FastQC n.d. <https://github.com/s-andrews/FastQC> (accessed April 20, 2021).
41. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014,30:2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
42. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 2019,37:540–6. <https://doi.org/10.1038/s41587-019-0072-8>.
43. Lin Y, Yuan J, Kolmogorov M, Shen MW, Chaisson M, Pevzner PA. Assembly of long error-prone reads using de Bruijn graphs. *Proc Natl Acad Sci U S A* 2016,113:E8396–405. <https://doi.org/10.1073/pnas.1604560113>.
44. Song L, Shankar DS, Florea L. Rascaf: Improving Genome Assembly with RNA Sequencing Data. *Plant Genome* 2016,9. <https://doi.org/10.3835/plantgenome2016.03.0027>.
45. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 2017,27:722–36. <https://doi.org/10.1101/gr.215087.116>.
46. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 2013,10:563–9. <https://doi.org/10.1038/nmeth.2474>.
47. Chin C-S, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, et al. Phased diploid genome assembly with single-molecule real-time sequencing. *Nat Methods* 2016,13:1050–4. <https://doi.org/10.1038/nmeth.4035>.
48. Chakraborty M, Baldwin-Brown JG, Long AD, Emerson JJ. Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. *Nucleic Acids Res* 2016,44:e147. <https://doi.org/10.1093/nar/gkw654>.
49. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS One* 2014,9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
50. Przytycki LP, Gabaldón T. Redundans: an assembly pipeline for highly heterozygous genomes. *Nucleic Acids Res* 2016,44:e113–e113. <https://doi.org/10.1093/nar/gkw294>.
51. Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. MUMmer4: A fast and versatile genome alignment system. *PLoS Comput Biol* 2018,14:e1005944. <https://doi.org/10.1371/journal.pcbi.1005944>.

52. O'Leary NA, Wright MW, Brister JR, Ciufu S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 2016,44:D733–45. <https://doi.org/10.1093/nar/gkv1189>.
53. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. *Genome Res* 2007,17:377–86. <https://doi.org/10.1101/gr.5969107>.
54. Du B-Z, Niu F-F, Wei S-J. The complete mitochondrial genome of the predatory bug *Orius sauteri* (Poppius) (Hemiptera: Anthocoridae). *Mitochondrial DNA Part A* 2016,27:777–8. <https://doi.org/10.3109/19401736.2014.915537>.
55. Holt C, Yandell M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 2011,12:1–14. <https://doi.org/10.1186/1471-2105-12-491>.
56. Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res* 2004,32:W309–12. <https://doi.org/10.1093/nar/gkh379>.
57. Lomsadze A, Ter-Hovhannisyan V, Chernoff YO, Borodovsky M. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res* 2005,33:6494–506. <https://doi.org/10.1093/nar/gki937>.
58. Solovyev V. Statistical approaches in eukaryotic gene prediction. In: D.J. Balding, M. Bishop, C. Cannings, editor. *Handbook of Statistical Genetics*, Chichester: John Wiley & Sons Ltd, 2001, p. 83–127.
59. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, et al. Automated eukaryotic gene structure annotation using EvidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol* 2008,9:1–22. <https://doi.org/10.1186/gb-2008-9-1-r7>.
60. Smit AFA, Hubley R. RepeatModeler Open-1.0 2008-2015. <http://www.repeatmasker.org>.
61. Wheeler TJ, Eddy SR. nhmmer: DNA homology search with profile HMMs. *Bioinformatics* 2013,29:2487–9. <https://doi.org/10.1093/bioinformatics/btt403>.
62. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. *Nucleic Acids Res* 2014,42:D222–30. <https://doi.org/10.1093/nar/gkt1223>.
63. Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0. 2013-2015.
64. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 2015,12:357–60. <https://doi.org/10.1038/nmeth.3317>.
65. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 2015,33:290–5. <https://doi.org/10.1038/nbt.3122>.
66. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 2011,29:644–52. <https://doi.org/10.1038/nbt.1883>.
67. Gilbert D. EvidentialGene - Evidence Directed Gene Construction for Eukaryotes 2013. <https://sourceforge.net/projects/evidentialgene/>.
68. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 2008,36:3420–35. <https://doi.org/10.1093/nar/gkn176>.
69. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 2013,29:2933–5. <https://doi.org/10.1093/bioinformatics/btt509>.
70. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsche G, et al. MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol* 2013,69:313–9. <https://doi.org/10.1016/j.ympev.2012.08.023>.
71. Bradnam K. A script to calculate a basic set of metrics from a genome assembly 2011. http://korflab.ucdavis.edu/Datasets/Assemblathon/Assemblathon2/Basic_metrics/assemblathon_stats.pl.
72. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015,31:3210–2. <https://doi.org/10.1093/bioinformatics/btv351>.
73. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019,20:1–14. <https://doi.org/10.1186/s13059-019-1832-y>.
74. Emms DM, Kelly S. STAG: Species Tree Inference from All Genes. *bioRxiv* 2018:267914. <https://doi.org/10.1101/267914>.
75. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014,30:1312–3. <https://doi.org/10.1093/bioinformatics/btu033>.
76. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol* 2011,29:24–6. <https://doi.org/10.1038/nbt.1754>.
77. Nelson DR. The Cytochrome P450 Homepage. *Hum Genomics* 2009,4:1–7. <https://doi.org/10.1186/1479-7364-4-1-59>.
78. UGT Committee. UGT Committee Home. UGT Committee Home n.d. <https://prime.vetmed.wsu.edu/resources/udp-glucuronosyltransferase-homepage> (accessed March 25, 2021).
79. Rosenfeld JA, Reeves D, Brugler MR, Narechania A, Simon S, Durrett R, et al. Genome assembly and geospatial phylogenomics of the bed bug *Cimex lectularius*. *Nat Commun* 2016,7:1–10. <https://doi.org/10.1038/ncomms10164>.
80. Mesquita RD, Vionette-Amaral RJ, Lowenberger C, Rivera-Pomar R, Monteiro FA, Minx P, et al. Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc Natl Acad Sci U S A* 2015,112:14936–41. <https://doi.org/10.1073/pnas.1506226112>.
81. Dermauw W, Van Leeuwen T. The ABC gene family in arthropods: Comparative genomics and role in insecticide transport and resistance. *Insect Biochem Mol Biol* 2014,45:89–110. <https://doi.org/10.1016/j.ibmb.2013.11.001>.

