Whole-genome resequencing in the sea louse *Caligus rogercresseyi* uncovers gene duplications and copy number variants associated with pesticide resistance

Gustavo Núñez-Acuña[1,2]*, Constanza Sáez-Vera[1,2], Diego Valenzuela-Miranda[1,2], Valentina Valenzuela-Muñoz[1,2] and Cristian Gallardo-Escárate[1,2]

[1]Interdisciplinary Center for Aquaculture Research, University of Concepción, O'Higgins, Concepción, Chile, [2]Laboratory of Biotechnology and Aquatic Genomics, Department of Oceanography, Center of Biotechnology, University of Concepción, Concepción, Chile

The sea louse *Caligus rogercresseyi* is a marine ectoparasite that constitutes one of the major threats to the salmon farming industry, where the primary control strategy is the use of delousing drugs through immersion treatments. The emergence of pharmacological resistance in this copepodid species has previously been described using transcriptome data. However, the molecular mechanisms underlying chromosome rearrangements have not yet been explored. This study aimed to identify structural genomic variations and gene expression in *C. rogercresseyi* associated with pesticide sensitivity. In this study, genome resequencing was conducted using Oxford Nanopore Technology on lice strains with contrasting sensitivity to azamethiphos to detect genome duplications. Transcriptome profiling of putative gene duplications was performed by Illumina sequencing. Copy Number Variants (CNVs) were identified through comparative coverage, and collinear/tandem gene duplications over all the chromosomal regions by sequence homology. Duplications or CNVs in functional genes were primarily identified in transposable elements and genes related to the drug response, with differential expression values calculated by RNA-seq analyses of the same strains. Notably, differentially duplicated genes were found in coding regions related to cuticle proteins, suggesting that a putative resistance mechanism may be associated with cuticular structure formation and the proteins involved. Collectively, the results revealed that the intensive use of pesticides on sea lice populations increases the frequency of gene duplication, expanding the molecular elements involved in drug response. This study is the first to report an association between genome rearrangements and pharmacological resistance in sea lice populations.

KEYWORDS

genome resequencing, copy number variants, *Caligus rogercresseyi*, pharmacological resistance, *cuticle proteins*, allatostatin

1 Introduction

The emergence of pharmacological resistance in parasite species is a growing problem in animal production (Råberg et al., 2009; Srivastava and Misra-Bhattacharya, 2015; Horowitz and Ishaaya, 2016). Drug resistance in insect species is defined as the ability of most individuals in a certain population to tolerate lethal doses of a toxicant compound (Stafford and Coles, 2017). There are diverse molecular mechanisms associated with pharmacological resistance, including punctual mutations in the target proteins for a specific drug, inhibiting or reducing the drug-target binding, and then blocking or decreasing the activity of the toxicant compound (Brogdon and McAllister, 1998). Another resistance mechanism is a decrease in toxicant activity by enhancing detoxification systems of cells (Liu, 2015). The activation of proteins related to the absorption, distribution, metabolism, and excretion (ADME) of drugs from the cell is pivotal in this resistance mechanism, and has been described in various parasite species (Schröder et al., 2013). Other resistance mechanisms are associated with physiological changes that may cause adaptation to tolerate toxic compounds in individuals from a population. For instance, in the fruit fly, *Drosophila melanogaster* individuals have developed a mechanism for degrading the deltamethrin drug by the action of a trypsin protein (Xiong et al., 2014). Another example is the cuticle engrossment in the exoskeleton of mosquitoes, preventing the penetration of drugs (Lin et al., 2012). Despite the description of these mechanisms, there are various ectoparasites, including marine parasites, whose resistance mechanisms remain unknown or unclear.

The identification of genomes or chromosome structural variations associated with pharmacological resistance was elucidated before the introduction of molecular biology techniques. Molecular cytogenetics techniques have enabled the identification of genome duplications in response to pesticide resistance in various species (Jugulam and Gill, 2018). However, molecular biology and "omics" have allowed the identification of more variations in genomes associated with drug resistance. Various genome-wide association studies (GWAS) have identified numerous single-nucleotide polymorphisms (SNPs) that may explain drug-resistant genotypes in different species (Anderson et al., 2011; Oppong et al., 2019; Wan et al., 2019). Most of these SNPs are present in target genes for drugs or genes with detoxification functions, where some of them might induce differential expression of these genes, causing changes in drug sensitivity (Mahmood et al., 2016). Nonetheless, advances in genomic sciences have facilitated the discovery of gene amplification by duplication in the genomes associated with pharmacological resistance, which are well-studied in detoxifying enzymes and target genes for drugs in arthropods (Bass and Field, 2011). These duplications may involve gene expansions or limitations associated with drug sensitivity (Revie et al., 2018) or the upregulation of genes related to drug response (Karunaratne et al., 2018).

One of the origins of gene duplication in genomes is the presence of Copy Number Variants (CNVs). These CNVs are structural genome variants that imply duplications or deletions in nucleotide regions, which can vary among populations, individuals, or even at different tissues or cell lineages on the same individual (Klegarth and Eisenberg, 2018). CNVs have gained interest in recent decades because these are much more extensive than SNPs, altering a higher number of base pairs and impacting a wider proportion of the genome (Freeman et al., 2006). The association of CNVs with pesticide resistance has been described for various genes related to drug response; however, the correct identification of these structural variants depends on the transcriptome and genome sequencing technology applied (Heckel, 2022). Short-read transcriptome assemblies may overlook CNVs in the genomic regions. However, the availability of high-throughput long-read sequencing technologies could represents an unprecedented opportunity to discover and associate CNVs with pharmacological resistance in model and non-model species (Begum et al., 2021; Lavrichenko et al., 2021).

This study was conducted on the marine copepod *Caligus rogercresseyi*, which was selected as a model species to evaluate the presence of gene duplications and CNVs associated with differential pharmacological sensitivity. This species, commonly named sea louse, is a marine ectoparasite that constitutes the most significant threat among parasitic infectious diseases in Atlantic salmon in Chile (Costello, 2009; Dresdner et al., 2019), the second largest farmed salmon producing country. Sea louse is mainly controlled by drug treatments (*e.g.*, organophosphates, pyrethroids, and chitin inhibitors) directly applied by immersion treatments on salmon cages, but the efficacy of these antiparasitic compounds has been continuously reduced (Bravo et al., 2008; Marín et al., 2015; Arriagada et al., 2020). The intensive use of drug treatments has potentially led to pharmacological resistance in *C. rogercresseyi*, which has been studied from various perspectives. For example, SNP variants were found in the target gene for the organophosphate azamethiphos (Agusti-Ridaura et al., 2018) and in various genes related to drug response and the nervous system of *C. rogercresseyi* (Nuñez-Acuña et al., 2014). However, several transcriptome studies have suggested that the response of the sea louse to delousing drugs is not dependent on target genes or a reduced group of genes, but involves numerous genes with diverse functions, such as drug detoxification, as well as other metabolic and physiological roles (Valenzuela-Muñoz et al., 2015a; Valenzuela-Muñoz et al., 2015b; Núñez-Acuña et al., 2020). For example, *cuticle protein* genes that are not directly involved as target proteins for delousing drugs are differentially expressed in lice exposed to different drugs (Chávez-Mardones et al., 2016). Furthermore, the publication of the first draft of the whole genome of *C. rogercresseyi* at the chromosome scale (Gallardo-Escárate et al., 2021) is a novel opportunity to explore genomic variants associated with pharmacological resistance. This study aimed to identify novel duplications and CNVs associated with drug resistance in *C. rogercresseyi* and to infer the potential causes of the emergence of resistance genotypes in these marine parasites through a combination of long-read and short-read DNA and RNA sequencing technologies.

2 Materials and methods

2.1 Sea lice samples and culture conditions

Adult parasite samples of *C. rogercresseyi* corresponded to two previously characterized strains with contrasting sensitivities to the antiparasitic drug azamethiphos (Núñez-Acuña et al., 2020). The sensitivity to this drug was tested through a specific bioassay validated for this species, representing drug resistance/susceptibility as the EC_{50} index, effective drug concentration affecting 50% of the population (SEARCH C, 2006; Marín et al., 2018), and the efficacy in field treatments in salmon farms. The selected resistant strain had an EC_{50} index of 13.72 and azamethiphos treatment efficacy in salmon farms < 60%, while the susceptible population had EC_{50} of 0.4778 and in-field treatment efficacy > 92%. Bioassays for each population consisted of the exposure of ten parasites (five females and five males) to a gradient of different concentrations of azamethiphos, including 1, 3, 8, 10, 30, 100, 200, and 300 ppb (μg/L water) of a drug. In addition, control unexposed groups of ten parasites were also included. Exposures were applied in triplicates (three groups of ten parasites per drug concentration) in a petri dish for 30 min. This exposure time was selected because it corresponds to the therapeutic time used on salmon farms. All the concentrations were used to calculate the EC_{50} indexes, and the parasites exposed to 8 ppb and un-exposed controls were used to collect samples fixed in RNA Later solution (Ambion®, Thermo Fisher Scientific™, Waltham, MA, USA) for molecular analyses.

2.2 DNA and RNA libraries construction and sequencing

The same groups of sea lice samples were used to perform DNA and RNA sequencing, including the control (unexposed samples to azamethiphos) and parasites exposed to 8 ppb of azamethiphos. Three pools of males and females from the same conditions were used for each type of sequencing, combining five parasites of each sex. DNA was extracted using the DNeasy® Blood & Tissue Kit (QIAGEN, Hilden, Germany) following the manufacturer's protocol. Sample homogenization was performed using ceramic beads in a Mixer Mill MM 200 homogenizer (RETSCH, Haan, Germany) at a frequency of 24 Hz for 5 min. DNA concentrations were measured using a Qubit 4 fluorometer (Invitrogen, Thermo Fisher Scientific™, Waltham, MA, USA) and the Qubit 1X dsDNA BR Assay kit. The quality of the extracted DNAs was observed in the TapeStation 2200 instrument (Agilent Technologies Inc., Santa Clara, CA, USA) by calculating the DIN (DNA Integrity Number). Only samples that were more significant than 9 in DIN were used for library preparations. The initial amount of DNA samples for library preparation was 1 μg, and then the Genomic DNA by Ligation kit (SQK-LSK109) was used for Nanopore library construction (Oxford Nanopore Technologies, ONT, Oxford, United Kingdom) following the manufacturer's protocol and washing the samples applying the AMPure XP beads (Beckman Coulters, Brea, CA, USA). High-molecular weight on the DNA samples was expected. Thus 500 ng of each library was sequenced in the R.9.4. FLO-MIN106D flow cell on the MinION sequencing platform (ONT). Three sequencing replicates were conducted on each pool in separate runs.

