



Research

Cite this article: Bhattarai UR, Doherty J-F, Poulin R, Dowle E, Gemmell NJ. 2026 Revealing the genetic mechanisms underpinning the parasite-induced water-seeking behaviour of insects through RNA-seq. *Proc. R. Soc. B* **293**: 20261414.

<https://doi.org/10.1098/rspb.2026.1414>

Received: 1 February 2026

Accepted: 9 July 2026

Subject Category:

Genetics and genomics

Subject Areas:

behaviour, ecology, molecular biology

Keywords:

host manipulation, water-seeking behaviour, dual RNA-seq, *Forficula auricularia*, *Mermis nigrescens*

Author for correspondence:

Neil J. Gemmell

e-mail: neil.gemmell@otago.ac.nz

Supplementary material is available online at
<https://doi.org/10.6084/m9.figshare.c.8614245>.

Revealing the genetic mechanisms underpinning the parasite-induced water-seeking behaviour of insects through RNA-seq

Upendra R. Bhattarai¹, Jean-François Doherty^{2,3}, Robert Poulin⁴, Eddy Dowle⁶ and Neil J. Gemmell⁵

¹Center for Translational Neuroscience, Brown University, Providence, RI 02912, USA

²Department of Zoology and ³Michael Smith Laboratories, Department of Biochemistry & Molecular Biology, The University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada

⁴Department of Zoology and ⁵Department of Anatomy, University of Otago, Dunedin 9016, New Zealand

⁶New Zealand Institute for Bioeconomy Science Limited, Lincoln 7608, New Zealand

J-FD, 0000-0003-4766-9417; RP, 0000-0003-1390-1206; NJG, 0000-0003-0671-3637

Water-seeking behaviour in arthropods infected by mermithid nematodes or nematomorphs is a classic example of host manipulation, yet its molecular basis remains poorly understood. We investigated mechanisms underlying this aberrant behaviour in European earwigs (*Forficula auricularia*) infected with the nematode *Mermis nigrescens* using comparative RNA-seq across stages of infection and behavioural manipulation in both host and parasite. We detected 12 876 and 9722 expressed genes in earwigs and nematodes, respectively. Differential expression analyses identified 673 upregulated and 593 downregulated genes in earwigs, and 2672 upregulated and 2293 downregulated genes in nematodes across comparisons of infection stage and parasite size. Temporal clustering revealed six earwig and four nematode gene clusters. Gene ontology enrichment indicated increased signalling and sensory pathways in the host, while the parasite upregulated transport and secretory processes during manipulation, suggesting complementary and coordinated transcriptional changes between parasite and host. Notably, a shared GATA transcription factor motif was enriched among overexpressed genes in both organisms during manipulation. Together, these results identify candidate pathways and regulatory elements in host and parasite, providing new insights into the molecular regulation of parasite-induced water-seeking behaviour.

1. Introduction

Animals continuously make behavioural choices impacting their fitness and reproductive success. Contrary to the assumption that behaviours are self-determined and naturally selected for the animal's own fitness gain [1,2], instances exist where this is not the case [3–5]. Parasites often manipulate host behaviour to confer advantages upon themselves [6,7], a strategy that has independently evolved across all major parasite phylogenetic lineages [8]. Such manipulated hosts essentially serve as extended phenotypes of the parasites [9–11].

The water-seeking behaviour in arthropod hosts infected by parasitic worms (nematodes and nematomorphs) illustrates this phenomenon [12,13]. The

evolutionary significance of this aberrant behaviour is particularly striking because of the apparent convergent evolution of this behaviour in parasites from two different phyla [14,15]. Nematomorphs and mermithid nematodes must emerge from their terrestrial arthropod hosts into water or water-saturated substrates to avoid desiccation, reproduce and thus complete their life cycle [16].

Despite historical documentation dating back over 140 years [17], the underlying mechanisms of parasite-induced water-seeking remain underexplored. Early proteomic investigations highlighted significant parasite–host molecular crosstalk during behavioural modification, suggesting possible biochemical pathways involved in host manipulation [18]. Research by Thomas and colleagues [19] identified neurotransmitter and neuromodulator alterations in the brain of the cricket *Nemobius sylvestris* infected by the nematomorph *Paragordius tricuspidatus*. The authors [20,21] carried out proteomic studies in grasshoppers and crickets infected with the nematomorphs *Spinochondodes tellinii* and *P. tricuspidatus*, respectively, correlating proteins from the Wnt signalling pathway with behavioural change. Moreover, although dehydration superficially increases interactions with water, protein profiling does not support this as a driving mechanism by hairworms [22].

Even less is known about the mechanisms underpinning the similar behavioural manipulation induced by mermithid nematodes. Williams *et al.* [23] found that the mermithid nematode *Thaumamermis zealandica* causes its amphipod host *Talorchestia quoyana* to burrow deeper into sand, down to water-saturated layers, following an increase in haemolymph osmolality after infection, suggesting some thirst-driven mechanism facilitating nematode emergence, yet this observation lacks a genetic explanation.

Transcriptomics offers insights into gene expression and physiological mechanisms dictating behaviour [24]. While high-throughput sequencing has advanced host–parasite interaction studies [25], the lack of comprehensive genomic data limits its application in non-model organisms. Dual RNA-seq, which profiles gene expression of both host and parasite, has been a popular approach [26–28]; however, this strategy is often complicated by RNA quality disparities between the host and the parasite when extracted together [29]. To address these issues, we adjusted our protocols to process host and parasite RNA separately, facilitating accurate dissection and individual sequencing. We also used a developmental series of both the parasite and the corresponding host. The use of a developmental time series across infections is crucial to differentiate between gene expression changes associated with infection and those associated with host manipulation. This research uses the European earwig (*Forficula auricularia*) and its parasitic mermithid nematode (*Mermis nigrescens*) to explore gene expression across different infection stages, aiming to elucidate the genetic mechanisms driving the observed water-seeking behaviour. Behavioural changes in earwigs occur primarily when mermithids have reached maturity [30]. As many cases of host manipulation in insects involve the central nervous system (CNS) [31], we hypothesize that substantial changes in the transcriptome should occur there. When their development within the earwig host is complete, these nematodes induce their host to seek water and enter it [30], allowing the worm to emerge where it needs to be to reproduce. Although some proteomic changes in earwig brains following infection by mermithid nematodes have been identified [14], these only provide a partial window into the underlying mechanisms responsible for host manipulation. Our comprehensive genomic approach focusing on both the host and parasite across developmental stages aims to address this knowledge gap.

2. Material and methods

(a) Insect rearing

Adult earwigs (*F. auricularia*) were collected from the Dunedin Botanic Garden (−45°51′27.59″ S, 170°31′15.56″ E) and reared in a temperature-controlled room (temperature: cycling from 15 to 12°C, day/night; photoperiod of L : D 16 : 8) in the Department of Zoology, University of Otago, Dunedin, New Zealand. Earwigs were placed in a 20 l glass container with a dirt bed, dahlia flower petals and pieces of cardboard cuttings as cover. They were provided with food containing a 50 : 50 mixture of ground oats and commercial cat food *ad libitum*. Water was sprayed every second day to maintain moisture in the container. Earwigs were left for at least a week to acclimatize before sampling.

(b) Sampling

Naturally infected and uninfected (UI) earwigs were sampled across four time points (UI, early infection (EI), late infection (LI) and during manipulation (DM)) and nematodes across three time points (early maturation, late maturation and DM) (figure 1). Out of all individuals maintained as described above, all earwigs the behaviour of which was manipulated, that is those that entered water, were included in our analysis along with their nematode; for other stages, we used the first earwigs and nematodes we dissected that fell into these categories until we had matching numbers (see figure 1). Adult earwigs were snap-frozen in liquid N₂ and dissected in 1× phosphate-buffered saline (PBS) under a dissection microscope to confirm the state and stage of parasitism. Samples were dissected quickly on dry ice and stored at −80°C until further use. When present, worms were located in the earwig's body cavity. After being removed from their insect host, they were thoroughly cleaned by repeated rinsing with PBS solution to remove contaminants of host origin prior to being put in frozen storage. Categorization of the samples was done by measuring the largest nematode found in each earwig. UI earwigs have no detectable nematode infection. EI earwigs and early maturation nematodes correspond to an earwig infected by a nematode smaller than 2 mm. LI earwigs and late maturation nematodes correspond to an earwig infected by a nematode between 7 and 9 mm in length. For DM, samples were obtained by observing earwigs inside their rearing container with a water pool (6 cm diameter and 0.5–1 cm depth) in darkness for 5–6 h