82. Benoit JB, Adelman ZN, Reinhardt K, Dolan A, Poelchau M, Jennings EC, et al. Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nat Commun* 2016;7:1–10. <https://doi.org/10.1038/ncomms10165>.
83. Joe Hull J, Chaney K, Geib SM, Fabrick JA, Brent CS, Walsh D, et al. Transcriptome-Based Identification of ABC Transporters in the Western Tarnished Plant Bug *Lygus hesperus*. *PLoS One* 2014;9:e113046. <https://doi.org/10.1371/journal.pone.0113046>.
84. Rotenberg D, Baumann AA, Ben-Mahmoud S, Christiaens O, Dermauw W, Ioannidis P, et al. Genome-enabled insights into the biology of thrips as crop pests. *BMC Biol* 2020;18:142. <https://doi.org/10.1186/s12915-020-00862-9>.
85. Guo S-K, Cao L-J, Song W, Shi P, Gao Y-F, Gong Y-J, et al. Chromosome-level assembly of the melon thrips genome yields insights into evolution of a sap-sucking lifestyle and pesticide resistance. *Mol Ecol Resour* 2020;20:1110–25. <https://doi.org/10.1111/1755-0998.13189>.
86. Pan Y, Zeng X, Wen S, Gao X, Liu X, Tian F, et al. Multiple ATP-binding cassette transporters genes are involved in thiamethoxam resistance in *Aphis gossypii* glover. *Pestic Biochem Physiol* 2020;167:104558. <https://doi.org/10.1016/j.pestbp.2020.104558>.
87. Pym A, Singh KS, Nordgren Å, Emyr Davies TG, Zimmer CT, Elias J, et al. Host plant adaptation in the polyphagous whitefly, *Trialeurodes vaporariorum*, is associated with transcriptional plasticity and altered sensitivity to insecticides. *BMC Genomics* 2019;20:1–19. <https://doi.org/10.1186/s12864-019-6397-3>.
88. Tian L, Song T, He R, Zeng Y, Xie W, Wu Q, et al. Genome-wide analysis of ATP-binding cassette (ABC) transporters in the sweetpotato whitefly, *Bemisia tabaci*. *BMC Genomics* 2017;18:1–16. <https://doi.org/10.1186/s12864-017-3706-6>.
89. Ramsey JS, Rider DS, Walsh TK, De Vos M, Gordon KHJ, Ponnala L, et al. Comparative analysis of detoxification enzymes in *Acyrtosiphon pisum* and *Myzus persicae*. *Insect Mol Biol* 2010;19 Suppl 2:155–64. <https://doi.org/10.1111/j.1365-2583.2009.00973.x>.
90. Traverso L, Lavore A, Sierra I, Palacio V, Martinez-Barnette J, Latorre-Estivalis JM, et al. Comparative and functional triatomine genomics reveals reductions and expansions in insecticide resistance-related gene families. *PLoS Negl Trop Dis* 2017;11:e0005313. <https://doi.org/10.1371/journal.pntd.0005313>.
91. Gawande ND, Subashini S, Murugan M, Subbarayalu M. Molecular screening of insecticides with sigma glutathione S-transferases (GST) in cotton aphid *Aphis gossypii* using docking. *Bioinformatics* 2014;10:679–83. <https://doi.org/10.6026/97320630010679>.
92. Lumjuan N, Rajatileka S, Changsom D, Wicheer J, Leelapat P, Prapanthadara L-A, et al. The role of the *Aedes aegypti* Epsilon glutathione transferases in conferring resistance to DDT and pyrethroid insecticides. *Insect Biochem Mol Biol* 2011;41:203–9. <https://doi.org/10.1016/j.ibmb.2010.12.005>.
93. Vontas JG, Small GJ, Hemingway J. Glutathione S-transferases as antioxidant defence agents confer pyrethroid resistance in *Nilaparvata lugens*. *Biochem J* 2001;357:65–72. <https://doi.org/10.1042/0264-6021:3570065>.
94. Adelman ZN, Kilcullen KA, Koganemaru R, Anderson MAE, Anderson TD, Miller DM. Deep sequencing of pyrethroid-resistant bed bugs reveals multiple mechanisms of resistance within a single population. *PLoS One* 2011;6:e26228. <https://doi.org/10.1371/journal.pone.0026228>.
95. Aidlin Harari O, Santos-Garcia D, Musseri M, Moshitzky P, Patel M, Visendi P, et al. Molecular Evolution of the Glutathione S-Transferase Family in the *Bemisia tabaci* Species Complex. *Genome Biol Evol* 2020;12:3857–72. <https://doi.org/10.1093/gbe/evaa002>.
96. Schara R, Pedrini N, Juárez MP, Nelson DR, Torres AQ, Valle D, et al. *Rhodnius prolixus* supergene families of enzymes potentially associated with insecticide resistance. *Insect Biochem Mol Biol* 2016;69:91–104. <https://doi.org/10.1016/j.ibmb.2015.06.005>.
97. Xia J, Xu H, Yang Z, Pan H, Yang X, Guo Z, et al. Genome-Wide Analysis of Carboxylesterases (COEs) in the Whitefly, (*Gennadius*). *Int J Mol Sci* 2019;20. <https://doi.org/10.3390/ijms20204973>.
98. Karatolos N. Molecular mechanisms of insecticide resistance in the greenhouse whitefly, *Trialeurodes vaporariorum*. PhD. University of Exeter, 2011. <https://doi.org/https://ore.exeter.ac.uk/repository/bitstream/handle/10036/3350/KaratolosN.pdf>.
99. Oakeshott J, Claudianos C, Campbell PM, Newcomb RD, Russell RJ. Biochemical genetics and genomics of insect esterases. In: Gilbert LI, Gill SS, Iatrou K, editors. *Insect Pharmacology: Channels, Receptors, Toxins and Enzymes*, Elsevier, 2010, p. 229–81.
100. Flores AE, Albedafño-Vázquez W, Salas IF, Badii MH, Becerra HL, Garcia GP, et al. Elevated α -esterase levels associated with permethrin tolerance in *Aedes aegypti* (L.) from Baja California, Mexico. *Pestic Biochem Physiol* 2005;82:66–78. <https://doi.org/10.1016/j.pestbp.2004.12.007>.
101. Orihuela PLS, Vassena CV, Zerba EN, Picollo MI. Relative Contribution of Monooxygenase and Esterase to Pyrethroid Resistance in *Triatoma infestans* (Hemiptera: Reduviidae) from Argentina and Bolivia. *J Med Entomol* 2014;45:298–306. <https://doi.org/10.1093/jmedent/45.2.298>.
102. Balanza V, Mendoza JE, Cifuentes D, Bielza P. Selection for resistance to pyrethroids in the predator *Orius laevigatus*. *Pest Manag Sci* 2021;77:2539–46. <https://doi.org/10.1002/ps.6288>.
103. Ganesh KN, Vijayan VA, Urmila J, Gopalan N, Prakash S. Role of esterases and monooxygenase in the deltamethrin resistance in *Anopheles stephensi* Giles (1908), at Mysore. *Indian J Exp Biol* 2002;40:583–8.
104. Prasad KM, Raghavendra K, Verma V, Velamuri PS, Pande V. Esterases are responsible for malathion resistance in *Anopheles stephensi*: A proof using biochemical and insecticide inhibition studies. *J Vector Borne Dis* 2017;54:226. <https://doi.org/10.4103/0972-9062.217613>.
105. Jyoti, Singh NK, Singh H, Singh NK, Rath SS. Multiple mutations in the acetylcholinesterase 3 gene associated with organophosphate resistance in *Rhipicephalus (Boophilus) microplus* ticks from Punjab, India. *Vet Parasitol* 2016;216:108–17. <https://doi.org/10.1016/j.vetpar.2015.12.004>.
106. Zhang Y, Li S, Xu L, Guo HF, Zi J, Wang L, et al. Overexpression of carboxylesterase-1 and mutation (F439H) of acetylcholinesterase-1 are associated with chlorpyrifos resistance in *Laodelphax striatellus*. *Pestic Biochem Physiol* 2013;106:8–13. <https://doi.org/10.1016/j.pestbp.2013.03.002>.
107. Zhao J, Xu L, Sun Y, Song P, Han Z. UDP-Glycosyltransferase Genes in the Striped Rice Stem Borer, (*Walker*), and Their Contribution to Chlorantraniliprole Resistance. *Int J Mol Sci* 2019;20. <https://doi.org/10.3390/ijms20051064>.
108. Tian F, Wang Z, Li C, Liu J, Zeng X. UDP-Glycosyltransferases are involved in imidacloprid resistance in the Asian citrus psyllid, *Diaphorina citri* (Hemiptera: Liviidae). *Pestic Biochem Physiol* 2019;154:23–31. <https://doi.org/10.1016/j.pestbp.2018.12.010>.