Whole-transcriptome data corresponded to an RNA-sequencing study that was previously published using the same samples of this present study (Núñez-Acuña et al., 2020), SRA accession numbers: SRX9398676, SRX9398677, SRX9398678, SRX9398679, SRX9398646, SRX9398647, SRX9398644, SRX9398645, SRX9398656, SRX9398667, SRX9398674, SRX9398675). Briefly, total RNA from pools of parasites exposed and unexposed to azamethiphos drug was extracted by the Ribopure™ kit (Ambion®, Thermo Fisher Scientific™, Waltham, MA, USA). Next, RNA integrity was calculated in a TapeStation 2200 instrument (Agilent Technologies Inc., Santa Clara, CA, USA), calculating the RNA Integrity Number (RIN) using the R6K Screentape kit. Only samples with RIN > 8.0 were used for dscDNA library construction, which was prepared from 2 μg of high-quality RNA of each sample using the TruSeq Total RNA kit (Illumina®, San Diego, CA, USA). Then, RNA-sequencing was carried out in a MiSeq NGS sequencer (Illumina®, San Diego, CA, USA) in a 2 x 250 paired-end format on three replicates for each experimental condition.

2.3 Genome mapping and bioinformatic analyses

Adapter sequences from long DNA reads obtained by Nanopore sequencing were trimmed out using the Porechop software with default parameters and including a search in both forward/reverse directions and adapters in the middle of reads (https://github.com/rrwick/Porechop). Trimmed reads were imported to the CLC Genomic Workbench software (Qiagen Bioinformatics, Hilden, Germany) for mapping to the *C. rogercresseyi* genome previously published for this species (Gallardo-Escárate et al., 2021) (GenBank genome assembly accession number: ASM1338718v1). In parallel, RNA-seq reads obtained from the Illumina sequencing platform were imported into the same software and trimmed by quality, keeping only reads with Q30 values > 30. The Illumina reads had greater quality (fewer sequencing errors) than Nanopore reads; thus, polishing of DNA MinION reads was achieved by the MiSeq reads using a combination of the racon program (Vaser et al., 2017) and minimap long-read assembler, as was previously described (Li, 2016). The minimum sequence length was 1, the POA windows size was 500 in the racon polishing algorithm.

The DNA polished reads were mapped to the reference genome using the "long read mapping" tool in CLC using the following parameters: match score = 2; mismatch cost = 4; gap open cost = 4; gap extend cost = 2; long gap open cost = 24; long gap extend cost = 1; score bonus for global alignment = 0. Mapped reads on the genome were used to find genes in each mapping by using the "transcript discovery tool" in CLC with the following parameters:

minimum length of ORF = 100; gene merging distance = 500; minimum reads in gene = 10; minimum predicted length = 250, exon merging distance = 100. Gene lists for the susceptible and resistant strains were saved in separate files.

2.4 Duplication analyses by collinearity and tandem-duplicated genes

Duplicated genes, including collinear genes (duplicated in different chromosomes – or different regions from the same chromosome) and tandem-repeated genes, were identified through an evaluation of the presence of syntenic blocks within the *C. rogercresseyi* genome. First, the Blastall software was used to compare the homology of all the predicted genes in the resistant and the susceptible strains, then MCScanX software (Wang et al., 2012) was used by each strain to establish significant alignments on genes considering an E-value < 1E-5 from the BLAST results. Significant collinear blocks were considered if the duplicated region on the genome involved at least five predicted genes and the E-value of the whole region was < 1E-5. Alignments were associated with chromosome positions in MCScanX by exporting a simplified GFF file from CLC Genomics, including the chromosome number, gene name (correlative number as ID), start and end positions of the genes for each strain. Tandem genes were identified if at least two consecutive genes had an E-value of < 1E-20 in the respective chromosome for each strain.

A set of collinear and tandem genes was exported by each sea louse strain, and the sequence was retrieved to identify gene functions and GO terms. Blast-X analyses were conducted on the collinear and tandem genes using a simplified version of the UniProt database (http://uniprot.org), considering all the protein sequences for arthropod species and, in parallel, compared to the SwissProt revised protein databases for all eukaryote species. Sequences with lower E-values in the two databases were used for further analyses. Sequences with significant BLAST hits were used for Gene Ontology annotation and Enrichment of GO terms on biological processes, molecular functions, and cellular components in the TBtools package (Chen et al., 2020).

The genes with significant BLAST hits by each strain were used as a reference to evaluate the expression levels by mapping back the RNA-seq Illumina reads corresponding to each strain. The "RNA-seq analysis" tool in CLC Genomics was used to calculate the TPM values of the duplicated genes of untreated and parasites exposed to azamethiphos. Expression differences between these conditions and between resistant and susceptible groups were evaluated using GLM-linear models in CLC Genomics, considering as significant those genes with variations in |Fold Change| > 4 and P-value < 0.01. Volcano plots and heatmaps based on a hierarchical cluster of TPM values were constructed using TBtools to visualize gene expression differences. Heatmaps were clustered based on Manhattan's distance and with an average linkage. P-values were corrected through false discovery rate (FDR) calculation following the procedure of Benjamini and Hochberg (1995), applying multiple testing criteria with a minor modification named "automatic independent filtering", which was previously reported (Love et al., 2014).

2.5 Evaluation of duplications on *cuticle protein* genes

Due to the importance of discovering novel potential mechanisms for pharmacological resistance in the sea louse *C. rogercresseyi*, genes related to cuticle formation and penetration resistance (Balabanidou et al., 2018) were evaluated in the genomes obtained by DNA resequencing. *Cuticle protein* genes and similar genes associated with cuticle formation were searched in the BLAST results from the duplicated genes in resistant and susceptible strains. Three genes related to this function were found in the resistant strain and two in the susceptible group. The MEME Suite package (Bailey et al., 2015) was used to observe the positions of the coding sequence, UTR regions, and a specific motif of the complete *cuticle protein* genes. Syntenic alignments involving these genes were involved were highlighted in the SynVisio Synteny Browser software (Bandi, 2020). Expression analyses of cuticle genes were conducted using the same methodology applied in the previous point for duplicated genes.

2.6 Identification of Copy-Number Variants in resistant and susceptible strains

Copy Number Variants (CNVs) were identified in the genome mappings obtained by Nanopore reads through the "depth-of-coverage" method, applying the CONTRA algorithm (Li et al., 2012) incorporated in the CLC Genomics Workbench. In this analysis, mapped reads on the genome from the resistant strains were defined as "case" and mapped reads from the susceptible strain as "control" to compare the two lice populations. As the target genes for CNVs prediction, we used the overlapping genes presented in both strains with an average coverage of > 20, which corresponded to 10,330 genes. The "graining level" parameter for the prediction of CNVs was "intermediate", and the statistical cutoff for significant CNV calls was P-value < 0.05, and |fold change| > 2.0. P-values in this section were also corrected through FDR correction using the method mentioned in Section 2.4. This means that if the fold change was > 2.0, it represented an amplification in the resistant strain, and if the fold change was < -2.0, it represented a deletion in this population.

Expression analyses of genes containing significant CNVs were conducted by applying the same method for duplicated genes described in Section 2.4. In addition, an association between the position of significant CNVs with the gene density in the sea louse reference genome, mapping coverage by each strain, and collinear genes obtained in the resistant strain was conducted by constructing a CIRCOS plot in the TBTools package (Chen et al., 2020).

3 Results

3.1 Whole genome resequencing and genome mapping

A resequencing strategy was performed, including DNA and RNA sequencing from two previously described *C. rogercresseyi*

strains with contrasting drug sensitivities to azamethiphos (Núñez-Acuña et al., 2020; Núñez-Acuña et al., 2021b). One strain was resistant to this drug, and the other was susceptible. Genomic DNA reads obtained through Nanopore sequencing yielded 581,919 high-quality trimmed reads with an average length of 2,327 bp, representing 1.352 Gbases of DNA reads for the resistant strain. Transcriptome RNA sequencing obtained from the same strain using Illumina sequencing yielded 99,389,810 trimmed reads with an average length of 150.07 bp. Illumina sequencing has higher quality than third-generation sequencing technologies, such as Nanopore, so it was used to polish long reads obtained from the MinION sequencer to improve long DNA reads. The obtained polished window was 22.52% of the original Nanopore reads for the resistant strain. In this polished window, the quality of nanopore sequences was increased from an average PHRED score 17.22 to 39.65 after correction. After the polishing, 470,394 DNA reads (80.83% of the total number of reads) from the resistant strain were successfully mapped to the *C. rogercresseyi* draft genome as a reference. Nevertheless, the percentage of base pairs mapped to the reference was 93.68% because the mapped reads were longer than the 111,525 unmapped reads (2,693 versus 767 bp, respectively). The number of base pairs covered by this genome mapping was 1.267 Gbases, which represented 2.5X of the genome size, considering that the reference length was 0.506 Gbases. Regarding the full-gene sequences, the average coverage of the genes from this strain was 6.94X.

Regarding the susceptible strain, the genomic DNA obtained from Nanopore sequencing yielded 840,131 reads with an average length of 2,408 bp, representing 2.023 Gbases. Transcriptome sequencing of the susceptible strain using Illumina sequencing yielded 100,992,178 reads with an average length of 149.73 bp. Reads polishing resulted in a window of 21.71% with respect to the original number of base pairs from Nanopore reads, increasing the quality from 18.1 to 36.1 in the PHRED score. Therefore, 82.88% of polished reads were correctly mapped to the *C. rogercresseyi* genome, corresponding to 92.88% of the total length. The average length of the mapped 696,294 reads was 2,699 bp, and 1,002 bp for the 143,837 not-mapped reads. Considering this number, 1.879 Gbases of Nanopore DNA sequencing were mapped, representing 3.7 X of the reference genome. Furthermore, the average coverage of the predicted genes in this strain was 6.51X.