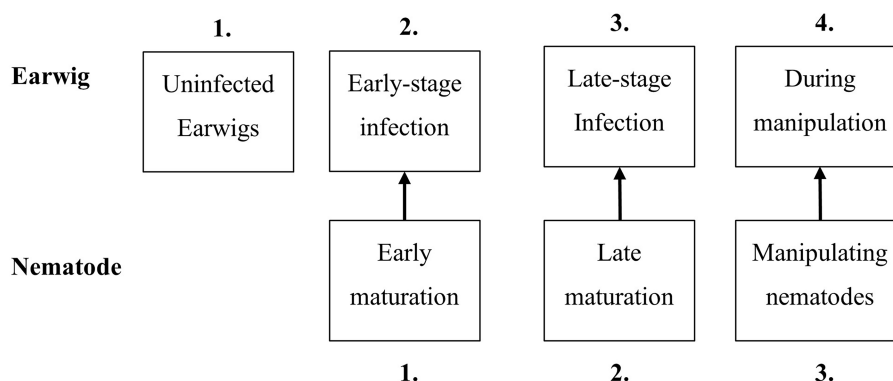


Figure 1. Sampling stages for RNA-seq studies of water-seeking behaviour. Earwigs were sampled at four time points: uninfected earwigs (UI), early-stage infection (EI), late-stage infection (LI) and during manipulation (DM), whereas nematodes were sampled at three time points: small nematodes (SN) extracted from earwigs at early stage of infection, large nematodes (LN) extracted from earwigs at late stage of infection, and manipulating nematodes (MN) extracted from earwigs during manipulation.

using a red light. Once a nematode was seen coming out of an earwig in water, both specimens were individually snap-frozen in liquid N₂. All samples ($N = 3$ or 4 per stage, see figure 1) were collected between January and July 2020, except for one sample in the EI which was collected the following year from the same field population following the same sampling procedure. One LI sample was infected with multiple nematodes, all between 7 and 9 mm long, and only the largest nematode from this sample was processed for sequencing. Earwigs were sexed based on their morphological characteristics; however, nematodes were not, owing to technical difficulties.

(c) Nucleotide extraction

Total RNA extraction was carried out using the Direct-zol RNA MicroPrep Kit (Zymo Research, USA) with an on-filter DNase treatment following the manufacturer's instructions. Samples were ground in 300 μ l 1 $\times$ PBS buffer with blue sterile tips in a 2 ml Eppendorf tube kept half dipped inside a mortar filled with liquid N₂ to prepare for extraction. For earwigs, extractions were performed from their head (after removing the black eye pigments on the outer eye) to get the whole CNS, while for nematodes extractions were performed from their whole body. Quality and quantity of the extracted nucleotides were measured on a Qubit 2.0 Fluorometer (Life Technologies, USA) and NanoDrop. High-quality RNA samples were stored at -80°C until use.

(d) RNA sequencing

RNA integrity was evaluated on a Fragment Analyzer (Advanced Analytical Technologies, USA), and total RNA sequencing library was prepared with the Zymo-Seq RiboFree Total RNA Library Kit (Zymo Research, USA). Prepared libraries were quality checked on a MiSeq and sequenced across two lanes of Illumina HiSeq 2500 V2 Rapid Sequencing with 100 bp paired-end (PE) reads at the Otago Genomics Facility, University of Otago, Dunedin, New Zealand.

(e) Bioinformatics pipeline for RNA-seq data

The RNA-seq reads were quality checked with FastQC and trimmed with Trimmomatic v. 0.38 [32] with the following parameters: ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:20 MINLEN:36. Reads were then mapped to their respective draft genomes [33,34] using two-pass mapping with the option `-sjdbOverhang` set to 99 in STAR v. 2.7.9 [35]. RSEM v. 1.3.3 [36] was used to generate read counts per gene with the `rsem-calculate-expression` command.

(f) Differential expression analysis and clustering

To understand how gene expression changed in both the earwig host and nematode parasite during infection, a differential expression analysis was carried out in R (v. 4.3.3) using the DESeq2 package [37]. The sample variation for expression profiles was assessed using principal component analysis (PCA) with regularized log-transformed data. For both earwig and nematode datasets, genes were retained if their expression count was >10 in at least three replicates. The DESeq2 analysis was modelled to regress out the batch effects and account for both the sex and time points (design: $\sim$ batch + sex + time points) for earwig and time points as main effect for nematode dataset (design: $\sim$ batch + time points). The pairwise differential expression was calculated between time points in both datasets using the Wald test. Genes were considered differentially expressed when the corrected p -value using the Benjamini–Hochberg method was below 0.05 and log₂ fold change (FC) was more than or equal to 2. A non-duplicated list of all differentially expressed genes from each pairwise comparison was used for clustering analysis of

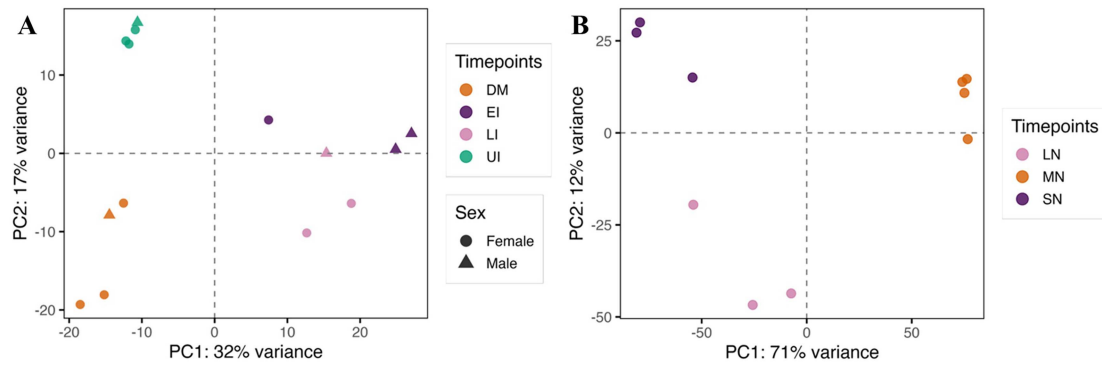


Figure 2. Principal component analysis (PCA) of the normalized RNA-seq data. (A) Earwig samples showing groupings based on infection status, with PC1 explaining 32% and PC2 17% of the variation. (B) Nematode samples coloured by their progressive developmental stages, as determined by size and change induced in behaviour, show grouping based on the time points, with PC1 explaining 71% and PC2 12% of the variance. Time points: for earwigs, UI = uninfected, EI = early-stage infection, LI = late-stage infection and DM = during manipulation; for nematodes, SN = early maturation, LN = late maturation and MN = manipulating worm.

expression patterns across time points using the DIANA algorithm in the DEGreport package (v. 1.38.5) [38] in R with a minimal cluster size of 20 genes.

(g) Functional analysis

An over-representation analysis (ORA) was performed with gene ontology (GO) terms for the genes in each cluster to identify common functional roles. GO annotation was extracted from the genome annotation [33,34], and an org.db database was created for each species using the makeOrgPackage function from the AnnotationForge (v. 1.44.0) package. ORA was performed with enrichGO from the clusterProfiler package (v. 4.0) [39] with q -value cutoff of 0.05 and multiple correction with Benjamini–Hochberg with all other default parameters unless otherwise stated. For genes in clusters 3, 5 and 6, default parameters did not yield enriched terms, so we used more lenient thresholds with multiple corrections set to none and q -value cutoff of 1. The universal gene set included all the genes in the differential gene expression (DGE) analysis.

(h) Motif analysis

Motif enrichment analyses were used to identify common transcription factor binding sites that may be involved in host manipulation. Analyses focused on the cluster of genes upregulated DM compared to UI, EI and LI stages. Motif enrichment analysis was carried out using the findMotifs.pl tool from the HOMER package (v. 5.1) [40] in FASTA mode. For the genes of interest, we took 500 bp on both sides of the transcription start site as foreground, and the same genomic window from all the expressed genes was used as the background.

3. Results

(a) Differential gene expression analysis

A total of 213.1 million 100 bp PE reads were generated from earwig samples and 155.7 million 100 bp PE reads for nematode samples. PCA showed that the samples clustered primarily by infection or development stage for both earwigs and nematodes along PC1 and PC2, while no clustering was observed based on earwig sex (figure 2A,B). In earwigs, 32% of sample variation was explained by PC1 and 17% by PC2, whereas in nematodes, 71% of variation was explained by PC1 and 12% by PC2. Nematode samples, particularly those from the manipulation stages (MN), exhibited clear separation from other stages, while earwig samples formed three separate clusters: UI, EI and LI stages and DM.