109. Cui X, Wang C, Wang X, Li G, Liu Z, Wang H, et al. Molecular Mechanism of the UDP-Glucuronosyltransferase 2B20-like Gene (AccUGT2B20-like) in Pesticide Resistance of *Apis cerana cerana*. *Front Genet* 2020,11:592595. <https://doi.org/10.3389/fgene.2020.592595>.
110. Pan Y, Xu P, Zeng X, Liu X, Shang Q. Characterization of UDP-Glucuronosyltransferases and the Potential Contribution to Nicotine Tolerance in *Myzus persicae*. *Int J Mol Sci* 2019,20:3637. <https://doi.org/10.3390/ijms20153637>.
111. Xie W, He C, Fei Z, Zhang Y. Chromosome-level genome assembly of the greenhouse whitefly (*Trialetrodes vaporariorum* Westwood). *Mol Ecol Resour* 2020,20:995–1006. <https://doi.org/10.1111/1755-0998.13159>.
112. Karunker I, Benting J, Lueke B, Ponge T, Nauen R, Roditakis E, et al. Over-expression of cytochrome P450 CYP6CM1 is associated with high resistance to imidacloprid in the B and Q biotypes of *Bemisia tabaci* (Hemiptera: Aleyrodidae). *Insect Biochem Mol Biol* 2008,38:634–44. <https://doi.org/10.1016/j.ibmb.2008.03.008>.
113. Liang X, Xiao D, He Y, Yao J, Zhu G, Zhu KY. Insecticide-mediated up-regulation of cytochrome P450 genes in the red flour beetle (*Tribolium castaneum*). *Int J Mol Sci* 2015,16:2078–98. <https://doi.org/10.3390/ijms16012078>.
114. Puinean AM, Foster SP, Oliphant L, Denholm I, Field LM, Millar NS, et al. Amplification of a cytochrome P450 gene is associated with resistance to neonicotinoid insecticides in the aphid *Myzus persicae*. *PLoS Genet* 2010,6:e1000999. <https://doi.org/10.1371/journal.pgen.1000999>.
115. Yang T, Liu N. Genome analysis of cytochrome P450s and their expression profiles in insecticide resistant mosquitoes, *Culex quinquefasciatus*. *PLoS One* 2011,6:e29418. <https://doi.org/10.1371/journal.pone.0029418>.
116. Main BJ, Everitt A, Cornel AJ, Hormozdiari F, Lanzaro GC. Genetic variation associated with increased insecticide resistance in the malaria mosquito, *Anopheles coluzzii*. *Parasit Vectors* 2018,11:1–9. <https://doi.org/10.1186/s13071-018-2817-5>.
117. Vlogiannitis S, Mavridis K, Dermauw W, Snoeck S, Katsavou E, Morou E, et al. Reduced proinsecticide activation by cytochrome P450 confers coumaphos resistance in the major bee parasite *Varroa destructor*. *Proc Natl Acad Sci U S A* 2021,118. <https://doi.org/10.1073/pnas.2020380118>.
118. Ilias A, Lagnel J, Kapantaidaki DE, Roditakis E, Tsigenopoulos CS, Vontas J, et al. Transcription analysis of neonicotinoid resistance in Mediterranean (MED) populations of *B. tabaci* reveal novel cytochrome P450s, but no nAChR mutations associated with the phenotype. *BMC Genomics* 2015,16:1–23. <https://doi.org/10.1186/s12864-015-2161-5>.
119. Bai X, Mamidala P, Rajarapu SP, Jones SC, Mittapalli O. Transcriptomics of the Bed Bug (*Cimex lectularius*). *PLoS One* 2011,6:e16336. <https://doi.org/10.1371/journal.pone.0016336>.
120. Zhu F, Gujar H, Gordon JR, Haynes KF, Potter MF, Palli SR. Bed bugs evolved unique adaptive strategy to resist pyrethroid insecticides. *Sci Rep* 2013,3:1–8. <https://doi.org/10.1038/srep01456>.
121. Feyereisen R. Evolution of insect P450. *Biochem Soc Trans* 2006,34:1252–5. <https://doi.org/10.1042/BST0341252>.
122. Pereira RR, Picanço MC, Santana PA Jr, Moreira SS, Guedes RNC, Corrêa AS. Insecticide toxicity and walking response of three pirate bug predators of the tomato leaf miner *Tuta absoluta*. *Agric For Entomol* 2014,16:293–301. <https://doi.org/10.1111/afe.12059>.
123. Shi L, Shi Y, Zhang Y, Liao X. A systemic study of indoxacarb resistance in *Spodoptera litura* revealed complex expression profiles and regulatory mechanism. *Sci Rep* 2019,9:1–13. <https://doi.org/10.1038/s41598-019-51234-5>.
124. Martinez-Torres D, Foster SP, Field LM, Devonshire AL, Williamson MS. A sodium channel point mutation is associated with resistance to DDT and pyrethroid insecticides in the peach-potato aphid, *Myzus persicae* (Sulzer) (Hemiptera: Aphididae). *Insect Mol Biol* 1999,8:339–46. <https://doi.org/10.1046/j.1365-2583.1999.83121.x>.
125. Alon M, Alon F, Nauen R, Morin S. Organophosphates' resistance in the B-biotype of *Bemisia tabaci* (Hemiptera: Aleyrodidae) is associated with a point mutation in an ace1-type acetylcholinesterase and overexpression of carboxylesterase. *Insect Biochem Mol Biol* 2008,38:940–9. <https://doi.org/10.1016/j.ibmb.2008.07.007>.
126. Lueke B, Douris V, Hopkinson JE, Maiwald F, Hertlein G, Papapostolou K-M, et al. Identification and functional characterization of a novel acetyl-CoA carboxylase mutation associated with ketoenol resistance in *Bemisia tabaci*. *Pestic Biochem Physiol* 2020,166:104583. <https://doi.org/10.1016/j.pestbp.2020.104583>.
127. Zuo Y-Y, Ma H-H, Lu W-J, Wang X-L, Wu S-W, Nauen R, et al. Identification of the ryanodine receptor mutation I4743M and its contribution to diamide insecticide resistance in *Spodoptera exigua* (Lepidoptera: Noctuidae). *Insect Sci* 2020,27:791–800. <https://doi.org/10.1111/1744-7917.12695>.
128. Richardson EB, Troczka BJ, Gutbrod O, Davies TGE, Nauen R. Diamide resistance: 10 years of lessons from lepidopteran pests. *J Pest Sci* 2020,93:911–28. <https://doi.org/10.1007/s10340-020-01220-y>.
129. Nauen R, Steinbach D. Resistance to Diamide Insecticides in Lepidopteran Pests. In: Horowitz AR, Ishaaya I, editors. *Advances in Insect Control and Resistance Management*, Cham: Springer International Publishing, 2016, p. 219–40. https://doi.org/10.1007/978-3-319-31800-4_12.
130. Dáder B, Colomer I, Adán Á, Medina P, Viñuela E. Compatibility of early natural enemy introductions in commercial pepper and tomato greenhouses with repeated pesticide applications. *Insect Sci* 2020,27:1111–24. <https://doi.org/10.1111/1744-7917.12723>.