Putative genes were predicted through the "Transcript Discovery" tool of the CLC Genomics workbench from the mappings obtained from Nanopore reads to the reference genome. The resultant number of genes for the resistant strain was 18,438, with an average length of 15,677 bp, whereas for the susceptible strain, the number of genes was 18,525 with an average length of 15,575 bp. In addition, differential expression analyses were conducted from RNA sequencing analyses for each strain, comparing transcriptomes from parasites exposed to the azamethiphos drugs versus non-exposed lice (control), resulting in 631 differentially expressed genes (|fold change| > 4 and P-value < 0.01) in the resistant strain and 754 in the susceptible strain (Figure 1, first track).

3.2 Duplication of genes by collinearity and tandem repetitions

Gene duplications were identified in the lice genome by sequence complementarity of predicted genes, by evaluating the microsynteny arrangements throughout all the chromosomes, and then categorized as collinear or tandem genes. From the 18,438 predicted genes in the resistant strain, 602 were collinear corresponding to 3.26%, and were grouped in 42 alignments between two chromosomes consisting of at least five genes per local arrangement (see red links in Figure 1). Additionally, tandem repeated genes, defined as genes subsequently duplicated in the same chromosome, corresponded to 274 genes or 1.49% of the total number of genes in this strain. Regarding the susceptible strain, the collinear genes were much lower, corresponding to 84 genes out of 18,525 genes, being 0.45% and grouped in six different alignments of at least five genes (see green links in Figure 1). Tandem duplicated genes were 520, corresponding to the 2.81% from the 18,525 total genes.

Gene categories according to their origin were similar in both strains, except for some genes with origins by Whole-Genome
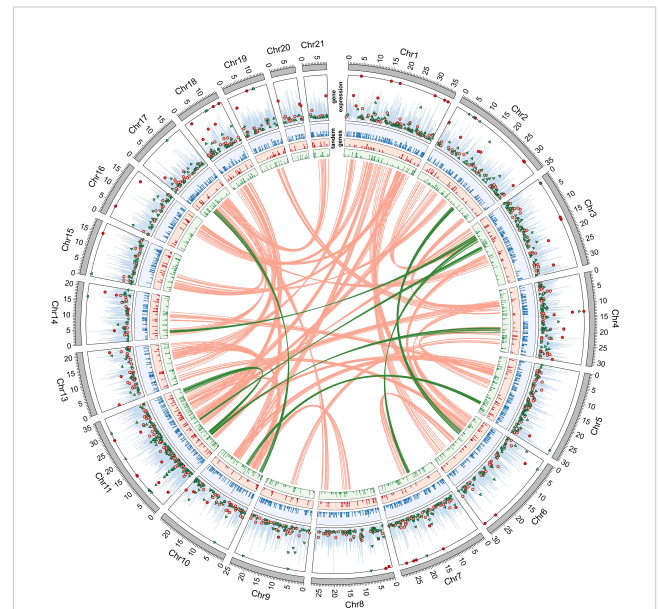

FIGURE 1
CIRCOS plot of duplicated genes in *Caligus rogercresseyi* resistant (R) and susceptible (S) strains and the chromosome regions implicated in the duplications. The first track corresponds to the gene expression of susceptible and resistant strains calculated by the Fold Change of TPM values of parasites treated with azamethiphos versus control samples. Only significant expression changes are shown (|Fold change| > 4.0; P-value < 0.01). Red circles correspond to resistant strain, green triangles to susceptible strain, and blue lines to the reference TPM values of the sea lice genome. The second track corresponds to the position of duplicated genes in tandem. Blue tiles correspond to the tandem genes in the reference genome, red to the resistant strain, and green to the susceptible strain. Links correspond to duplicated collinear genes in different pairs of genome regions. Red links correspond to collinear genes in the resistant strain and green to the susceptible strain.

Duplications (WGD) or Segmental events, which were higher in the resistant strain (21.16%) than in the susceptible strain (18.3%). These categories also implied a slight reduction in the % of the Dispersed genes in the resistant strain compared to the susceptible lice (Figure 2A). The collinear genes are among the Whole-Genome Duplications (WGD) or Segmental events. Collinear and tandem duplicated genes were annotated by Gene Ontology (Figure 2B). Enrichment analyses of GO terms in all the duplicated genes on both strains showed that the nucleus is a cellular compartment enriched for collinear and tandem duplicated genes. Regarding molecular functions, the most enriched GO terms were similar for both types of duplicated genes, most of which were related to DNA integration, metabolic process and recombination, and transposition. Regarding this last term, the transposable elements (TEs) were highly abundant among the duplicated genes after Blast results (Supplementary File 1). Molecular functions were also similar for both tandem and collinear duplicated genes, being different terms related to catalytic activity on DNA, nucleases, and polymerase activity.

3.3 Expression analyses of collinear and tandem duplicated genes from resistant and susceptible sea lice strains

Transcriptome expression patterns were analyzed in the R and S lice strain, focusing on collinear and tandem-duplicated genes (Figure 3). The 686 collinear genes (including those collinear genes in R and those in the S strains) and the 794 tandem-duplicated genes had expression patterns with differential expression trends in R/S strains and samples exposed to azamethiphos (Figures 3A, D). Hierarchical clustering of TPM (transcripts per million of reads) values for the four experimental groups was observed to be associated with the drug exposure and not according to the resistant/susceptible strain. However, the clustering of features observed in the heatmaps showed various groups of genes that were differentially expressed when comparing the resistant and susceptible lice strains. Statistical comparisons were conducted in different pairs of samples to validate these differences (Figures 3B, C, E, F). Among all the collinear genes,
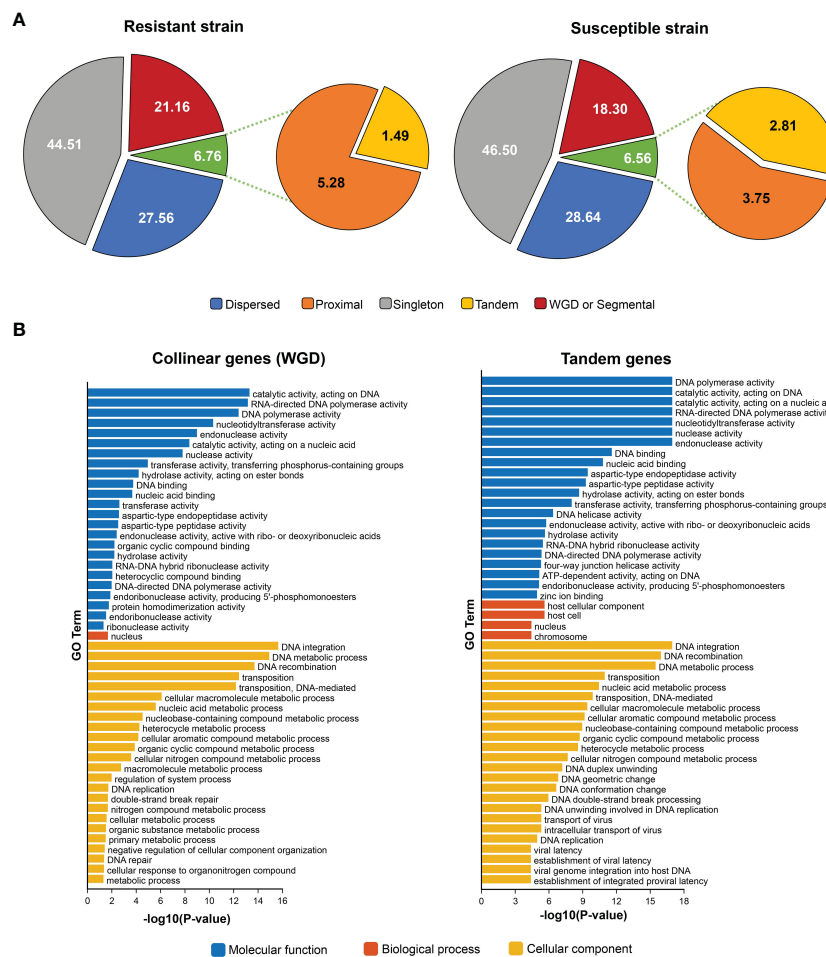