A total of 12876 and 9722 genes were detected in earwig and nematode datasets, respectively. Among these, 673 genes were significantly upregulated and 593 genes downregulated in earwigs, and 2672 genes were upregulated and 2293 genes downregulated in nematodes across all pairwise comparisons (Wald tests, false discovery rate (FDR) < 0.05 and $|\log_2 \text{FC}| \geq 2$). Heat maps revealed clear clustering of samples by developmental stages in both species (figure 3A,B).

Pairwise differential expression analysis in earwigs showed 34 upregulated and 93 downregulated genes between UI and EI, whereas only 1 gene was upregulated and 5 downregulated between EI and LI. Comparisons between LI and DM showed 187 upregulated and 46 downregulated genes. Similarly, 41 genes were up- and 75 were downregulated between UI and LI, 162 up- and 55 downregulated between UI and DM, and 256 up- and 75 downregulated between EI and DM.

In the nematode dataset, 181 genes were upregulated and 1111 genes downregulated between small nematode (SN) and large nematode (LN), 609 up- and 805 downregulated between LN and manipulating nematode (MN), and 761 up- and 1730 downregulated between SN and MN. Volcano plots illustrate the differences in magnitude and number of differentially expressed genes

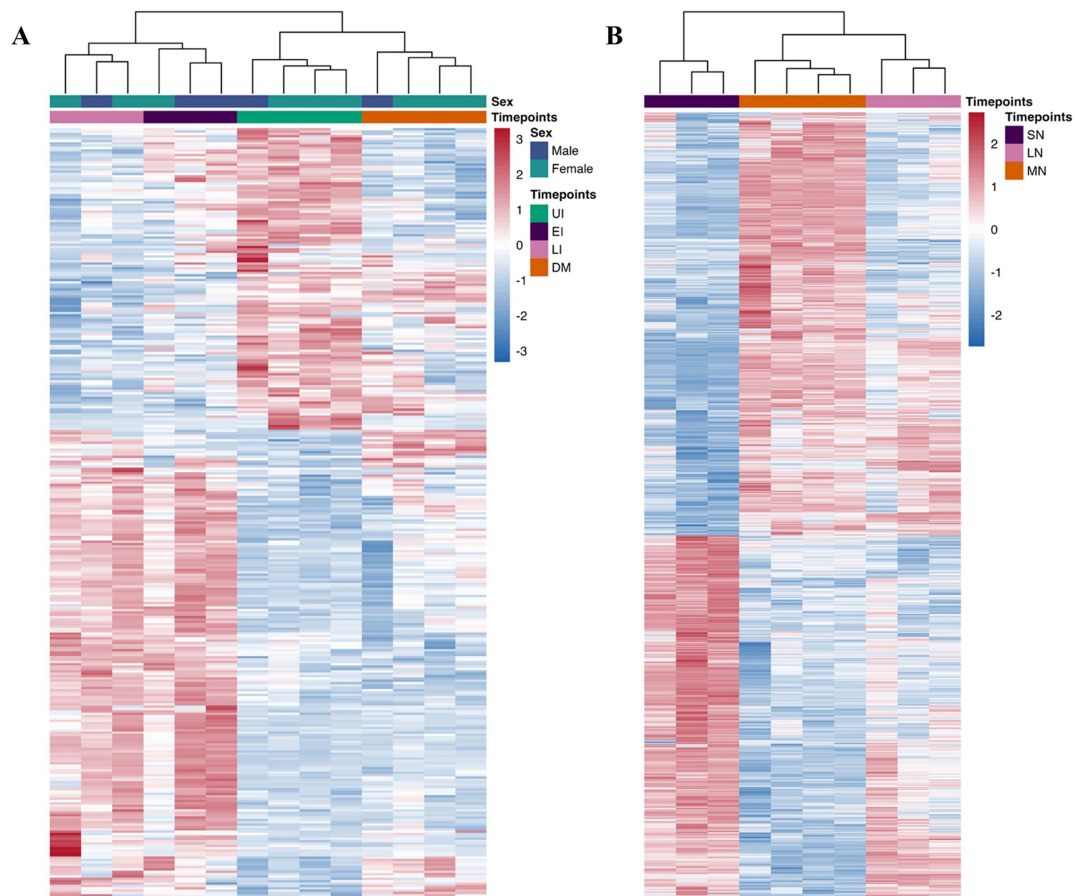


Figure 3. Heat map showing normalized, regularized logarithm (rlog)-transformed read counts of top 500 differentially expressed genes ($p_{\text{adj.}} < 0.05$) for earwig samples (A) and nematode samples (B). Each row of the heat map represents a gene differentially expressed across all samples, colour-coded according to Wald statistic test scores, where red indicates upregulation and blue downregulation of a gene expression (gradient colour bar on the right). Each column represents an individual. Colours above the heat map indicate the sample categories as indicated by legends on the right. Time points: for earwigs, UI = uninfected, EI = early-stage infection, LI = late-stage infection and DM = during manipulation; for nematodes, SN = early maturation, LN = late maturation and MN = manipulating worm.

for each comparison (supplementary material, figure S1). In both species, the changes of gene expression heightened during the manipulation stage.

(b) Clustering of differentially expressed genes based on expression patterns

The heat maps of the differentially expressed genes (DEGs) revealed distinct gene groupings across the various time points in both the earwig and nematode datasets (figure 3). To identify genes likely associated with host manipulation rather than those associated with infection or general nematode development, DEGs were clustered based on their expression profiles across time points (figure 4A,B).

In the earwig dataset, six gene clusters with distinct patterns were identified. Clusters 1, 2 and 4 are characterized by reduced expression during the manipulation stage compared with earlier time points. In contrast, cluster 5 includes genes showing a steady increase in expression from UI earwigs through all stages of infection (EI and LI) and DM. Expression of the genes in cluster 3 reduced after infection to LI and recovered slightly at DM, whereas the genes in cluster 6 peaked at DM after a slight decrease from UI to EI and LI. Cluster sizes range from 25 to 153 genes, with cluster 1 containing 153 genes, cluster 2 with 77, cluster 3 with 25, cluster 4 with 151, cluster 5 with 55 and cluster 6 with 37 genes. Of particular interest are cluster 2, which shows a pronounced drop in expression at manipulation stage, and cluster 6 which shows decreased expression following infection but rises DM. These clusters may provide insights into the gene expression changes driving behavioural manipulation.

In the nematode dataset, DEG clustering yielded four groups with distinct temporal dynamics. Cluster 1 consists of genes that are consistently upregulated from the early maturation (SN) to the late maturation (LN) and manipulative (MN) stages, whereas cluster 2 shows the opposite pattern, with genes that are consistently downregulated across these stages. Clusters 3 and 4 are notable in that their genes exhibit decreased expression from SN to LN but then show a significant increase from LN to the manipulation stage (MN). The gene counts in these clusters are 1725 for cluster 1, 907 for cluster 2, 64 for cluster 3 and 209 for cluster 4. Since our primary focus is on gene changes occurring during the manipulation stage, clusters 3 and 4 with their

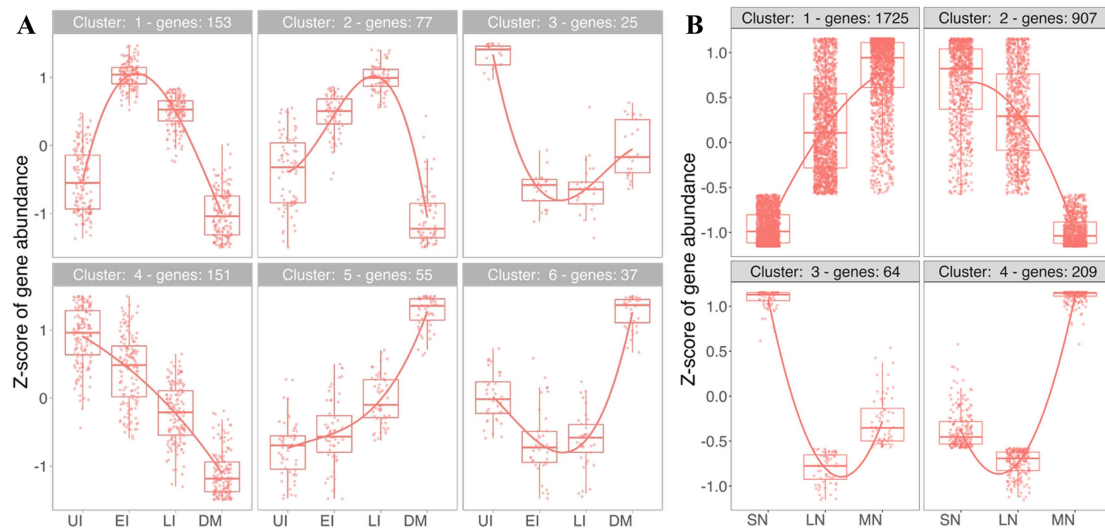


Figure 4. Cluster analysis of significantly expressed genes for earwig samples (A) and nematode samples (B). The y-axis shows the Z-score of gene abundance; the values are centred to the mean and scaled to the standard deviation for each gene. The x-axis shows different sampling stages. Each dot represents the expression of a gene; their relative expression values are summarized by box plots and whiskers at each sampling point. The curved lines represent the expressed trends. The number of genes in each cluster is given at the top of each cluster. Time points: for earwigs, UI = uninfected, EI = early-stage infection, LI = late-stage infection and DM = during manipulation; for nematodes, SN = early maturation, LN = late maturation and MN = manipulating worm.

marked upregulation at MN compared with the previous infection stage represent particularly promising groups for functional analysis.