Figures

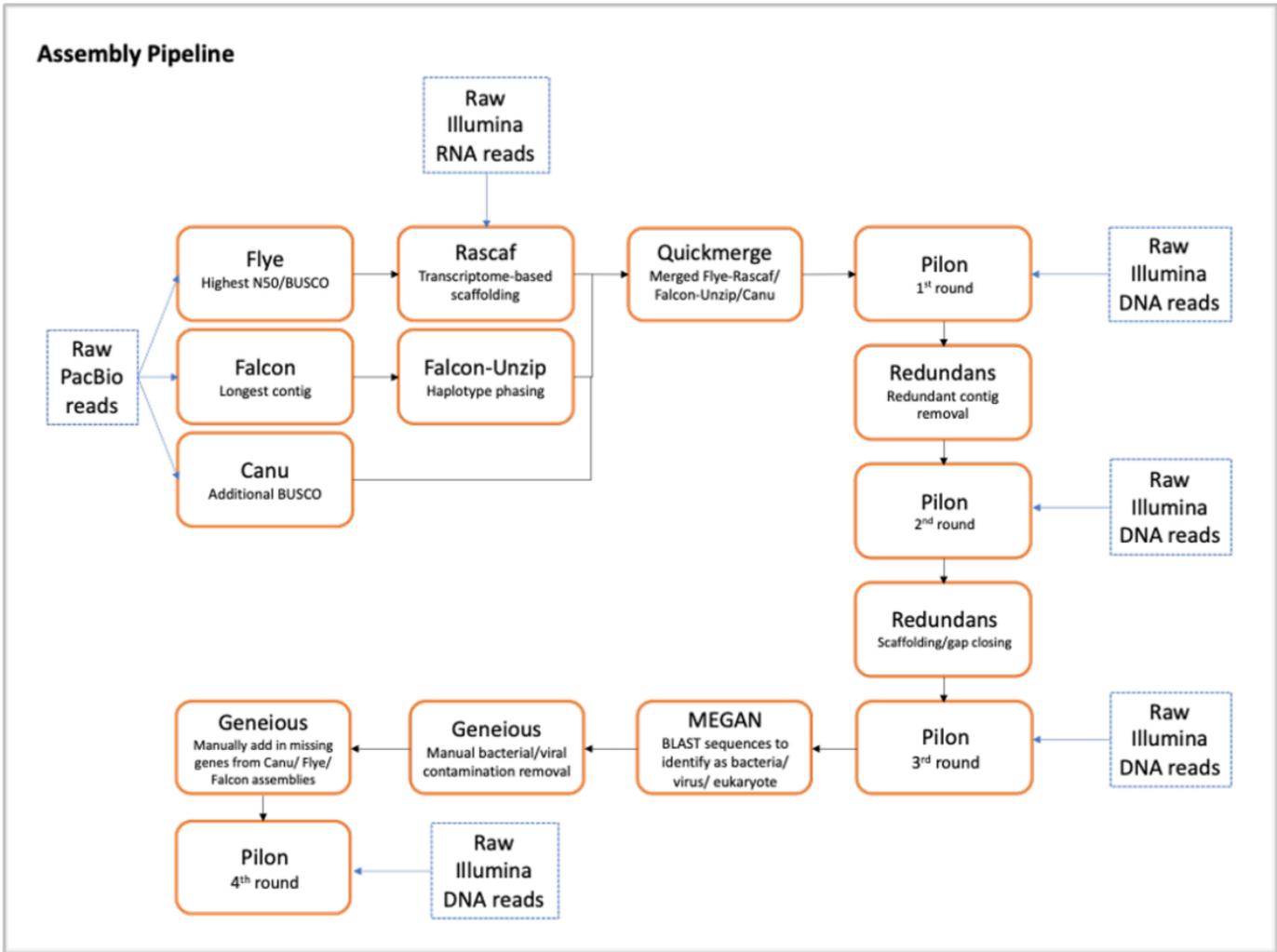


Figure 1

The assembly pipeline for the *Orius laevigatus* genome

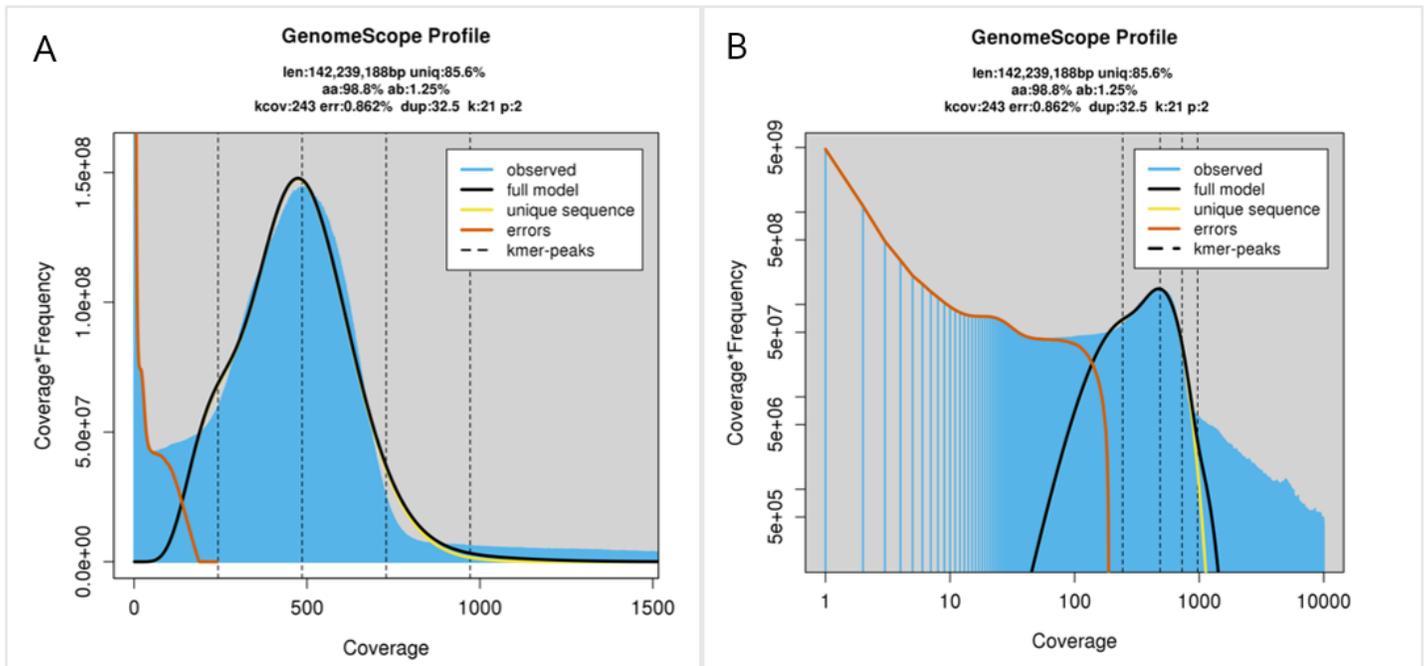


Figure 2

GenomeScope v2.0 profile plots of A: a transformed linear plot of k-mer frequency and B: a transformed log plot of k-mer coverage, at a k-mer length of 21 and a maximum k-mer coverage of 10,000.