FIGURE 2
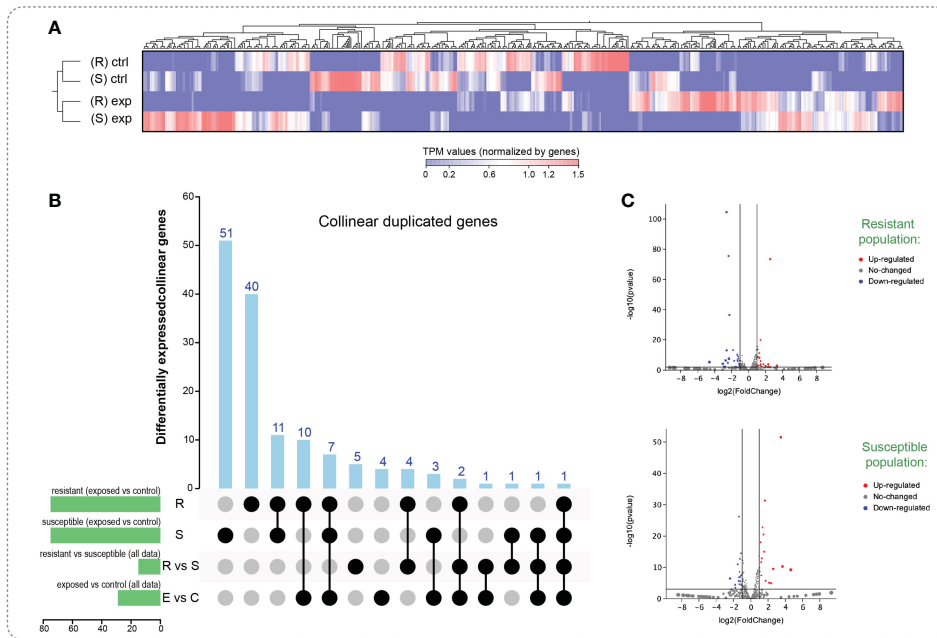Gene types and enrichment of GO annotation of duplicated genes. (A): gene type categories found in the resistant and susceptible strains after syntenic alignments evaluation by MCScanX. (B): Enrichment of GO annotations according to Molecular Function, Biological Process, and Cellular Component categories (level 3) of collinear and tandem duplicated genes in the resistant and susceptible populations.
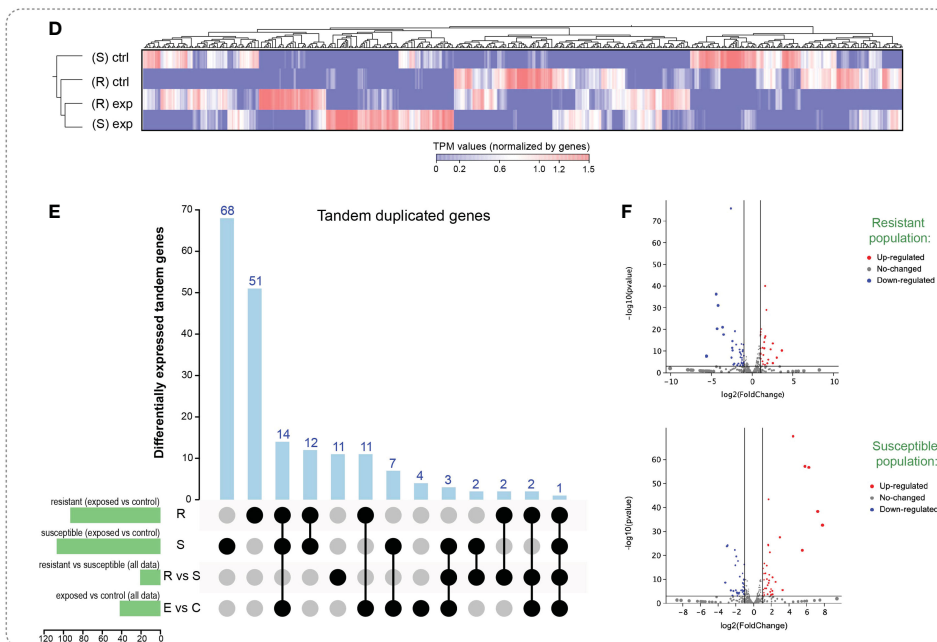
FIGURE 3

Expression analyses of collinear and tandem duplicated genes in the resistant and susceptible strains of sea lice. **(A, D)**: Hierarchical clustering of TPM values of collinear genes **(A)** and tandem genes **(D)** based on Manhattan's distances with an average linkage. Red color corresponds to high expression values and blue to low expression values. (R): resistant strain; (S): susceptible strain; ctrl: untreated groups; exp: parasites treated with azamethiphos. **(B, E)**: Upset plot counting the number of differentially expressed genes for collinear **(B)** and tandem **(E)** duplicated genes for all the statistical comparisons conducted in this study. Differentially expressed genes cutoff was |Fold change| > 4.0; P-value < 0.01. R: statistical comparison on exposed vs. control lice in the resistant strain; S: the same comparison for susceptible lice; R vs. S: statistical comparison between TPM values from the resistant strain vs. the susceptible strain; E vs. C: statistical comparison between TPM values from the exposed parasites versus control samples independently of the strain. **(C, F)**: Volcano plot of differentially expressed genes for each population for collinear **(C)** and tandem **(F)** genes.

75 were differentially expressed (|fold change| > 4; P-value < 0.01) in the resistant population when comparing the exposed and control samples; the other 75 genes were differentially expressed when conducting this comparison in the susceptible population. However, these collinear genes were not the same, with only 19

differentially expressed in exposed vs. control samples in both populations (Figure 3B). Other statistical comparisons consisted of the combination of the TPM values for exposed and control samples and then comparing the fold change between resistant and susceptible populations (R vs. S in Figures 3B, E); and combining

TPM values for both populations and then comparing then fold change between exposed and control samples (E vs. C in Figures 3B, E). The R vs. S comparison resulted in only 15 differentially expressed collinear genes, whereas the E vs. C comparison resulted in 29 differentially expressed collinear genes. Finally, 40 collinear genes were exclusively differentially expressed in the exposed vs. control comparison in the lice from the resistant strain, and 51 collinear genes were exclusively expressed in the susceptible strain (Figure 3B). Concerning tandem-duplicated genes, 93 were differentially expressed in the exposed vs. control comparison for the resistant strain and 107 were differentially expressed in the susceptible strain (Figure 3E). From these genes, 26 were differentially expressed according to this comparison in both strains. The R vs. S comparison for tandem genes resulted in 21 differentially expressed genes and the E vs. C comparison for 42 genes. Exclusive differentially expressed genes in the resistant population were 51, whereas 68 were for the susceptible lice. There was only one collinear gene and one tandem gene with differential expression values for all possible comparisons, which was an annotated gene for collinear genes (not found in public databases) and a retrotransposable element ORF2 for tandem genes.

3.4 Duplications associated with *cuticle protein* genes

One group of genes of particular interest was the *cuticle protein* genes, and another with functions on the cuticle formation was evaluated to identify duplication in lice genomes (Figure 4). Five collinear duplicated *cuticle* genes were found in both strains (white links in Figure 4A, D). Three cuticle-formation genes were collinear in the sea lice-resistant strain: *cuticle protein 4*, present in chromosome 4 and with duplication in chromosome 9; *larval cuticle protein 65*, in chromosome 5 and duplicated in chromosome 4; and *cuticlin 4*, in chromosome 10 with duplication in chromosome 5 (Figure 4A). In susceptible lice population, there were two collinear *cuticle genes*: *cuticle protein 9* and *cuticle protein 19*, both present in chromosome 11 and duplicated in chromosome 4 (Figure 4D). The full sequence length of these genes ranged from 8,030 to 32,756 bp, with a coding sequence ranging from 378 to 1,566 bp. Additionally, conserved Motif 1 was repeatedly found throughout all the *cuticle protein* genes (Figure 4B). The expression patterns of these five genes differed depending on the strain. In the resistant strain, *cuticle*
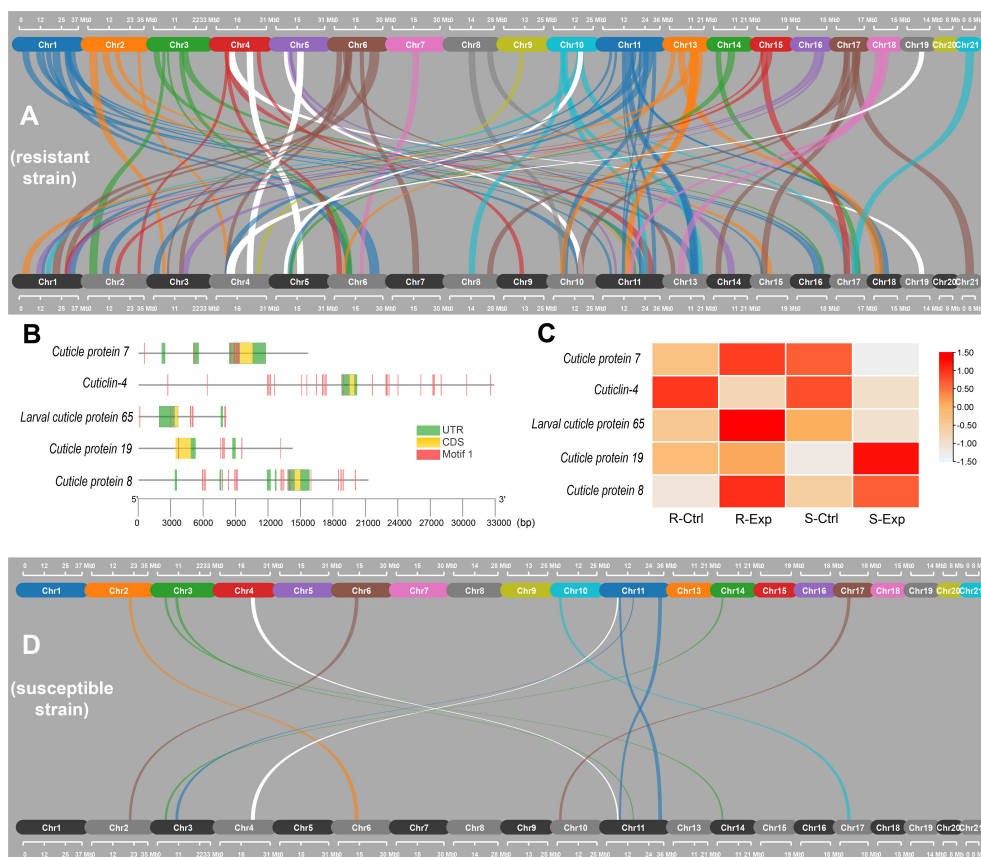
FIGURE 4
Analysis of gene duplications in *cuticle protein* genes in susceptible and resistant strains of *C. rogercresseyi*. **(A, D)** collinear duplicated genes in resistant **(A)** and susceptible **(D)** strains found in SynVisio software. White links corresponded to alignments involving *cuticle protein* genes. **(B)** Gene structure of *cuticle protein* genes duplicated in the resistant or the susceptible populations. Motif 1 was obtained through the MEME suite with default parameters. **(C)** gene expression of *cuticle protein* genes (normalized TPM values) for control samples (Ctrl) and samples exposed to azamethiphos (Exp) in both strains.

proteins 7, 8, and the *larval cuticle protein 65* were upregulated in sea lice exposed to the delousing drug azamethiphos. In contrast, only *cuticle protein 8* showed a similar pattern in the susceptible lice after exposure to the same drug (Figure 4C). The other two genes were downregulated in the susceptible strain. The *cuticlin-4* gene was downregulated in both strains after exposure to azamethiphos, while the *cuticle protein 19* showed no significant change in the resistant strain, but was upregulated in the susceptible lice (Figure 4C).