(c) Functional analysis of significantly expressed genes in the earwig host

An ORA of within-cluster gene GO terms was used to identify key functions associated with each of the clusters. Earwig cluster 1 had significant enrichment in three GO terms. The most prominent term was extracellular region ($p_{\text{adj.}} 1.48\text{e-}08$), which included genes *cda1*, *obst-e*, *chit1*, *cht10* and *ForAur_00000656*. These same genes were also enriched for the chitin binding category ($p_{\text{adj.}} 2.82\text{e-}07$). Additionally, the carbohydrate metabolic process term was enriched ($p_{\text{adj.}} 2.86\text{e-}03$), with genes including *chit1*, *cht10*, *cda8*, *cda1*, *cht2* and *gusb*.

For cluster 2, extracellular region ($p_{\text{adj.}} 0.004$) and chitin binding ($p_{\text{adj.}} 0.012$), incorporating *cda1*, *chit1* and *obst-e* genes, were again enriched. Other terms enriched in this cluster include iron ion binding, haem binding and tetrapyrrole binding. There was notable overlap of genes across these terms, particularly *cyp15a1*, *cyp9e2*, *cyp4c1*, *agmo* and *obst-e* (figure 5B). Cluster 4 presented 15 enriched GO terms, with many enriched terms overlapping with cluster 2 (figure 5D).

Clusters exhibiting increased expression during the manipulation stage compared with EI and LI, namely clusters 3 and 5, initially did not yield significantly enriched terms under the q -value cutoff of <0.05 and Benjamini–Hochberg multiple correction. Using a more lenient threshold ($p\text{AdjustMethod} = \text{'none'}$, $q\text{valueCutoff} = 1$), cluster 3 returned eight enriched GO terms, four of which belonged to the biological process category. Each featured a gene ratio of 0.2, with *vg2* associated with lipid transport and localization and *hmgcs-1* involved in phospholipid biosynthetic and metabolic processes. The remaining four enriched terms were from the molecular function category, involving genes such as *gld*, *beta*, *vg2*, *nrf-6*, *oacyl* and *hmgcs-1* (figure 5C). Cluster 5 also showed increased expression throughout infection and manipulation stages. Enriched GO terms included extracellular region, transition metal ion binding, oxidoreductase activity, response to external stimuli, response to stress, and others (figure 5E).

The most intriguing findings emerged from cluster 6, where gene expression was high during the manipulation stage after being downregulated during EI and LI (compared with UI stage). Most of the enriched terms in this cluster are related to signalling, particularly transmembrane signalling. These included signal transduction, cell communication, G protein-coupled receptor (GPCR) activity and pathway, transmembrane signalling receptor activity and molecular transducer activity. Notably, just three genes, *hcrtr2*, *adra1a* and *pde10a*, were involved across all these enriched terms, and each exhibited higher expression during DM than at other time points (figure 6).

(d) Motif analysis of earwig genes putatively involved in host manipulation

Motif analyses were used within clusters with gene expression changes DM to identify common gene regulatory regions. Enrichment of common motifs across candidate manipulation genes could suggest that the regulatory control of behavioural manipulations is linked to specific transcription factors. Motif analysis of genes in cluster 6 revealed nine known motifs (supplementary material, table S1). Several of these motifs, including NF κ B-p65, NF κ B-p50, Fra2 and TF3A, are known regulators of immune

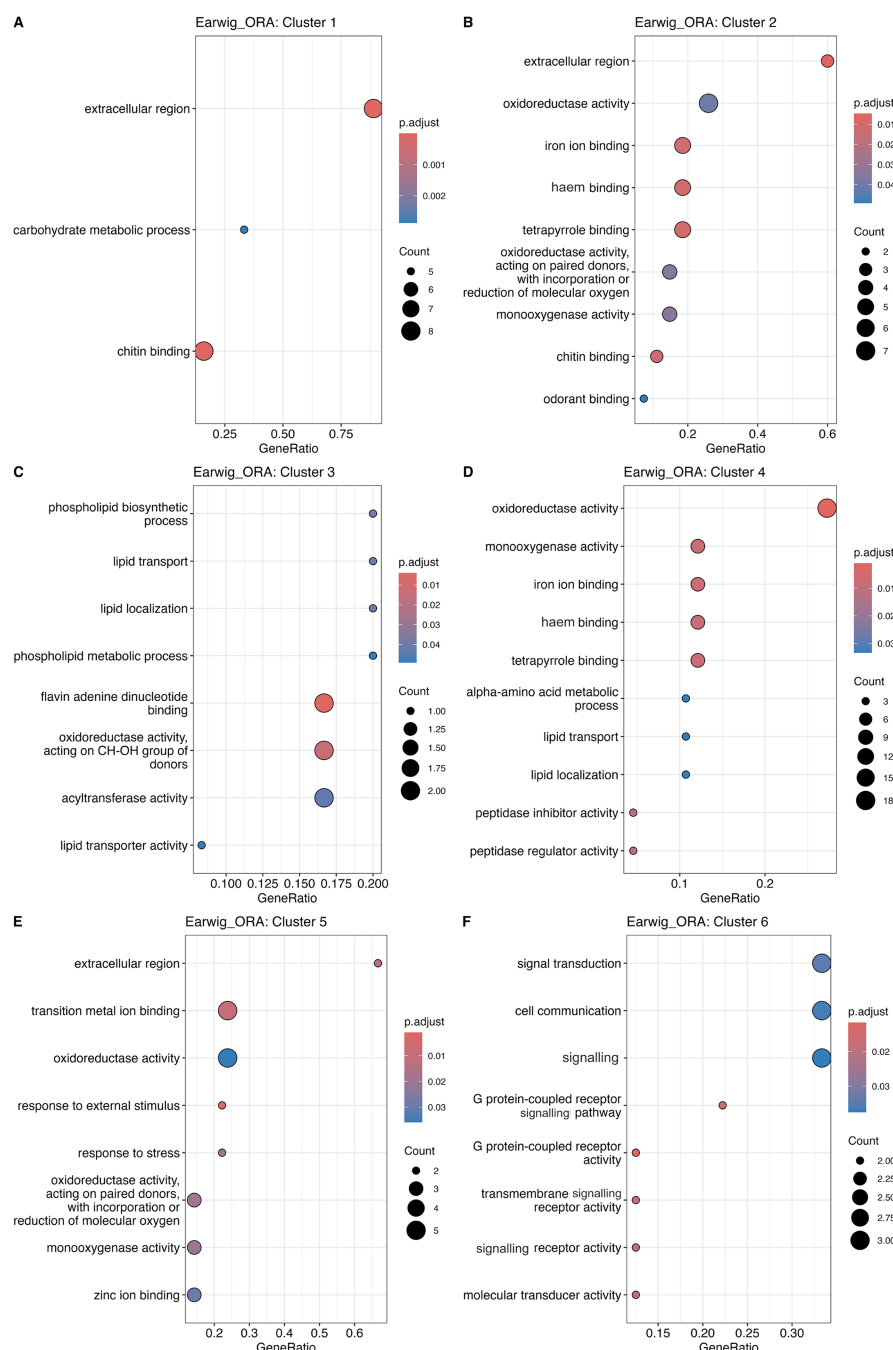


Figure 5. Dot plots showing top 10 enriched GO terms from the over-representation analysis of genes in earwig clusters 1 (A), 2 (B), 3 (C), 4 (D), 5 (E) and 6 (F) (see figure 4 for an explanation of clusters). Adjusted p -values are shown by the colour gradient, red to blue for low to high; the size of the dot indicates the gene count of each term. The x -axis represents gene ratio (scale varies between plots), and enriched terms are listed on the y -axis.

responses and inflammation [41–44]. Among them, NF κ B-p65 was the most enriched motif associated with 11 genes. ORA of GO terms for this gene set indicated enrichment for lipid metabolic process ($p_{\text{adj.}}$ 0.03) and lipase activity ($p_{\text{adj.}}$ 0.04), each represented by a single gene (*lip3* and *plbd2*, respectively) (figure 7A).