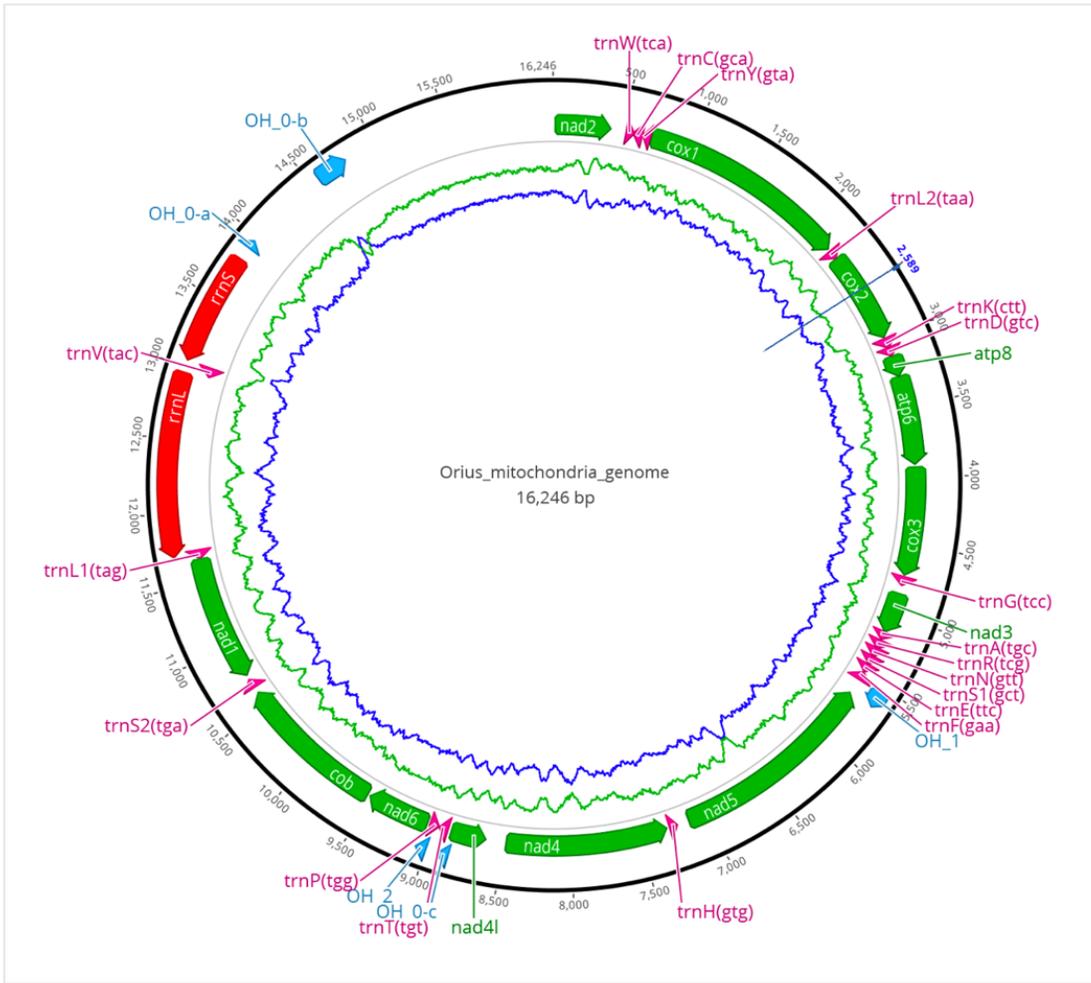


Figure 3

The mitochondrial genome for *Orius laevigatus*, visualised using Geneious and annotation track obtained using MITOS2. The innermost graphs represent AT content shown in green, and GC content shown in blue.