3.5 Copy Number Variants associated with resistant and susceptible strains of the sea lice

Coverage normalization analyses resulted in 10,330 genes with standard coverage for Copy Number Variants (CNVs) identification among both resistant and susceptible strains (Figures 5A, B). These 10,330 common genes in both strains were used for coverage comparisons using the susceptible strain as a reference and obtaining 134 genes with differential coverages indicating the presence of CNVs. Establishing a cutoff of | adjusted Fold Change| > 2 and P-value < 0.01, a total of 106 genes containing CNVs with duplication in the resistant population (Gain) and 28 genes containing CNVs with duplication in the susceptible lice (Loss) were found (Figure 5C). Most of the CNVs were in chromosome 6 (16 CNVs), 4 and 7 (12 CNVs in each one), which was not directly associated with the gene density on those chromosomes or the presence of collinear genes (Figure 5D). Additionally, there were 17 region CNVs, which were those variants involving wider genome regions, including more than a single-gene portion of the genome, and were distributed throughout all the chromosomes in the genome ("Gain" and "Loss" annotations in Figure 5D). The highest fold change value for a region CNV was 6.39 (P-value = 2.23 E – 308) and corresponded to a region including the gene *Rotatin*, located in chromosome 6, amplified in the resistant strain (Gain). This gene was located on the first chromosome region, in-between an area of the genome that is also collinear with other genomic regions from chromosomes 3 and 4 (Figure 5D).

3.6 Differential expression analyses of genes containing CNVs in different strains

Genes located in areas with CNVs in the genome were extracted for expression analyses using the Illumina RNA-seq data for these strains (Figure 6). Fold change values were calculated for both the strains comparing control vs. samples exposed to azamethiphos, resulting in some genes with the most significant differences for both populations, such as *Rotatin, Allatostatin-A, Gag-Pol polyprotein*, and *Ribonuclease H*, which were the genes with the most differential values (Figures 6A, C). Gene ontology analyses from the genes containing CNVs between susceptible and resistant populations showed that the nucleus was the only enriched cellular component, and various polymerase/catalytic activities were the

most enriched molecular functions. The most enriched biological process was DNA metabolic process, followed by the cellular macromolecule metabolic process and DNA integration/recombination process (Figure 6B).

4 Discussion

Loss of antiparasitic sensitivity or pharmacological resistance in *C. rogercresseyi*, has emerged as one of the primary concerns in salmon farming. The molecular mechanisms associated with the R and S phenotypes are a priority for developing sustainable aquaculture worldwide. As in many other model species, such as insects, where pharmacological resistance mechanisms have been widely studied, identifying structural gene variants on the essential proteins related to drug response is a common starting point. One of the first approaches was to infer which were the most relevant genes in the drug response by whole-transcriptome studies on sea lice exposed to the commonly applied chemotherapeutics in the salmon industry (Gallardo-Escárate et al., 2019a; Gallardo-Escárate et al., 2019b). One of the first studies where azamethiphos exposure was tested to infer genes related to potential molecular mechanisms associated with resistance/susceptibility to this drug have shown that a low dose of 3 ppb (the concentration used in the salmon industry is 100 ppb) could up-regulate the expression of genes related to drug detoxification, such as glutathione S-transferase, metalloproteinase, and ABC transporters; and genes with other functions such as trypsins and cuticle proteins (Valenzuela-Muñoz et al., 2015a). ABC transporters from the subfamilies ABCB and ABCC might be involved in azamethiphos detoxication systems, as was inferred through transcriptome studies (Valenzuela-Muñoz et al., 2015b), which are subfamilies that have also been studied in other salmon louse species *Lepeophtheirus salmonis*, but under treatment with another chemical compound (Igboeli et al., 2014). Particular attention has been paid to these subfamilies because of their described role as multidrug resistance proteins (MDRPs), with specific binding sites for various drugs (Kaur, 2002). Regarding the *cuticle proteins* genes, *C. rogercresseyi* exposed individuals to azamethiphos, and deltamethrin drugs showed an up-regulation of many cuticle synthesis-related genes in a previous study (Chávez-Mardones et al., 2016), indicating a potential drug-response role of these proteins that constitute a physical barrier for protection against external agents in arthropod species. The *trypsin* genes were also validated in further studies, which were abundant in *C. rogercresseyi*, and various isoforms were up-regulated in lice exposed to azamethiphos (Valenzuela-Miranda and Gallardo-Escárate, 2016). These genes were also differentially expressed in resistant and susceptible strains of *C. rogercresseyi* to azamethiphos and the pyrethroids deltamethrin and cypermethrin, particularly the *trypsin 2* and *trypsin 5* genes (Núñez-Acuña et al., 2020; Núñez-Acuña et al., 2021a), which have been previously validated as molecular markers associated to bioassays to infer sensitivity to azamethiphos (Sáez-Vera et al., 2021). These genes and families were, in part, the focus for finding potential gene duplications on the *C. rogercresseyi* genomes from lice obtained from strains susceptible and resistant strains to azamethiphos.
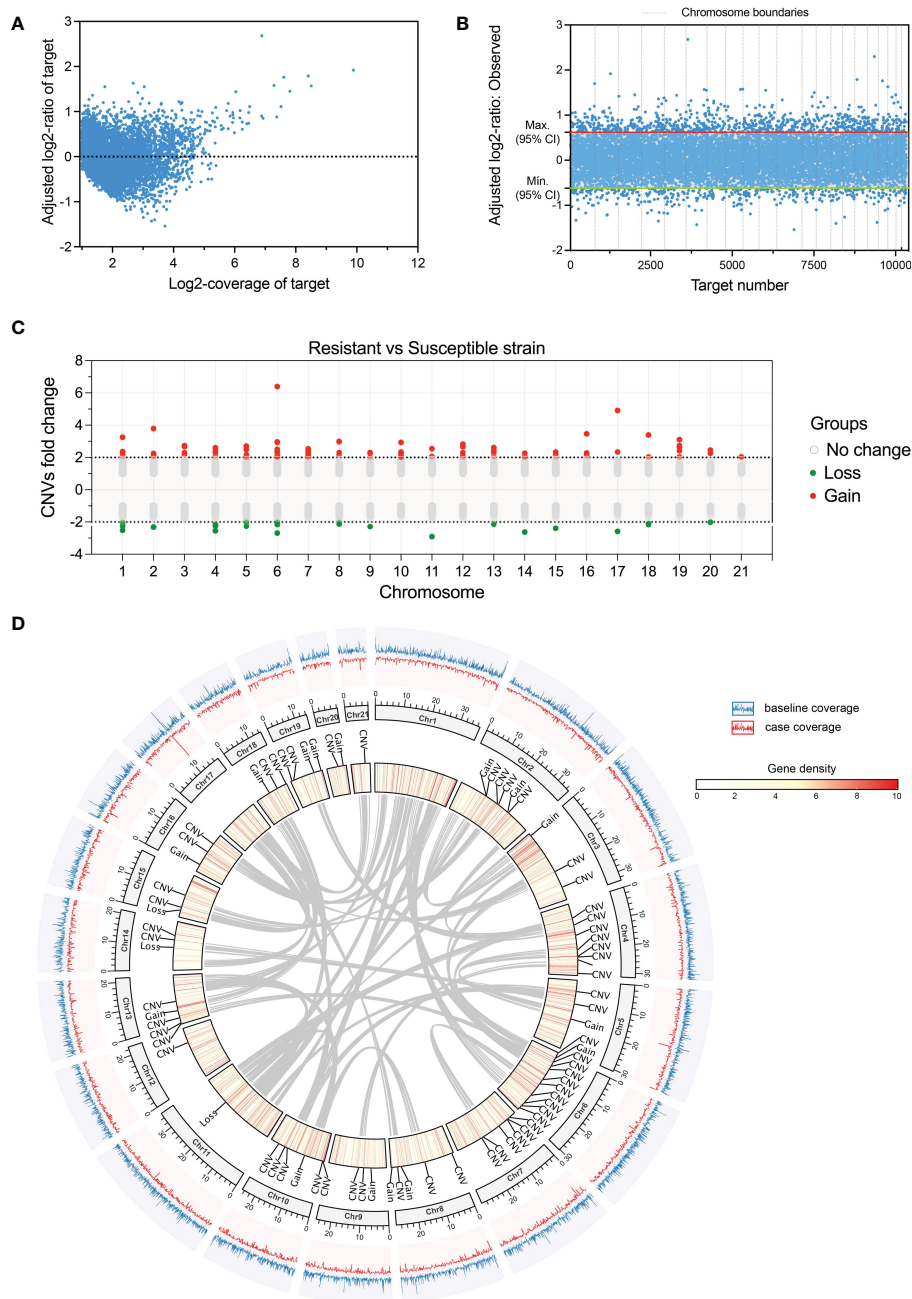
FIGURE 5
Copy Number Variations (CNVs) identified in the comparison of the resistant and the susceptible strain of sea lice. **(A)** "depth of coverage" evaluation of genes successfully mapped to the reference genome in both strains. **(B)** statistical significance of fold change obtained in CNV detection according to each gene (target number). **(C)** Fold change comparison for CNV identification in coverages from the resistant vs. the susceptible strain. **(D)** CIRCOS plot indicating the position of each CNV from the resistant strain and colocalization in the *C. rogercresseyi* genome according to gene density (track in the middle), coverage (outer track), and collinear genes (gray links).