The GATA transcription factor ELT-3, a known regulator of ageing in worms [45], is surprisingly one of the enriched motifs in cluster 6 genes in the earwig. ORA using a lenient threshold has earlier identified two enriched GO terms: hydrolase activity, acting on carbon–nitrogen bonds ($p_{\text{adj.}}$ 0.03), and lipase activity (0.04), each represented by one gene (*pgrp* and *plbd2*, respectively) (figure 7B).

(e) Functional analysis of significantly expressed genes in the nematode

Clusters 1 and 2 in the nematode dataset showed distinct expression trends across time, with cluster 1 genes exhibiting progressively increased expression and cluster 2 genes showing a decrease from SN to MN. ORA identified 49 significantly enriched GO terms for cluster 1 and 23 for cluster 2. For cluster 1, top terms highlighted robust involvement in signalling and cellular

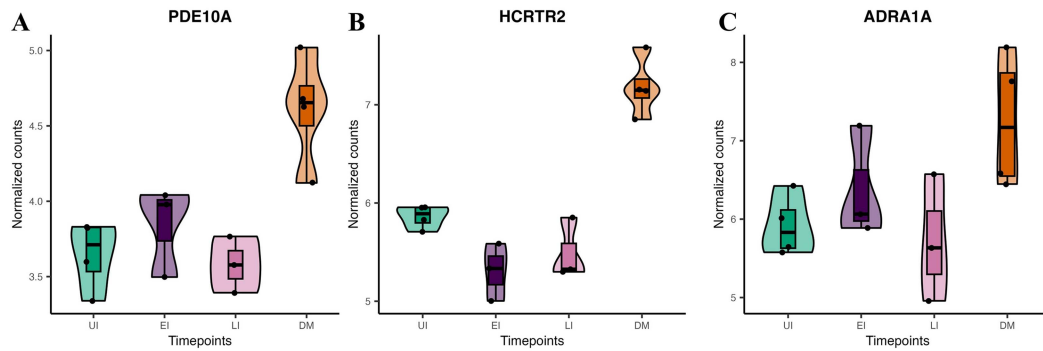


Figure 6. Expression of genes from enriched pathways in cluster 6 of the earwig dataset (see figure 4 for explanation of cluster 6). (A) *pde10a*, (B) *hcrtr2* and (C) *adra1a* gene expression. The y-axis represents log₂-normalized gene counts, and the x-axis represents the infection stages of the earwig. UI = uninfected, EI = early infection stage, LI = late infection stage and DM = during manipulation.

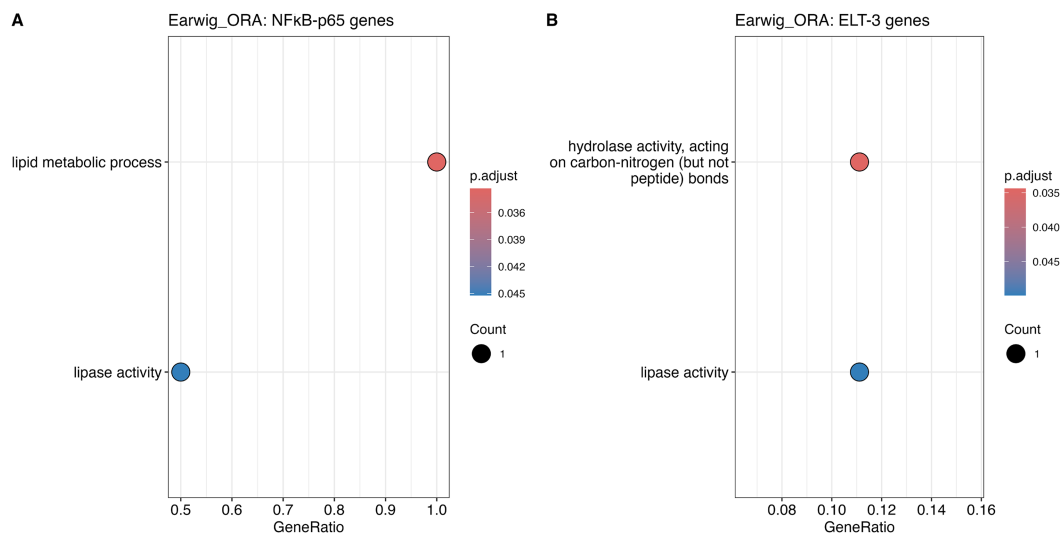


Figure 7. Dot plots showing enriched GO terms from the over-representation analysis of genes potentially linked to manipulation from two of the over-represented motifs. (A) Genes with NFκB-p65 motif enrichment and (B) genes with ELT-3 motif enrichment. Adjusted p -values are shown by the colour gradient, red to blue for low to high; the size of the dot indicates the gene count of each term. The x-axis represents gene ratio, and enriched terms are listed on the y-axis.

communication, including pathways related to GPCR signalling and transmembrane transport ($p_{\text{adj.}} 9.8\text{e-}06$), cell communication ($p_{\text{adj.}} 1.52\text{e-}05$), signalling ($p_{\text{adj.}} 1.6\text{e-}05$) and signal transduction ($p_{\text{adj.}} 4.33\text{e-}05$). Additional terms pointed to heightened cellular responses to external stimuli ($p_{\text{adj.}} 2.41\text{e-}03$). Together, these results suggest a pronounced increase in transmembrane signalling activity in this group.

The most prominent features of cluster 2 were terms associated with biosynthetic and metabolic processes. The most significant terms included translation ($p_{\text{adj.}} 4.69\text{e-}19$), and the following terms were peptide and amide biosynthetic and metabolic processes. This pattern indicates widespread downregulation of biosynthetic and metabolic activity in cluster 2 (figure 8).

Cluster 3, characterized by a modest increase in expression at the manipulation stage (MN) following a decrease from SN to LN, yielded a single significantly over-represented term, cell surface ($p_{\text{adj.}} 0.01$), including the genes *ttr-52* and *ttr-2*. In cluster 4, where expression strongly increased at MN after a slight decline from SN to LN, the standard threshold did not yield significant enrichment. However, by relaxing the threshold as before, 25 enriched terms were identified. The most notable among these were transmembrane transport ($p_{\text{adj.}} 0.005$), regulation of DNA-templated transcription and RNA biosynthetic processes ($p_{\text{adj.}} 0.014$), and various metabolic and gene regulatory processes.

Within the transmembrane transport category, the majority of enriched genes encoded solute carrier proteins, such as *slc15a2*, *slc2a3*, *slc17a5*, *slc18a2* and *slc39a10*, as well as other transport-related proteins like *abcc3*, *oatp74d*, *pept-3*, *trpc2*, *kha1*, *orct* and *cac*. Many top enriched terms also consistently featured a set of regulatory genes, including *lrrfip2*, *meis3p2*, *ehf*, *foxp1*, *foxo2*, *st18*, *sgf1* and *eip74ef* (figure 8).

(f) Motif analysis of nematode genes putatively involved in host manipulation

Motif analysis of nematode cluster 4 genes identified six enriched known motifs: ABF1, GATA (Zf) IR3, AGL95 (ND), NFE2L2 (bZIP), Bach1 (bZIP) and PHA-4 (forkhead) (supplementary material, table S2). Among these, GATA-IR3, a zinc finger