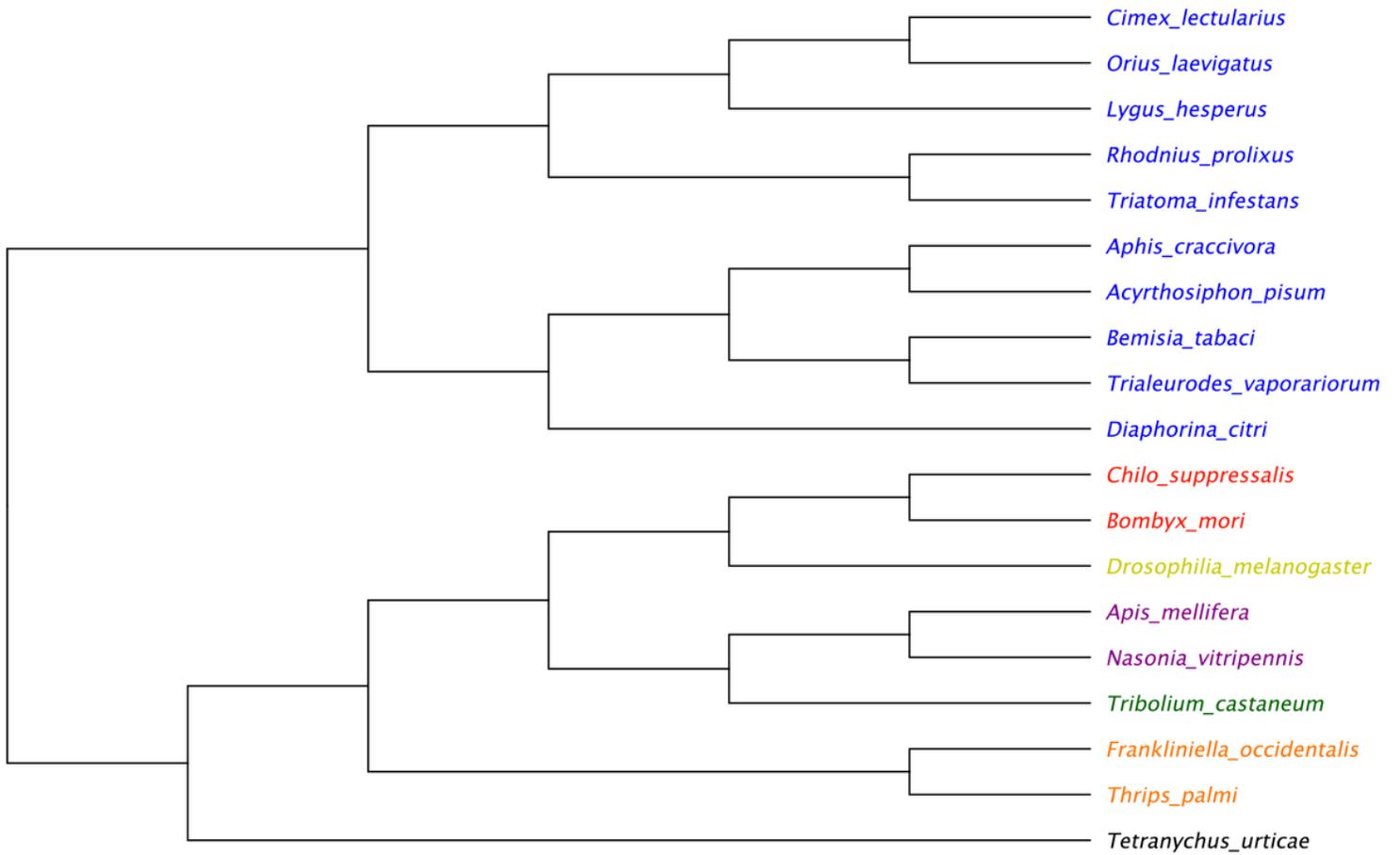


Figure 4

Phylogeny and divergence of Insecta Nodes are coloured by order, blue=Hemiptera, red=Lepidoptera, yellow=Diptera, purple=Hymenoptera, green=Coleoptera, orange=Thysanoptera, black=Chelicerata. Produced using the STAG tree inference method and full proteomes of the following species: *C. lectularius*: PRJNA167477, *L. hesperus*: PRJNA284294, *R. prolixus*: PRJNA13648, *T. infestans*: PRJNA589079, *A. craccivora*: PRJNA558689, *A. pisum*: PRJNA13657, *B. tabaci*: PRJNA312470, *T. vaporariorum*: PRJNA553773, *D. citri*: PRJNA2944, *C. suppressalis*: PRJNA506136, *B. mori*: PRJNA205630, *D. melanogaster*: PRJNA13812, *A. mellifera*: PRJNA471592, *N. vitripennis*: PRJNA575073, *T. castaneum*: PRJNA12540, *F. occidentalis*: PRJNA203209, *T. palmi*: PRJNA607431, *T. urticae*: PRJNA315122.

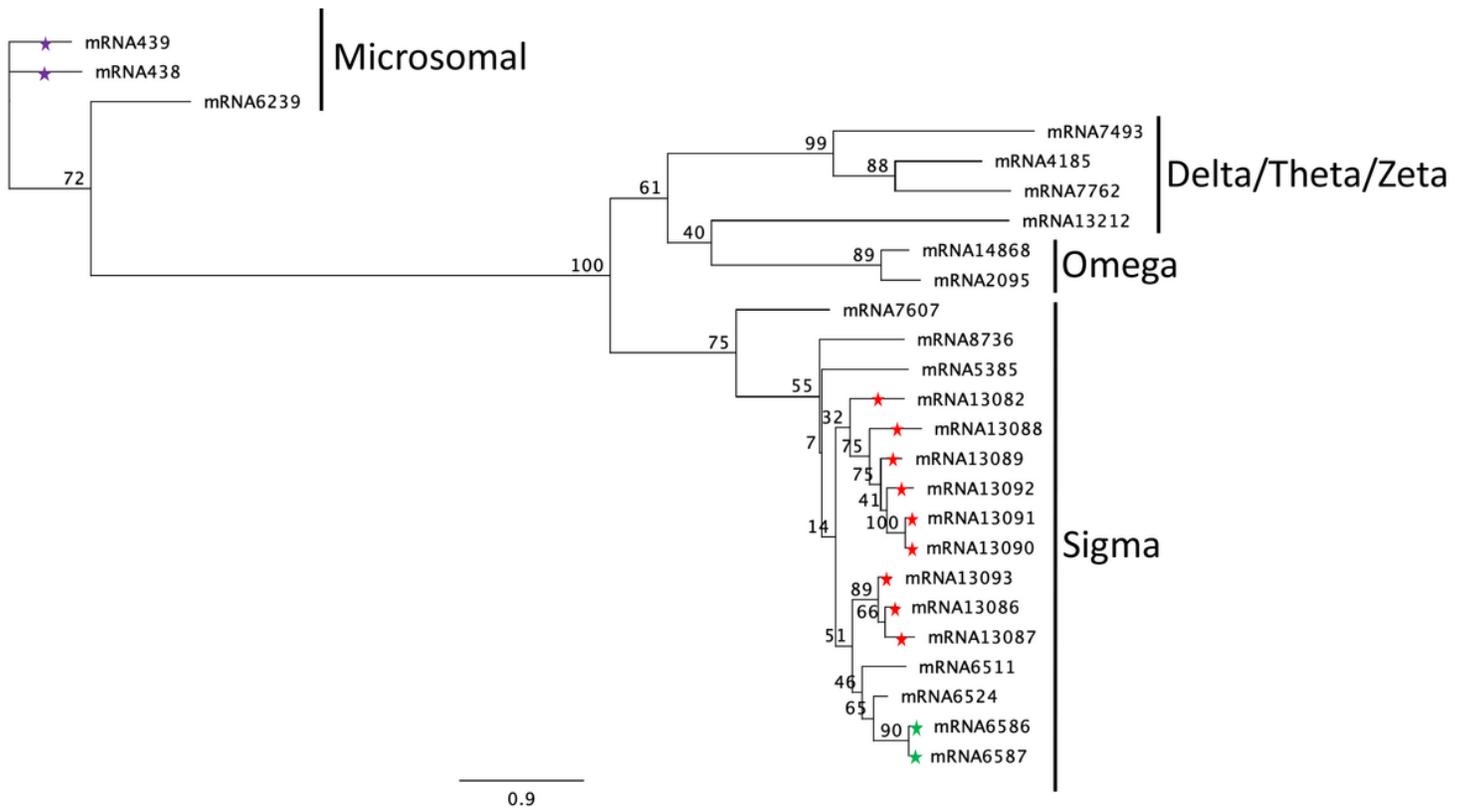


Figure 5

Phylogenetic tree of the *Orius laevigatus* glutathione S-transferases. Amino acid sequences were aligned using MAFFT and analysed using RAxML (the GAMMA LG protein model was used). The bootstrap consensus tree was inferred from 100 replicates. Coloured stars on branches indicate tandem duplications, with each colour representing a different scaffold/set of tandem duplications in the *O. laevigatus* assembly.

