

RESEARCH ARTICLE

Open Access

Transcriptional variation and divergence of host-finding behaviour in *Steinernema carpocapsae* infective juveniles



Neil D. Warnock[†], Deborah Cox[†], Ciaran McCoy, Robert Morris and Johnathan J. Dalzell^{* ID}

Abstract

Background: *Steinernema carpocapsae* is an entomopathogenic nematode that employs nictation and jumping behaviours to find potential insect hosts. Here we aimed to investigate the transcriptional basis of variant host-finding behaviours in the infective juvenile (IJ) stage of three *S. carpocapsae* strains (ALL, Breton and UK1), with a focus on neuronal genes known to influence behaviour in other nematode species. Identifying gene expression changes that correlate with variant host-finding behaviours will further our understanding of nematode biology.

Results: RNA-seq analysis revealed that whilst up to 28% of the *S. carpocapsae* transcriptome was differentially expressed ($P < 0.0001$) between strains, remarkably few of the most highly differentially expressed genes ($> 2 \log_2$ fold change, $P < 0.0001$) were from neuronal gene families. *S. carpocapsae* Breton displays increased chemotaxis toward the laboratory host *Galleria mellonella*, relative to the other strains. This correlates with the up-regulation of four srsx chemosensory GPCR genes, and a sodium transporter gene, *asic-2*, relative to both ALL and UK1 strains. The UK1 strain exhibits a decreased nictation phenotype relative to ALL and Breton strains, which correlates with co-ordinate up-regulation of neuropeptide like protein 36 (*nlp-36*), and down-regulation of an srt family GPCR gene, and a distinct *asic-2-like* sodium channel paralogue. To further investigate the link between transcriptional regulation and behavioural variation, we sequenced microRNAs across IJs of each strain. We have identified 283 high confidence microRNA genes, yielding 321 predicted mature microRNAs in *S. carpocapsae*, and find that up to 36% of microRNAs are differentially expressed ($P < 0.0001$) between strains. Many of the most highly differentially expressed microRNAs ($> 2 \log_2$ fold, $P < 0.0001$) are predicted to regulate a variety of neuronal genes that may contribute to variant host-finding behaviours. We have also found evidence for differential gene isoform usage between strains, which alters predicted microRNA interactions, and could contribute to the diversification of behaviour.

Conclusions: These data provide insight to the transcriptional basis of behavioural variation in *S. carpocapsae*, supporting efforts to understand the molecular basis of complex behaviours in nematodes.

Keywords: Nematode, Transcriptome, microRNA, Isoform

Background

Parasitic nematodes employ a variety of behaviours that maximise opportunity for host contact and invasion. These behaviours vary across species, ranging from passive reliance on host ingestion, through pro-active host-finding by migration, nictation, and even jumping [1–3]. Each of these behavioural strategies rely on the incorporation of multiple

sensory inputs, spanning chemosensory, olfactory, mechanosensory, thermosensory and hygroscopic circuits [2, 4]. Nematode host-finding strategies are also remarkably plastic, varying in response to experience and environment [5]. Despite the obvious importance of parasite host-finding behaviour to medical, veterinary and agricultural interests, we know relatively little about the genes involved in regulating these behaviours. It may be possible to develop new approaches to parasite control by targeting components of the parasite host-finding apparatus.

* Correspondence: j.dalzell@qub.ac.uk

[†]Neil D. Warnock and Deborah Cox contributed equally to this work.
School of Biological Sciences, Queen's University Belfast, Belfast, Northern Ireland



The evident complexity of nematode behaviour belies their relative neuroanatomical simplicity. It is thought that the neurochemical complexity of nematodes is central to their diverse behavioural capacity and adaptability [6]; neuropeptides in particular are highly enriched and conserved between nematode species with diverse life history traits [7]. It has also been demonstrated that FMRFamide-like peptide (*flp*) genes are co-ordinately up-regulated in host-finding stages of diverse parasitic nematodes [8], and that they contribute to various host-finding behaviours [8, 9], along with insulin-like peptides [10] and neuropeptide-like proteins [11]. Diverse neuronal gene families also contribute to the surprising behavioural enrichment of such simple organisms [12]. Variation in gene transcript abundance and sequence identity is central to the phenotypic plasticity of cells, tissues and organisms, underpinning behavioural variation.