Unexpectedly, most of the duplications evaluated in this study were not found on genes associated with the drug detoxification system (see Supplementary Material 1 for blast results). In arthropods, gene amplification and upregulation of the expression owing to duplications in drug detoxification system genes have been vastly described (Bass and Field, 2011; Ffrench-Constant, 2013). However, it is not possible to discard potential duplications of these genes in *C. rogercresseyi* just because we could not detect those in this present study. Furthermore, this study only used two

laboratory-produced strains of sea lice, one resistant and one susceptible to azamethiphos. Therefore, we suggest increasing the number of lice populations/strains for further studies to discover novel duplications in their genomes. Another critical point is that we used the first draft and the first annotation for the sea louse reference genome *C. rogercresseyi* (Gallardo-Escárate et al., 2021), which was different from the initial assembly from the *L. salmonis* salmon louse genome (Joshi et al., 2022). In parallel with this publication, we are working on a second draft of the reference
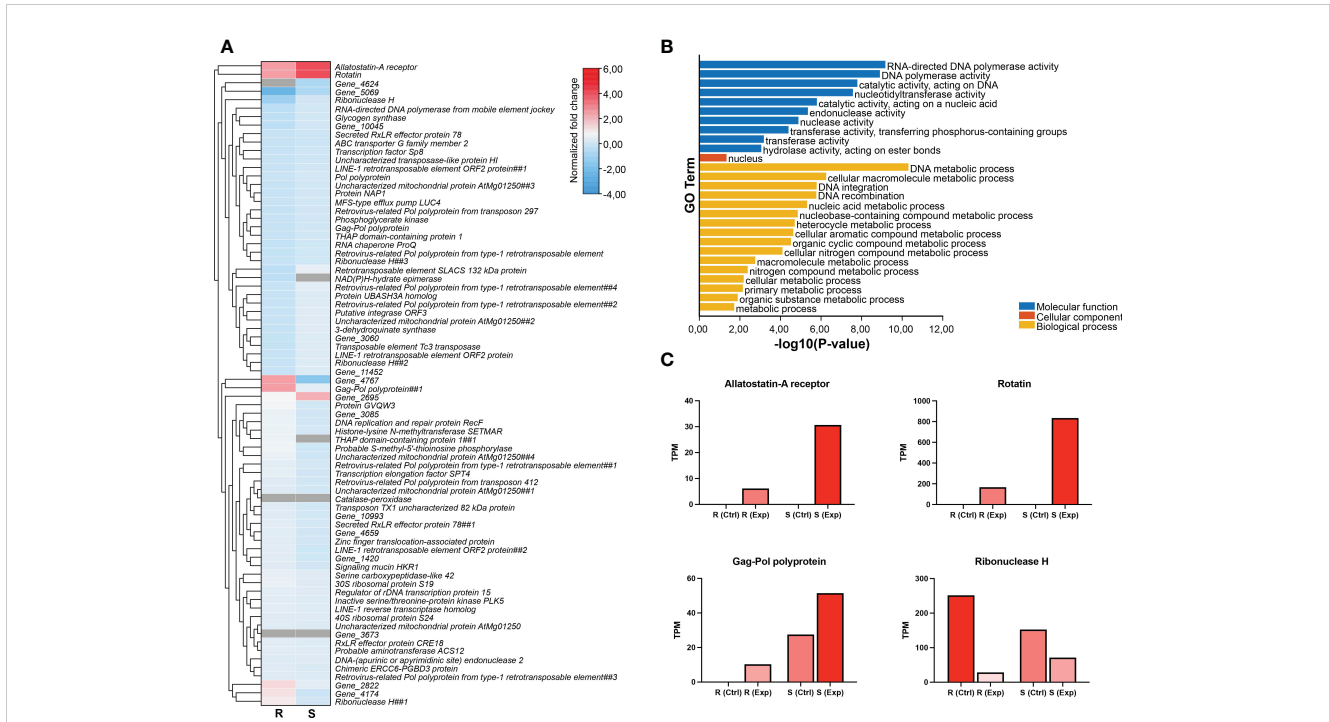
FIGURE 6

Expression analyses of genes containing CNVs in the resistant vs. susceptible strains comparison. **(A)** Hierarchical clustering of fold change values (obtained from exposed vs. control comparison for each strain) based on Manhattan's distances with an average linkage. R: the resistant strain; S: the susceptible strain. **(B)** GO Enrichment for the genes containing CNVs in the resistant vs. susceptible comparison. **(C)** selection of particular genes of interest with significant changes in TPM values.

genome and annotation to avoid the abundant "uncharacterized proteins" in public databases for *C. rogercresseyi*, which were found in many duplicated genes in this study (Supplementary Material 1). Nonetheless, we obtained valuable information due to the high coverage in the data for large genome regions. Therefore, as we could not discard the presence of undiscovered duplicated genes, we could confirm the presence of duplications in various genes related to transposable elements, cuticle proteins, and other functional genes in the obtained genomes by resequencing (Figures 2–5).

Transposable elements of the genome (TEs) are abundant in the *C. rogercresseyi* genome (Gallardo-Escárate et al., 2021) and may have been overlooked in other species in the past, but there is a growing interest in the potential regulatory functions of TEs and their consequences. Although some authors have described TEs without relevant function in the genomes (Strobel et al., 1979) or with deleterious effects on genome organization (Hickey, 1982), there is evidence of the crucial roles of TEs in the evolutionary dynamics of eukaryotic genomes (González and Petrov, 2009). There were diverse retrotransposons that were components of coding regions of functional genes or part of regulatory regions in the arthropod model species *Drosophila melanogaster* (Franchini et al., 2004). In the present study, most collinear duplicated genes that had a positive blast (discarding the uncharacterized proteins) were coding genes for different transposons, retrotransposons, or transposases. The first significant difference in this study was the unprecedented differential number of duplicated genes (obtained by collinearity) between resistant and susceptible lice to azamethiphos

(5.16% vs. 0.45% of collinear genes). The significantly higher number of duplications on the resistant strain, and the substantial difference in the percentage of genes originated by Whole-Duplication or Segmented Duplications events (WGD events were 27.47% of the total genes in resistant vs. 18.3% in the susceptible strain), suggests that mobile elements of the genome may be involved in pharmacological resistance for this species. This hypothesis is supported by the fact that we found multiple TEs duplicated genes in the resistant strain and the reported antecedents indicating that TEs may be involved in pharmacological resistance in other arthropod species (Chénais et al., 2012; Rostant et al., 2012). Indeed, some TEs have been described as the cause of duplications or deletions of coding genes with critical roles in pharmacological resistance, such as the drug detoxication *Cytochrome p450* gene (Chung et al., 2007). The extent of this study did not consider an evaluation regarding which genes were affected due to duplicated TEs. However, it sheds light on potential research lines considering these regions of the genome and the potential outcome concerning drug sensitivity in marine parasite species.

Interestingly, evidence of duplications in the comparative analyses between resistant and susceptible strains of sea lice was found in genes related to cuticle formation (Figure 4). These genes are of particular interest for this study due to their pivotal role in insecticide resistance in other arthropods (Balabanidou et al., 2018) and previous studies in *C. rogercresseyi* showing a higher number of *cuticle protein* genes in this species and an overexpression after

exposure of delousing drugs (Chávez-Mardones et al., 2016). Along with the modification of target proteins and drug detoxication systems, penetration resistance involved another mechanism referring to the change in the cuticle thickening or composition that reduced the penetration of xenobiotic drugs within the organism's body (Dang et al., 2017). Penetration pharmacological resistance has been described in various arthropods, species with many *cuticle protein* genes, which code for structural components of the cuticle. For example, in the mosquito *Anopheles gambiae*, 2% of the total coding genes belong to the family of cuticle proteins (Cornman et al., 2008). The role of these genes and cuticle structure in pharmacological resistance has been documented in several insect species treated with various drugs (Balabanidou et al., 2018). In our study, five duplicated genes associated with cuticle formation were differentially expressed in both strains in samples treated with azamethiphos against control lice. These duplicated genes were different in both strains (*cuticle protein 7*, *cuticlin-4*, and *larval cuticle protein 65* in the resistant strain, whereas *cuticle proteins 19* and *8* in the susceptible lice). Furthermore, these duplicated genes were up-regulated in the resistant strain (except for *cuticlin-4*), and those that were duplicated in the resistant strain were down-regulated in the susceptible lice. This result supports the hypothesis of a penetration-resistant mechanism in *C. rogercresseyi* that must be further investigated to help the control strategies for these parasites in salmon farms. In addition, the expression results obtained in the susceptible strain were concordant with the previous study in *C. rogercresseyi* exposed to the same antiparasitic drug (Chávez-Mardones et al., 2016).

Copy Number Variants (CNVs) analysis showed similar results to collinearity analyses, indicating that most CNVs-containing genes had enriched GO terms associated with DNA metabolic process and polymerase activity and transposon-related genes (Figure 6). In addition, resistant strain presented higher fold change values with respect to susceptible lice, similar to the results of previous collinearity analyses (Figure 5), and thus suggesting that the resistant strain might have an adaptative advantage in response to the delousing drug azamethiphos exposure through duplication in mobile elements of the genome. Other relevant genes with relevant functions were found in the differential CNVs analysis, such as the *ABC transporter G* gene, a member of a family described previously (Glavinas et al., 2004) with a fold change value of 2.14, indicating a higher number of copies in the resistant strain. However, this gene did not exhibit significant variation in the expression pattern in resistant and susceptible strains and was slightly downregulated in both strains after exposure to azamethiphos (Figure 6). Further studies will address whether this ABC transporter or other genes from this family have a potential role in pharmacological resistant due to the presence of CNVs in different populations of *C. rogercresseyi*. Another relevant gene was the *Allatostatin-A receptor* having differential expression patterns in the strains and an evident up-regulation in exposed samples vs. control (Figure 6C). This gene is the receptor for allatostatin, which is a family of neuropeptides previously described in insects and crustaceans and plays a role in development and reproduction (Stay and Tobe, 2007; Mayoral

et al., 2010). Allatostatin receptors are the homologs for the mammalian somatostatin receptor (a family of G protein-coupled receptors, GPCR), that have not yet been described in *C. rogercresseyi*. However, some other GPCRs have been described in sea lice with differential expression patterns in parasites exposed to delousing drugs (Núñez-Acuña et al., 2016), and due to some structural similarities, there is the potential to have a role in drug response. Furthermore, *allatostatin receptor* genes have been described multiple times as potential targets for insecticides in other arthropods (Bai and Palli, 2013; Audsley and Down, 2015; Liu et al., 2021). The antecedents for other arthropods and the results obtained in this present study suggest the hypothesis that the *Allatostatin-A receptor* gene and its genomic variations could have an essential role in pharmacological resistance in *C. rogercresseyi*.