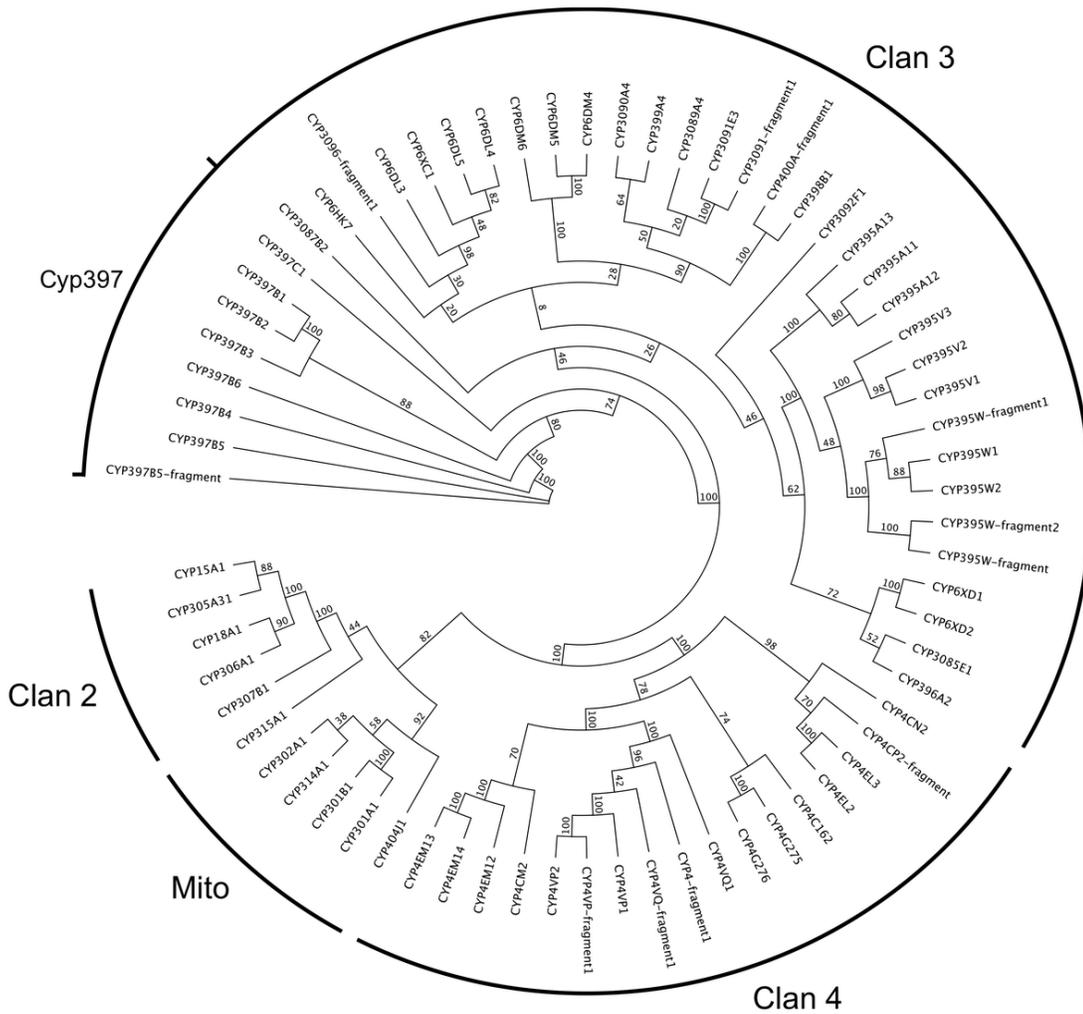
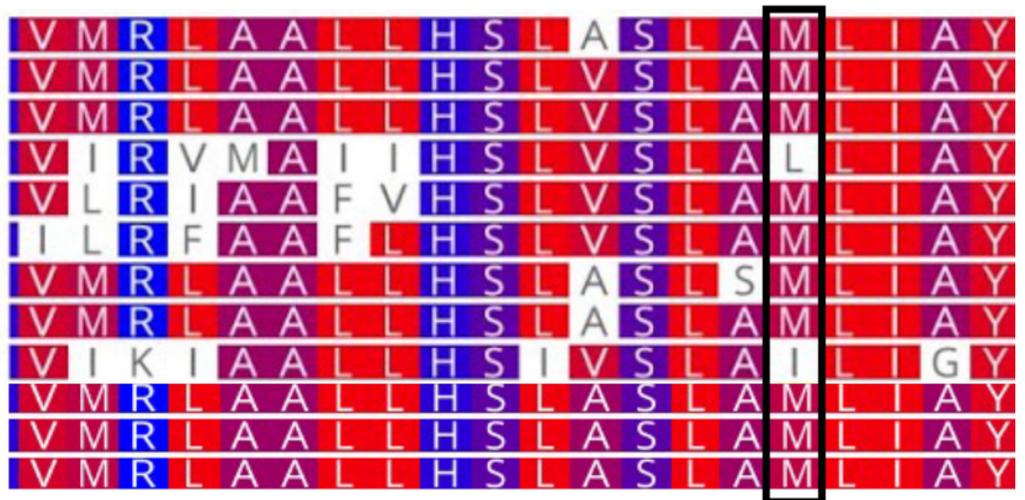


Figure 6

Phylogenetic tree of the *Orius laevigatus* cytochrome P450s. Amino acid sequences were aligned using MAFFT and analysed using RAxML (the GAMMA LG protein model was used). The bootstrap consensus tree was inferred from 100 replicates.

1. *Orius laevigatus*
2. *Cimex lectularius*
3. *Triatoma infestans*
4. *Lygus hesperus*
5. *Dialeurodes citri*
6. *Bemisia tabaci*
7. *Toxoptera citricida*
8. *Sipha flava*
9. *Plutella xylostella*
10. *Myzus persicae*
11. *Acyrtosiphon pisum*
12. *Aphis glycines*



I4790M

Figure 7

An alignment of amino acid sequences to compare transmembrane domain 3 of the conserved ryanodine receptor (RyR) from *Plutella xylostella* with hemipteran species. The box indicates the I4790M RyR point mutation linked to diamide resistance (numbered according to PxRyR). A strain of *Plutella xylostella* without the point mutation was used for the alignment. RyR sequences were obtained from the UniProt database, excluding *Orius laevigatus*. UniProt entry names are as follows: *Cimex lectularius*: A0A7E4RNZ4_CIMLE, *Triatoma infestans*: A0A023F678_TRIIF, *Lygus hesperus*: A0A146M8X0_LYGHE, *Dialeurodes citri*: A0A141BN13_DIACT, *Bemisia tabaci*: A0A1U9JHP1_BEMTA, *Toxoptera citricida*: A0A0H3XSN1_TOXCI, *Sipha flava*: A0A2S2QG40_9HEMI, *Plutella xylostella*: I3NWW8_PLUXY, *Myzus persicae*: A0A0A7RS32_MYZPE, *Acyrtosiphon pisum*: X1WXB1_ACYPI, *Aphis glycines*: A0A6G0U418_APHGL.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.docx](#)
- [Additionalfile2.docx](#)
- [Additionalfile3.docx](#)
- [Additionalfile4.xlsx](#)
- [Additionalfile5.pdf](#)