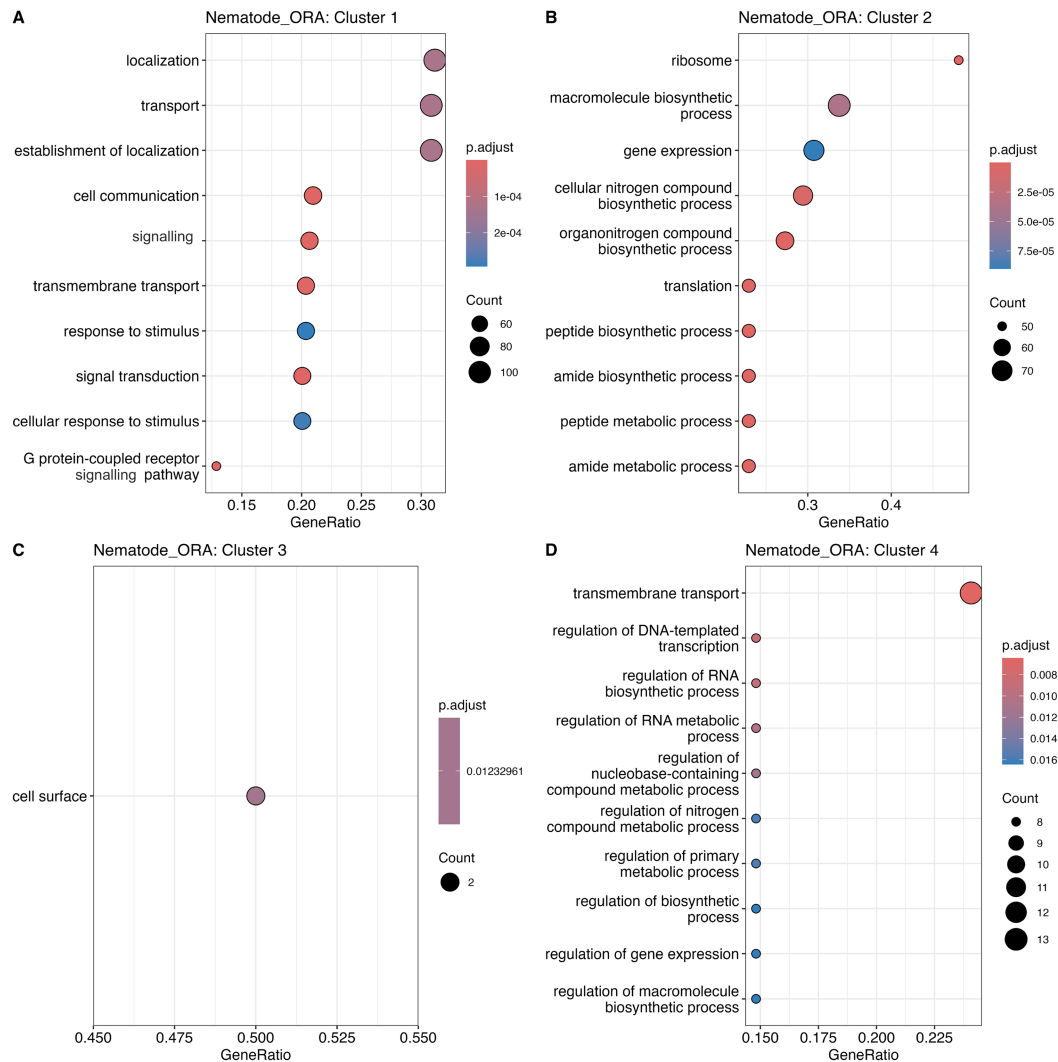


Figure 8. Dot plots showing the top 10 enriched GO terms from the over-representation analysis of genes in nematode clusters 1 (A), 2 (B), 3 (C) and 4 (D). Adjusted *p*-values are shown by the colour gradient, red to blue for low to high; the size of the dot indicates the gene count of each term. The x-axis represents gene ratio, and enriched terms are listed on the y-axis.

DNA-binding motif, though evolutionary distant, shares fundamental structural similarities with ELT-3, the GATA zinc finger motif enriched in earwig cluster 6 genes. ORA using this motif revealed seven enriched GO terms at a lenient threshold as earlier (figure 9), each represented by a single gene. These included: UDP-glycosyltransferase activity, extracellular ligand-gated monoatomic ion channel activity and oligosaccharide metabolic process-containing gene *tre1*; ligand-gated monoatomic ion channel activity, ligand-gated channel activity and hydrolase activity acting on glycosyl bond-containing gene *exp-1*; and hydrolase activity, hydrolysing O-glycosyl-compound-containing gene *ugt3*.

4. Discussion

We hypothesized that the nematode manipulates earwig behaviour by interfering with the host's CNS, thereby altering information processing and decision-making by the host [31]. In earwigs, EI versus LI produced relatively few DEGs. This limited change may reflect the nematode's growth inside the abdomen having only subtle effects on the CNS until the parasite reaches the manipulation stage. In contrast, nematode samples represented distinct developmental stages, which likely explains the larger number of DEGs observed between stages. This pattern supports the notion that nematodes remain 'under the radar' until they are fully developed and ready to emerge, an adaptive strategy critical for completing their life cycle.

Focusing on the genes differentially expressed DM, functional analysis revealed enriched signalling processes in both species. In earwigs, genes involved in transmembrane signalling were upregulated in the CNS, while nematodes showed increased activity of genes linked to carrier functions, molecular transport and regulatory processes. These complementary changes suggest a coordinated interaction in which parasite-driven transport and regulatory mechanisms may modulate host signalling, ultimately producing the behavioural shifts observed.

Keeping in mind that we used more lenient thresholds for enrichment and thus that the findings should be interpreted with caution, several host genes nevertheless emerged as particularly noteworthy. In cluster 6 of the earwig data, *hcrtr2*, *adra1a* and

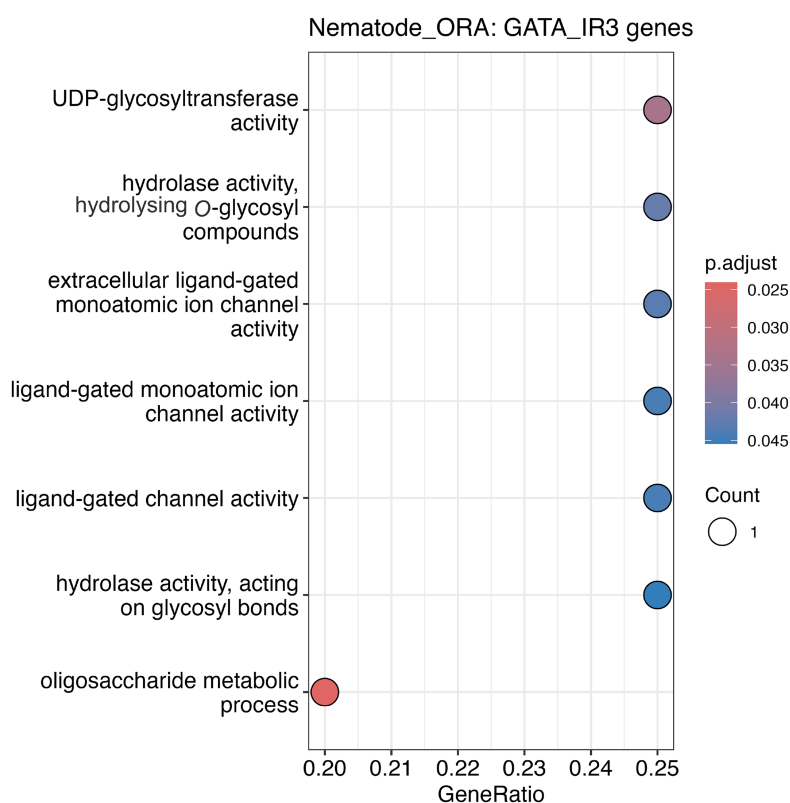


Figure 9. Dot plots showing enriched GO terms from the over-representation analysis of genes with GATA-IR3 motif enrichment. Adjusted p -values are shown by the colour gradient, red to blue for low to high; the size of the dot indicates the gene count of each term. The x -axis represents gene ratio, and enriched terms are listed on the y -axis.

pde10a were strongly associated with enriched GO terms. The gene *pde10a* encodes phosphodiesterase 10A, a cyclic nucleotide phosphodiesterase involved in signal transduction and the dopamine receptor-associated second messenger pathway [46,47]. Its role in cyclic guanosine monophosphate (cGMP) hydrolysis is well established in both vertebrates and invertebrates [47,48], and loss of *pde10a* expression has been linked to reduced locomotion and cognitive defects in mice [49,50]. Notably, earwigs exhibit increased walking during the manipulative stage (personal observation), which may increase their chances of encountering water. This behavioural correlation highlights the *pde10a* gene as an interesting candidate gene for future investigation. The other genes, *adra1a*, encoding the adrenoceptor alpha 1A protein, and *hcrtr2*, encoding hypocretin receptor 2 proteins, are involved in GPCRs, a family of transmembrane proteins involved in diverse physiological and regulatory functions, including immune responses and sensory perception in insects [51,52]. Their expression further underscores the role of CNS signalling pathways in host manipulation.

On the parasite side, cluster 4 nematode genes were enriched for transmembrane transport and transcriptional and translational regulation. These processes are consistent with the secretion of molecules that could influence earwig signalling. Motif analysis revealed enrichment of conserved regulatory elements, including the zinc finger GATA motif in both species and NF- κ B-p65 as the most enriched in earwig. The lipid metabolism-related pathways were enriched in the earwig genes with those motifs. Lipid metabolism is tightly linked to neural function and cognitive processes [53]. One of the genes enriching lipase activities, *lip3*, is a regulator of metabolic responses to starvation and ageing [54]. The host's candidate genes with ELT-3 motifs enriching the functional GO terms are *pgrp* and *plbd2*. The *pgrp* gene encodes peptidoglycan recognition proteins (PGRPs); these are highly conserved pattern recognition proteins and are known to activate immune and neurobehavioural pathways [55,56].