Small non-coding RNAs also contribute to gene regulation, as a factor of developmental stage, and in response to the environment. Small RNAs have been implicated in driving phenotypic novelty and adaptation within and between species [13–16]. MicroRNAs are small non-coding RNAs that negatively regulate target gene expression across higher organisms [17], and have been shown to modulate neuronal connectivity, synaptic remodelling [18] and memory within the olfactory system of other invertebrates [19]. In vivo cross-linking and immunoprecipitation of microRNA-specific argonaute proteins coupled with sequencing of bound mRNA transcripts in *Caenorhabditis elegans*, demonstrates an enrichment of neuronal gene families. This confirms that many neuronal genes with known involvement in behaviour are biologically relevant microRNA targets [20]. Key to understanding the contribution of neuronal gene function and microRNA regulation to host-finding behaviour in parasitic nematodes is the development of a suitable model system through provision of foundational datasets.

Steinernema spp. nematodes are obligate entomopathogens that invade and kill insect hosts through coordinated action with commensal *Xenorhabdus* bacteria [21]. *Steinernema* infective juveniles (IJs) display qualitatively different host-finding strategies between species [22], representing a unique resource for the comparative analysis of behaviour. *Steinernema carpocapsae* is generally considered to employ an ‘ambushing’ strategy, characterised by nictation and jumping behaviours. Nictation is enacted by nematodes that stand upright, waving their anterior in the air [23]. During nictation, the nematode can respond to sensory stimuli in one of three ways: (i) it can cease nictation and transition to a migratory phase; (ii) it can engage in a torpid ‘standing’ phenotype that may enhance opportunity for host attachment; and (iii) it may jump directionally in response to volatile and

mechanosensory input [8, 23]. Whilst the jumping behaviour is thought to be unique to a small number of *Steinernema* spp., nictation is shared amongst many economically important animal parasitic nematodes, alongside the model nematode *C. elegans*, for which nictation represents a long-range phoretic dispersal behaviour that is regulated by IL2 neurons [24]. In this study our aim was to profile the host-finding behaviours of *S. carpocapsae* strains, and to identify protein-coding and non-coding RNAs that are differentially expressed in strains with variant behaviours. These data will underpin future efforts to understand nematode host-finding behaviour.

Methods

S. carpocapsae culture

Steinernema carpocapsae strains (ALL, Breton and UK1) were maintained in *Galleria mellonella* at 23 °C. *S. carpocapsae* strains were gifted by Prof Ali Mortazavi (University of California, Irvine), Prof Nelson Simões (University of the Azores, Portugal), and BASF UK, respectively. IJs were collected by White trap [25] in a solution of Phosphate Buffered Saline (PBS). Freshly emerged IJs were used for each experiment. Individual biological replicates and RNAseq libraries described below were derived from a mixed population of IJs that emerged from multiple *G. mellonella* cadavers on the same day.

Behavioural assays

Galleria mellonella host-finding assays and dispersal assays were conducted as published [9]; in both instances five biological replicates were assayed across three technical replicates each. A chemosensory index (CI) was calculated following host-finding assays, giving a measure of relative attraction for participating IJs [8]. For the nictation assays, micro-dirt chips were made from a PDMS mould [26], with 3.5% ddH₂O agar. 20 IJs suspended in 1.5 µl PBS were pipetted onto the micro-dirt chip, under a binocular light microscope. Once the liquid had evaporated and the IJs could move freely, the number of nictating IJs was counted at 1, 2.5 and 5 min intervals. Nictation assays were conducted over five biological replicates, each with five technical replicates of 20 IJs each. Each dataset was assessed by Brown-Forsythe and Bartlett’s tests to examine homogeneity of variance between groups. One way ANOVA and Tukey’s multiple comparison tests were then used to assess statistically significant differences in mean across experimental groups. Tests were conducted in Graphpad Prism 7.02.

RNA-seq, differential expression and isoform variant analysis

Six biological replicates of each strain were prepared from approximately 10,000 individuals (80 µl packed volume after centrifugation at 2000 rpm for 2 min) each.