In summary, this study contributes significantly to the understanding of potential pharmacological resistance mechanisms in the sea louse *C. rogercresseyi*. Through a combination of long-read DNA and short-read RNA resequencing, novel genomic markers putatively associated with drug sensitivity were found to be promising new pharmacological resistance assessment tools because they involved wider genomic regions than other punctual markers. Furthermore, the novel evidence obtained in this study suggests the presence of previously undiscovered drug resistance mechanisms associated with penetration resistance, which implies the role of duplicated *cuticle protein* genes. Finally, this study has paved the way for identifying and evaluating duplications on mobile elements of the genome (e.g., transposable elements) as a strategy for adaptation and drug response in these parasite species.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

GN-A and CG-E conceived the idea. GN-A, CS-V and VV-M performed the Nanopore and Illumina sequencing. CS-V performed the bioassays and RNA/DNA extraction. GN-A performed the bioinformatic analyses. DV-M assisted in the bioinformatic analyses. GN-A led the manuscript draft writing and figures. CS-V, VV-M, DV-M and CG-E revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2023.1112691/full#supplementary-material

References

Agusti-Ridaura, C., Dondrup, M., Horsberg, T. E., Leong, J. S., Koop, B. F., Bravo, S., et al. (2018). *Caligus rogercresseyi* acetylcholinesterase types and variants: a potential marker for organophosphate resistance. *Parasites Vectors* 11 (1), 1–16. doi: 10.1186/s13071-018-3151-7

Anderson, T., Nkhoma, S., Ecker, A., and Fidock, D. (2011). How can we identify parasite genes that underlie antimalarial drug resistance? *Pharmacogenomics* 12 (1), 59–85. doi: 10.2217/pgs.10.165

Arriagada, G., Figueroa, J., Marín, S. L., Arriagada, A. M., Lara, M., and Gallardo-Escárate, C. (2020). First report of the reduction in treatment efficacy of the organophosphate azamethiphos against the sea lice *Caligus rogercresseyi* (Boxshall & bravo, 2000). *Aquaculture Res.* 51 (1), 436–439. doi: 10.1111/are.14334

Audsley, N., and Down, R. E. (2015). G Protein coupled receptors as targets for next generation pesticides. *Insect Biochem. Mol. Biol.* 67, 27–37. doi: 10.1016/j.ibmb.2015.07.014

Bai, H., and Palli, S. R. (2013). *G Protein-coupled receptors as target sites for insecticide discovery. advanced technologies for managing insect pests* Vol. p (Berlin, Germany: Springer), 57–82.

Bailey, T. L., Johnson, J., Grant, C. E., and Noble, W. S. (2015). The MEME suite. *Nucleic Acids Res.* 43 (W1), W39–W49. doi: 10.1093/nar/gkv416

Balabanidou, V., Grigoraki, L., and Vontas, J. (2018). Insect cuticle: a critical determinant of insecticide resistance. *Curr. Opin. Insect Science.* 27, 68–74. doi: 10.1016/j.cois.2018.03.001

Bandi, V. K. (2020). *SynVisio: A multiscale tool to explore genomic conservation* (Saskatoon, Canada: University of Saskatchewan).

Bass, C., and Field, L. M. (2011). Gene amplification and insecticide resistance. *Pest Manage. Science* 67 (8), 886–890. doi: 10.1002/ps.2189

Begum, G., Albanna, A., Bankapur, A., Nassir, N., Tambi, R., Berdiev, B. K., et al. (2021). Long-read sequencing improves the detection of structural variations impacting complex non-coding elements of the genome. *Int. J. Mol. Sci.* 22 (4), 2060. doi: 10.3390/ijms22042060

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. society: Ser. B (Methodological).* 57 (1), 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bravo, S., Sevatdal, S., and Horsberg, T. E. (2008). Sensitivity assessment of *Caligus rogercresseyi* to emamectin benzoate in Chile. *Aquaculture.* 282, 7–12. doi: 10.1016/j.aquaculture.2008.06.011

Brogdon, W. G., and McAllister, J. C. (1998). Insecticide resistance and vector control. *Emerg. Infect. Dis.* 4 (4), 605–613. doi: 10.3201/eid0404.980410

Chávez-Mardones, J., Valenzuela-Muños, V., and Gallardo-Escárate, C. (2016). In silico transcriptome analysis of cuticle-related genes associated with delousing drug responses in the sea louse *caligus rogercresseyi*. *Aquaculture* 450, 123–135. doi: 10.1016/j.aquaculture.2015.07.017

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13 (8), 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chénais, B., Caruso, A., Hiard, S., and Casse, N. (2012). The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. *Gene* 509 (1), 7–15. doi: 10.1016/j.gene.2012.07.042

Chung, H., Bogwitz, M. R., McCart, C., Andrianopoulos, A., Ffrench-Constant, R. H., Batterham, P., et al. (2007). Cis-regulatory elements in the accord retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene Cyp6g1. *Genetics.* 175 (3), 1071–1077. doi: 10.1534/genetics.106.066597

Cornman, R. S., Togawa, T., Dunn, W. A., He, N., Emmons, A. C., and Willis, J. H. (2008). Annotation and analysis of a large cuticular protein family with the R&R consensus in anopheles gambiae. *BMC Genomics* 9 (1), 1–16. doi: 10.1186/1471-2164-9-22

Costello, M. J. (2009). The global economic cost of sea lice to the salmonid farming industry. *J. Fish Diseases* 32, 115–118. doi: 10.1111/j.1365-2761.2008.01011.x

Dang, K., Doggett, S. L., Veera Singham, G., and Lee, C.-Y. (2017). Insecticide resistance and resistance mechanisms in bed bugs, *Cimex* spp.(Hemiptera: Cimicidae). *Parasites Vectors* 10 (1), 1–31. doi: 10.1186/s13071-017-2232-3

Dresdner, J., Chávez, C., Quiroga, M., Jiménez, D., Artacho, P., and Tello, A. (2019). Impact of *Caligus* treatments on unit costs of heterogeneous salmon farms in Chile. *Aquacult. Econ. Manage.* 23 (1), 1–27. doi: 10.1080/13657305.2018.1449271

Ffrench-Constant, R. H. (2013). The molecular genetics of insecticide resistance. *Genetics* 194 (4), 807–815. doi: 10.1534/genetics.112.141895

Franchini, L. F., Ganko, E. W., and McDonald, J. F. (2004). Retrotransposon-gene associations are widespread among d. melanogaster populations. *Mol. Biol. Evolution.* 21 (7), 1323–1331. doi: 10.1093/molbev/msh116

Freeman, J. L., Perry, G. H., Feuk, L., Redon, R., McCarroll, S. A., Altshuler, D. M., et al. (2006). Copy number variation: new insights in genome diversity. *Genome Res.* 16 (8), 949–961. doi: 10.1101/gr.3677206

Gallardo-Escárate, C., Arriagada, G., Carrera, C., Gonçalves, A. T., Nuñez-Acuña, G., Valenzuela-Miranda, D., et al. (2019a). The race between host and sea lice in the Chilean salmon farming: a genomic approach. *Rev. Aquaculture.* 11 (2), 325–339. doi: 10.1111/raq.12334

Gallardo-Escárate, C., Valenzuela-Muñoz, V., Núñez-Acuña, G., Carrera, C., Gonçalves, A. T., Valenzuela-Miranda, D., et al. (2019b). Catching the complexity of salmon-louse interactions. *Fish Shellfish Immunol.* 90, 199–209. doi: 10.1016/j.fsi.2019.04.065

Gallardo-Escárate, C., Valenzuela-Muñoz, V., Nuñez-Acuña, G., Valenzuela-Miranda, D., Gonçalves, A. T., Escobar-Sepulveda, H., et al. (2021). Chromosome-scale genome assembly of the sea louse *Caligus rogercresseyi* by SMRT sequencing and Hi-c analysis. *Sci. Data.* 8 (1), 1–12. doi: 10.6084/m9.figshare.13186865

Glavinas, H., Krajcsi, P., Cserepes, J., and Sarkadi, B. (2004). The role of ABC transporters in drug resistance, metabolism and toxicity. *Curr. Drug Del.* 1 (1), 27–42. doi: 10.2174/1567201043480036

González, J., and Petrov, D. A. (2009). The adaptive role of transposable elements in the *Drosophila* genome. *Gene* 448 (2), 124–133. doi: 10.1016/j.gene.2009.06.008

Heckel, D. G. (2022). Perspectives on gene copy number variation and pesticide resistance. *Pest Manage. Science.* 78 (1), 12–18. doi: 10.1002/ps.6631

Hickey, D. A. (1982). Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* 101 (3-4), 519–531. doi: 10.1093/genetics/101.3-4.519

Horowitz, A. R., and Ishaaya, I. (2016). *Advances in insect control and resistance management* (Cham, Switzerland: Springer Cham).

Igboeli, O. O., Burka, J. F., and Fast, M. D. (2014). Sea Lice population and sex differences in p-glycoprotein expression and emamectin benzoate resistance on salmon farms in the bay of fundy, new Brunswick, Canada. *Pest Manage. Science* 70 (6), 905–914. doi: 10.1002/ps.3620

Joshi, J., Flores, A.-M., Christensen, K. A., Johnson, H., Siah, A., and Koop, B. F. (2022). An update of the salmon louse (*Lepeophtheirus salmonis*) reference genome assembly. *G3* 12 (6), jkac087. doi: 10.1093/g3journal/jkac087

Jugulam, M., and Gill, B. S. (2018). Molecular cytogenetics to characterize mechanisms of gene duplication in pesticide resistance. *Pest Manage. Science* 74 (1), 22–29. doi: 10.1002/ps.4665

Karunaratne, S., De Silva, W., Weeraratne, T. C., and Surendran, S. N. (2018). Insecticide resistance in mosquitoes: development, mechanisms and monitoring. *Ceylon J. Sci.* 47 (4), 299–309. doi: 10.4038/cjs.v47i4.7547

Kaur, P. (2002). Multidrug resistance: can different keys open the same lock? *Drug Resist. Updat.* 5 (2), 61–64. doi: 10.1016/S1368-7646(02)00020-1

Klegarth, A. R., and Eisenberg, D. T. (2018). "Copy number variants (CNV)," in *The international encyclopedia of biological anthropology* (Hoboken, NJ, USA: John Wiley & Sons), 1–2.