In nematodes, GATA-IR3 motif-enriched genes include *tre1*, which is a GPCR, essential for germ cell migration in *Drosophila* [57]; similarly, *exp-1* is important for nutrient uptake in parasitic worms [58]. The *ugt3* (UDP-glucuronosyltransferase) gene is involved in detoxification and longevity [59] and may also contribute to the nematode's ability to survive within the host and influence host signalling.

Our findings are derived from naturally infected earwigs. We used worm size as a proxy for developmental stage, with no information on how 'old' the infection was. An obvious next step would be to conduct a similar study with experimentally infected hosts, in order to control the timing and dose of infection and reduce inherent variability among individual hosts used in the study. Nevertheless, taken together, our findings reveal that both earwigs and nematodes undergo coordinated transcriptional changes during the manipulation stage. It is possible that some of these changes serve a different function, that is, preparing the nematode for the physiological adjustments required for its free-living adult stage in a water-saturated substrate. However, many of the transcriptional changes we observed seem aligned with a role in host manipulation. Host genes linked to CNS signalling and parasite genes involved in transport, secretion and regulation converge on shared functional pathways. The enrichment of conserved *cis*-regulatory motifs across both species suggests common molecular switches that may mediate cross-species

communication. This interplay between host signalling and parasite regulation provides a molecular framework for understanding the remarkable behavioural manipulation observed in this system and highlights candidate genes for future functional validation.

Ethics. No permits were required for specimen collection in this study.

Data accessibility. Raw sequence reads for both the host and the parasites are deposited in the SRA repository under the BioProject PRJNA1364760.

Supplementary material is available online [60].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. U.R.B.: data curation, formal analysis, investigation, validation, visualization, writing—original draft, writing—review and editing; J.F.D.: methodology, writing—review and editing; R.P.: conceptualization, funding acquisition, writing—review and editing; E.D.: conceptualization, data curation, formal analysis, investigation, writing—original draft, writing—review and editing; N.J.G.: conceptualization, funding acquisition, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed herein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This study was funded by a Royal Society Te Aparangi Marsden Fund grant (UOO1613, to N.J.G.) and a Departmental PhD scholarship from the Department of Anatomy, University of Otago (to U.R.B.).

Acknowledgements. We thank Joanne Gillum and Sara Ferreira for their technical assistance and support staff at the Dunedin Botanic Garden for their help in collecting earwigs. We acknowledge the use of New Zealand eScience Infrastructure (NeSI) high-performance computing facilities.

References

- McNamara JM, Trimmer PC, Houston AI. 2014 Natural selection can favour 'irrational' behaviour. *Biol. Lett.* **10**, 20130935. (doi:10.1098/rsbl.2013.0935)
- Travis J, Reznick DN. 2018 Natural selection: how selection on behavior interacts with selection on morphology. *Curr. Biol.* **28**, R882–R884. (doi:10.1016/j.cub.2018.07.006)
- Hafer N, Manfred M. 2016 Inter- and intraspecific conflicts between parasites over host manipulation. *Proc. R. Soc. B.* **283**, 20152870. (doi:10.1098/rspb.2015.2870)
- Hughes DP, Libersat F. 2019 Parasite manipulation of host behavior. *Curr. Biol.* **29**, R45–R47. (doi:10.1016/j.cub.2018.12.001)
- Poulin R. 2010 Parasite manipulation of host behavior. In *Advances in the study of behavior*, 1st edn. Amsterdam, The Netherlands: Elsevier. (doi:10.1016/S0065-3454(10)41005-0)
- Heil M. 2016 Host manipulation by parasites: cases, patterns, and remaining doubts. *Front. Ecol. Evol.* **4**, 80. (doi:10.3389/fevo.2016.00080)
- Poulin R, Maure F. 2015 Host manipulation by parasites: a look back before moving forward. *Trends Parasitol.* **31**, 563–570. (doi:10.1016/j.pt.2015.07.002)
- Brown SP. 1999 Cooperation and conflict in host-manipulating parasites. *Proc. R. Soc. Lond. B* **266**, 1899–1904. (doi:10.1098/rspb.1999.0864)
- Bailey NW. 2012 Evolutionary models of extended phenotypes. *Trends Ecol. Evol.* **27**, 561–569. (doi:10.1016/j.tree.2012.05.011)
- Dawkins R. 1982 *The extended phenotype: the gene as the unit of selection*. Oxford, UK: Oxford University Press. (doi:10.2307/1309359)
- Hughes DP, Kronauer DJC, Boomsma JJ. 2008 Extended phenotype: nematodes turn ants into bird-dispersed fruits. *Curr. Biol.* **18**, R294–R295. (doi:10.1016/j.cub.2008.02.001)
- Ponton F, Otálora-Luna F, Lefèvre T, Guerin PM, Lebarbenchon C, Duneau D, Biron DG, Thomas F. 2011 Water-seeking behavior in worm-infected crickets and reversibility of parasitic manipulation. *Behav. Ecol.* **22**, 392–400. (doi:10.1093/beheco/arq215)
- Thomas F, Schmidt-Rhaesa A, Martin G, Manu C, Durand P, Renaud F. 2002 Do hairworms (Nematomorpha) manipulate the water seeking behaviour of their terrestrial hosts? *J. Evol. Biol.* **15**, 356–361. (doi:10.1046/j.1420-9101.2002.00410.x)
- Herbison REH, Evans S, Doherty JF, Algje M, Kleffmann T, Poulin R. 2019 A molecular war: convergent and ontogenetic evidence for adaptive host manipulation in related parasites infecting divergent hosts. *Proc. R. Soc. B* **286**, 20191827. (doi:10.1098/rspb.2019.1827)
- Obayashi N, Iwatani Y, Sakura M, Tamotsu S, Chiu MC, Sato T. 2021 Enhanced polarotaxis can explain water-entry behaviour of mantids infected with nematomorph parasites. *Curr. Biol.* **31**, R777–R778. (doi:10.1016/j.cub.2021.05.001)
- Baker GL, Capinera JL. 1997 Nematodes and nematomorphs as control agents of grasshoppers and locusts. *Mem. Entomol. Soc. Can.* **129**, 157–211. (doi:10.4039/entm129171157-1)
- McCook HC. 1884 Note on the intelligence of a cricket parasitised by a *Gordius*. *Proc. Acad. Nat. Sci. Phila.* **36**, 293–294.
- Biron DG, Loxdale HD. 2013 Host–parasite molecular cross-talk during the manipulative process of a host by its parasite. *J. Exp. Biol.* **216**, 148–160. (doi:10.1242/jeb.073825)
- Thomas F, Ulitsky P, Augier R, Dusticier N, Samuel D, Strambi C, Biron DG, Cayre M. 2003 Biochemical and histological changes in the brain of the cricket *Nemobius sylvestris* infected by the manipulative parasite *Paragordius tricuspidatus* (Nematomorpha). *Int. J. Parasitol.* **33**, 435–443. (doi:10.1016/S0020-7519(03)00014-6)
- Biron DG, Marché L, Ponton F, Loxdale HD, Galéotti N, Renault L, Joly C, Thomas F. 2005 Behavioural manipulation in a grasshopper harbouring hairworm: a proteomics approach. *Proc. R. Soc. B* **272**, 2117–2126. (doi:10.1098/rspb.2005.3213)
- Biron DG *et al.* 2006 Suicide of crickets harbouring hairworms: a proteomics investigation. *Insect Mol. Biol.* **15**, 731–742. (doi:10.1111/j.1365-2583.2006.00671.x)
- Coates LM *et al.* 2025 Dehydration does not drive host behavioural manipulation by hairworms. *PLoS One* **20**, e0332641. (doi:10.1371/journal.pone.0332641)
- Williams CM, Poulin R, Sinclair BJ. 2004 Increased haemolymph osmolality suggests a new route for behavioural manipulation of *Talorchestia quoyana* (Amphipoda: Talitridae) by its mermithid parasite. *Funct. Ecol.* **18**, 685–691. (doi:10.1111/j.0269-8463.2004.00910.x)
- Han Y, Gao S, Muegge K, Zhang W, Zhou B. 2015 Advanced applications of RNA sequencing and challenges. *Bioinf. Biol. Insights* **9**, 29–46. (doi:10.4137/BBI.S28991)
- Goodwin S, McPherson JD, McCombie WR. 2016 Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17**, 333–351. (doi:10.1038/nrg.2016.49)
- Choi YJ, Aliota MT, Mayhew GF, Erickson SM, Christensen BM. 2014 Dual RNA-seq of parasite and host reveals gene expression dynamics during filarial worm–mosquito interactions. *PLoS Negl. Trop. Dis.* **8**, e2905. (doi:10.1371/journal.pntd.0002905)
- Ehret T, Spork S, Dieterich C, Lucius R, Heitlinger E. 2017 Dual RNA-seq reveals no plastic transcriptional response of the coccidian parasite *Eimeria falciformis* to host immune defenses. *BMC Genomics* **18**, 686. (doi:10.1186/s12864-017-4095-6)
- Mukherjee P, Burgio G, Heitlinger E. 2021 Dual RNA sequencing meta-analysis in *Plasmodium* infection identifies host–parasite interactions. *mSystems* **6**, e00182–21. (doi:10.1128/mSystems.00182-21)

29. O’Keeffe KR, Jones CD. 2019 Challenges and solutions for analysing dual RNA-seq data for non-model host–pathogen systems. *Methods Ecol. Evol.* **10**, 401–414. (doi:10.1111/2041-210X.13135)
30. Herbison REH, Evans S, Doherty J-F, Poulin R. 2019 Let’s go swimming: mermithid-infected earwigs exhibit positive hydrotaxis. *Parasitology* **146**, 1631–1635. (doi:10.1017/S0031182019001045)
31. Hughes DP, Libersat F. 2018 Neuroparasitology of parasite–insect associations. *Annu. Rev. Entomol.* **63**, 471–487. (doi:10.1146/annurev-ento-020117-043234)
32. Bolger AM, Lohse M, Usadel B. 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. (doi:10.1093/bioinformatics/btu170)
33. Bhattarai UR, Katuwal M, Poulin R, Gemmell NJ, Dowle E. 2022 Genome assembly and annotation of the European earwig *Forficula auricularia* (subspecies B). *G3* **12**, jkac199. (doi:10.1093/g3journal/jkac199)
34. Bhattarai UR, Poulin R, Gemmell NJ, Dowle E. 2024 Genome assembly and annotation of the mermithid nematode *Mermis nigrescens*. *G3* **14**, jkae023. (doi:10.1093/g3journal/jkae023)
35. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013 STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21. (doi:10.1093/bioinformatics/bts635)
36. Li B, Dewey CN. 2011 RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinf.* **12**, 323. (doi:10.1186/1471-2105-12-323)
37. Love MI, Huber W, Anders S. 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. (doi:10.1186/s13059-014-0550-8)
38. Pantano L. 2022 DEGreport: report of DEG analysis. R package version 1.32.0. See <http://lpantano.github.io/DEGreport/>.
39. Wu T *et al.* 2021 clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb.)* **2**, 100141. (doi:10.1016/j.xinn.2021.100141)
40. Heinz S *et al.* 2010 Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589. (doi:10.1016/j.molcel.2010.05.004)
41. Giridharan S, Srinivasan M. 2018 Mechanism of NF- κ B p65 and strategies for therapeutic manipulation. *J. Inflamm. Res.* **11**, 407–419. (doi:10.2147/JIR.S140188)
42. Patalano SD, Bass PF, Bass JF. 2023 Transcription factors in the development and treatment of immune disorders. *Transcription* **16**, 118–140. (doi:10.1080/21541264.2023.2294623)
43. Sun SC. 2018 The non-canonical NF- κ B pathway in immunity and inflammation. *Nat. Rev. Immunol.* **17**, 545–558. (doi:10.1038/nri.2017.52)
44. Waldschmidt R, Jahn D, Teichmann M, Jahn M, Meissner W, Seifart KH. 1990 Physical and immunological characterization of human transcription factor IIIA. *Eur. J. Biochem.* **194**, 167–174. (doi:10.1111/j.1432-1033.1990.tb19441.x)
45. Budovskaya YV, Wu K, Southworth LK, Jiang M, Tedesco P, Johnson TE, Kim SK. 2008 *Anelt-3/elt-5/elt-6* GATA transcription circuit guides aging in *C. elegans*. *Cell* **143**, 291–303. (doi:10.1016/j.cell.2008.05.044)
46. Bonate R, Kurek G, Hrabak M, Patterson S, Padovan-Neto F, West AR, Steiner H. 2022 Phosphodiesterase 10A (PDE10A): regulator of dopamine agonist-induced gene expression in the striatum. *Cells* **11**, 2214. (doi:10.3390/cells11142214)
47. Day JP, Dow JAT, Houslay MD, Davies S-A. 2005 Cyclic nucleotide phosphodiesterases in *Drosophila melanogaster*. *Biochem. J.* **388**, 333–342. (doi:10.1042/BJ20050057)
48. Francis SH, Turko IV, Corbin JD. 2000 Cyclic nucleotide phosphodiesterases: relating structure and function. *Prog. Nucleic Acid Res. Mol. Biol.* **65**, 1–52. (doi:10.1016/S0079-6603(00)65001-8)
49. Hebb ALO, Robertson HA, Denovan-Wright EM. 2008 Phosphodiesterase 10A inhibition is associated with locomotor and cognitive deficits and increased anxiety in mice. *Eur. Neuropsychopharmacol.* **18**, 339–363. (doi:10.1016/j.euroneuro.2007.08.002)
50. Menniti FS, Chappie TA, Schmidt CJ. 2021 PDE10A inhibitors: clinical failure or window into antipsychotic drug action? *Front. Neurosci.* **14**, 600178. (doi:10.3389/fnins.2020.600178)
51. Liu N, Li T, Wang Y, Liu S. 2021 G-protein coupled receptors (GPCRs) in insects: a potential target for new insecticide development. *Molecules* **26**, 2993. (doi:10.3390/molecules26102993)
52. Scott MM, Marcus JN, Pettersen A, Birnbaum SG, Mochizuki T, Scammell TE, Nestler EJ, Elmquist JK, Lutter M. 2011 Hcrtr1 and 2 signaling differentially regulates depression-like behaviors. *Behav. Brain Res.* **222**, 289–294. (doi:10.1016/j.bbr.2011.02.044)
53. Vallochi AL, Teixeira L, Oliveira KS, Maya-Monteiro CM, Bozza PT. 2018 Lipid droplet, a key player in host–parasite interactions. *Front. Immunol.* **9**, 1022. (doi:10.3389/fimmu.2018.01022)
54. Hanschke L, Heier C, Palacios SJM, Ozek HE, Thiele C, Bauer R, Kuhnlein RP, Bulow MH. 2022 *Drosophila* lipase 3 mediates the metabolic response to starvation and aging. *Front. Aging* **3**, 800153. (doi:10.3389/fragi.2022.800153)
55. Lanz-Mendoza H, Contreras-Garduno J. 2022 Innate immune memory in invertebrates: concept and potential mechanisms. *Dev. Comp. Immunol.* **127**, 104285. (doi:10.1016/j.dci.2021.104285)
56. Li X, Xiong G, Luo M, Mao S, Zhang R, Meng Z, Li J, Liao X. 2025 Peptidoglycan recognition protein PGRP-5 is involved in immune defence and neuro-behavioral disorders in zebrafish embryos. *PLoS One* **20**, e0315714. (doi:10.1371/journal.pone.0315714)
57. Kunwar PS, Starz-Gaiano M, Bainton RJ, Heberlein U, Lehmann R. 2003 Tre1, a G protein-coupled receptor, directs transepithelial migration of *Drosophila* germ cells. *PLoS Biol.* **1**, e80. (doi:10.1371/journal.pbio.0000080)
58. Mesén-Ramírez P, Bergmann B, Tran TT, Garten M, Stäcker J, Naranjo-Prado I, Höhn JK, Zimmerberg J, Spielmann T. 2019 EXP1 is critical for nutrient uptake across the parasitophorous vacuole membrane of malaria parasites. *PLoS Biol.* **17**, e3000473. (doi:10.1371/journal.pbio.3000473)
59. McElwee JJ, Schuster E, Blanc E, Thomas JH, Gems D. 2004 Shared transcriptional signature in *Caenorhabditis elegans* Dauer larvae and long-lived *daf-2* mutants implicates detoxification system in longevity assurance. *J. Biol. Chem.* **279**, 44533–44543. (doi:10.1074/jbc.M406207200)
60. Bhattarai UR, Doherty J-F, Poulin R, Dowle E, Gemmell NJ. 2026 Supplementary material from: Revealing the genetic mechanisms underpinning the parasite-induced water-seeking behaviour of insects through RNA-seq. Figshare. (doi:10.6084/m9.figshare.c.8614245)