Lavrichenko, K., Johansson, S., and Jonassen, I. (2021). Comprehensive characterization of copy number variation (CNV) called from array, long-and short-read data. *BMC Genomics* 22 (1), 1–15. doi: 10.1186/s12864-021-08082-3

Li, H. (2016). Minimap and miniasm: fast mapping and *de novo* assembly for noisy long sequences. *Bioinformatics* 32 (14), 2103–2110. doi: 10.1093/bioinformatics/btw152

Li, J., Lupat, R., Amarasinghe, K. C., Thompson, E. R., Doyle, M. A., Ryland, G. L., et al. (2012). CONTRA: copy number analysis for targeted resequencing. *Bioinformatics* 28 (10), 1307–1313. doi: 10.1093/bioinformatics/bts146

Lin, Y., Jin, T., Zeng, L., and Lu, Y. (2012). Cuticular penetration of β-cypermethrin in insecticide-susceptible and resistant strains of bactrocera dorsalis. *Pesticide Biochem. Physiol.* 103 (3), 189–193. doi: 10.1016/j.pestbp.2012.05.002

Liu, N. (2015). Insecticide resistance in mosquitoes: impact, mechanisms, and research directions. *Annu. Rev. Entomol.* 60 (1), 537–559. doi: 10.1146/annurev-ento-010814-020828

Liu, N., Li, T., Wang, Y., and Liu, S. (2021). G-Protein coupled receptors (GPCRs) in insects–a potential target for new insecticide development. *Molecules* 26 (10), 2993. doi: 10.3390/molecules26102993

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi: 10.1186/s13059-014-0550-8

Mahmood, K., Højland, D. H., Asp, T., and Kristensen, M. (2016). Transcriptome analysis of an insecticide resistant housefly strain: insights about SNPs and regulatory elements in cytochrome P450 genes. *PloS One* 11 (3), e0151434. doi: 10.1371/journal.pone.0151434

Marín, S., Ibarra, R., Medina, M., and Jansen, P. (2015). Sensitivity of *Caligus rogercresseyi* (Boxshall and bravo 2000) to pyrethroids and azamethiphos measured using bioassay tests–a large scale spatial study. *Prev. Vet. Med.* 122, 33–41. doi: 10.1016/j.prevetmed.2015.09.017

Marín, S., Mancilla, J., Hausdorf, M., Bouchard, D., Tudor, M., and Kane, F. (2018). Sensitivity assessment of sea lice to chemotherapeutants: Current bioassays and best practices. *J. Fish Diseases* 41 (6), 995–1003. doi: 10.1111/jfd.12768

Mayoral, J. G., Nouzova, M., Brockhoff, A., Goodwin, M., Hernandez-Martinez, S., Richter, D., et al. (2010). Allatostatin-c receptors in mosquitoes. *Peptides* 31 (3), 442–450. doi: 10.1016/j.peptides.2009.04.013

Núñez-Acuña, G., Boltaña, S., and Gallardo-Escárate, C. (2016). Pesticides drive stochastic changes in the chemoreception and neurotransmission system of marine ectoparasites. *Int. J. Mol. Sci.* 17 (6), 700. doi: 10.3390/ijms17060700

Núñez-Acuña, G., Sáez-Vera, C., Valenzuela-Muñoz, V., Valenzuela-Miranda, D., Arriagada, G., and Gallardo-Escárate, C. (2020). Tackling the molecular drug sensitivity in the Sea louse *Caligus rogercresseyi* based on mRNA and lncRNA interactions. *Genes* 11 (8), 857. doi: 10.3390/genes11080857

Núñez-Acuña, G., Valenzuela-Muñoz, V., Carrera-Naipil, C., Sáez-Vera, C., Benavente, B. P., Valenzuela-Miranda, D., et al. (2021a). Trypsin genes are regulated through the miRNA bantam and associated with drug sensitivity in the Sea louse *caligus rogercresseyi*. *Non-Coding RNA* 7 (4), 76. doi: 10.3390/ncrna7040076

Nuñez-Acuña, G., Valenzuela-Muñoz, V., and Gallardo-Escárate, C. (2014). High-throughput SNP discovery and transcriptome expression profiles from the salmon louse *Caligus rogercresseyi* (Copepoda: Caligidae). *Comp. Biochem. Physiol. Part D: Genomics Proteomics.* 10, 9–21. doi: 10.1016/j.cbd.2014.01.003

Núñez-Acuña, G., Valenzuela-Muñoz, V., Valenzuela-Miranda, D., and Gallardo-Escárate, C. (2021b). Comprehensive transcriptome analyses in Sea louse reveal novel delousing drug responses through MicroRNA regulation. *Mar. Biotechnol.* 4 (76), 1–17. doi: 10.1007/s10126-021-10058-z

Oppong, Y. E., Phelan, J., Perdigão, J., Machado, D., Miranda, A., Portugal, I., et al. (2019). Genome-wide analysis of *Mycobacterium tuberculosis* polymorphisms reveals lineage-specific associations with drug resistance. *BMC Genomics* 20 (1), 1–15. doi: 10.1186/s12864-019-5615-3

Råberg, L., Graham, A. L., and Read, A. F. (2009). Decomposing health: tolerance and resistance to parasites in animals. *Philos. Trans. R. Soc. London B: Biol. Sci.* 364 (1513), 37–49. doi: 10.1098/rstb.2008.0184

Revie, N. M., Iyer, K. R., Robbins, N., and Cowen, L. E. (2018). Antifungal drug resistance: evolution, mechanisms and impact. *Curr. Opin. Microbiol.* 45, 70–76. doi: 10.1016/j.mib.2018.02.005

Rostant, W. G., Wedell, N., and Hosken, D. J. (2012). Transposable elements and insecticide resistance. *Adv. Genet.* 78, 169–201. doi: 10.1016/B978-0-12-394394-1.00002-X

Sáez-Vera, C., Núñez-Acuña, G., and Gallardo-Escárate, C. (2021). Sensitivity assessment to azamethiphos by time-to-response bioassay and biomarkers in the sea louse *caligus rogercresseyi*. *Aquaculture* 546 (737340), 1–10. doi: 10.1016/j.aquaculture.2021.737340

Schröder, A., Klein, K., Winter, S., Schwab, M., Bonin, M., Zell, A., et al. (2013). Genomics of ADME gene expression: mapping expression quantitative trait loci relevant for absorption, distribution, metabolism and excretion of drugs in human liver. *Pharmacogenomics J.* 13 (1), 12. doi: 10.1038/tpj.2011.44

SEARCH C (2006). *Sea Lice resistance to chemotherapeutants: A handbook in resistance management* (Sacramento, CA: SEARCH).

Srivastava, M., and Misra-Bhattacharya, S. (2015). Overcoming drug resistance for macro parasites. *Future Microbiol.* 10 (11), 1783–1789. doi: 10.2217/fmb.15.73

Stafford, K. A., and Coles, G. C. C. (2017). "Drug resistance in ectoparasites of medical and veterinary importance," in *Antimicrobial drug resistance: Mechanisms of drug resistance*, vol. 1 . Eds. D. L. Mayers, J. D. Sobel, M. Ouellette, K. S. Kaye and D. Marchaim (Cham: Springer International Publishing), 735–744.

Stay, B., and Tobe, S. S. (2007). The role of allatostatins in juvenile hormone synthesis in insects and crustaceans. *Annu. Rev. Entomol.* 52, 277–299. doi: 10.1146/annurev.ento.51.110104.151050

Strobel, E., Dunsmuir, P., and Rubin, G. M. (1979). Polymorphisms in the chromosomal locations of elements of the 412, copia and 297 dispersed repeated gene families in drosophila. *Cell* 17 (2), 429–439. doi: 10.1016/0092-8674(79)90169-7

Valenzuela-Miranda, D., and Gallardo-Escárate, C. (2016). *Caligus rogercresseyi* serine proteases: Transcriptomic analysis in response to delousing drugs treatments. *Aquaculture* 465, 65–77. doi: 10.1016/j.aquaculture.2016.08.027

Valenzuela-Muñoz, V., Chavez-Mardones, J., and Gallardo-Escárate, C. (2015a). RNA-Seq analysis evidences multiple gene responses in *Caligus rogercresseyi* exposed to the anti-salmon lice drug azamethiphos. *Aquaculture* 446, 156–166. doi: 10.1016/j.aquaculture.2015.05.011

Valenzuela-Muñoz, V., Sturm, A., and Gallardo-Escárate, C. (2015b). Transcriptomic insights on the ABC transporter gene family in the salmon louse *caligus rogercresseyi*. *Parasites Vectors* 8 (1), 1–14. doi: 10.1186/s13071-015-0801-x

Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017). Fast and accurate *de novo* genome assembly from long uncorrected reads. *Genome Res.* 27 (5), 737–746. doi: 10.1101/gr.214270.116

Wan, F., Yin, C., Tang, R., Chen, M., Wu, Q., Huang, C., et al. (2019). A chromosome-level genome assembly of *Cydia pomonella* provides insights into chemical ecology and insecticide resistance. *Nat. Commun.* 10 (1), 1–14. doi: 10.1038/s41467-019-12175-9

Wang, Y., Tang, H., DeBarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40 (7), e49–e4e. doi: 10.1093/nar/gkr1293

Xiong, C., Fang, F., Chen, L., Yang, Q., He, J., Zhou, D., et al. (2014). Trypsin-catalyzed deltamethrin degradation. *PloS One* 9 (3), e89517. doi: 10.1371/journal.pone.0089517