Total RNA was extracted from IJs using TRIzol Reagent (Invitrogen) and DNase treated using the Turbo DNase kit (Ambion) following manufacturer's instructions. RNA quantity were assessed by gel electrophoresis and quantified using Qubit RNA BR Assay Kit (Life Technologies) as per manufacturer's instructions. A total of six transcriptome libraries (150 bp, paired end) were prepared for each *S. carpocapsae* strain, from 1 µg of total RNA each, using the TruSeq RNA Library Prep Kit v2 (Illumina) following manufacturer's instructions. Sequencing was performed on the HiSeq2500 instrument. Fastq files were assessed for quality using the FastQC (v. 0.11.3) package [27]. Adapters, low quality bases, and reads shorter than 36 bp were removed using the Trimmomatic (v. 0.35) package [28]. The trimmed data quality was then re-assessed using the FastQC package. The *S. carpocapsae* genomic contigs (PRJNA202318.WBPS9) and associated GFF file (PRJNA202318.WBPS9) were downloaded from <https://parasite.wormbase.org/index.html>. Annotations were converted to GTF format using Cufflinks (v. 2.2.2.20150701) [29]. High quality reads were then mapped to the *S. carpocapsae* genome [30] using the STAR (v. 2.5.3a) package [31]. Isoform expression levels were quantified using the RSEM (v. 1.2.19) package [32], and integrated EBseq package [33]. Raw read counts mapping to each gene in each sample were consolidated into a single count table. This process was repeated for each isoform. Downstream analyses was performed using R (v. 3.3.1) [34]. Differential expression of genes was quantified using the DESeq2 (v. 1.14.1) package [35]. Graphics were generated in R using RColorBrewer (v. 1.1–2) [36], gplots (v. 3.0.1) [37], geneplotter (v. 1.52) [38] and pheatmap (v. 1.0.8) [39] packages with custom R scripts. Additionally, we conducted a GO enrichment analysis of differentially expressed genes that correlated with behavioural variation between strains, using g:Profiler (<https://biit.cs.ut.ee/gprofiler/gost>), using standard parameters.

MicroRNA sequencing, discovery and quantification

Small RNA libraries were generated from the same RNA samples that were used for matched transcriptomes. A total of six small RNA libraries were prepared for each strain, from 1 µg of total RNA each, using the TruSeq Small RNA library Kit (Illumina) following manufacturer's instructions. 50 bp single-end libraries were sequenced on the HiSeq2500 instrument. Fastq files were assessed for quality using the FastQC (v. 0.11.3) package [27]. Adapters, low quality bases, and reads shorter than 13 bp were removed using the Cutadapt package (v. 1.8) [40]. Sequence reads without adapters were also discarded. Reads that passed QC were mapped to the genome sequence of *S. carpocapsae*, and microRNAs were identified by miRDeep2 (v. 2.0.0.8) [41], using a training set of mature and precursor microRNA sequences

downloaded from miRBase (<http://www.mirbase.org/>). Naming of microRNAs was preferentially aligned with *C. elegans*, as indicated by miRDeep2 output. Novel *S. carpocapsae* microRNAs were named and numbered sequentially, taking care to avoid overlap with any *C. elegans* microRNA. Differentially expressed microRNAs were identified as above, using the DESeq2 package, and were presented using RColorBrewer, gplots, geneplotter and pheatmap packages.

MicroRNA target gene prediction

Three and five prime UnTranslated Regions (UTRs) of computationally predicted *S. carpocapsae* genes were exported from wormbase parasite [30, 42] using the biomaRt function. Retrieved sequences were converted to fasta format and predicted microRNA binding sites were identified using miRanda [43]. Two separate miRanda analyses were performed, using i) default settings and ii) strict settings that require perfect conservation of seed site sequence complementarity between microRNA and target mRNA. Experimental verification of microRNA target predictions indicate that perfect complementarity between the microRNA seed region and target mRNA provides the highest degree of specificity and sensitivity [44]. However, Argonaute CLIP-seq analyses indicate that around 40% of all microRNA-mRNA interactions lack perfect seed region complementarity [45]; limiting analyses to perfect seed region requirements will lead to a substantial number of false negatives. MiRanda allows analyses that span canonical seed region complementarity, and non-canonical interactions, providing a robust overview of interactions that follow experimentally validated examples [44]. In each instance, we have included information of relative target site predictions using both strict and default target identification approaches.

Annotation of neuropeptide, neurotransmitter, GPCR, innexin and ion channel genes

S. carpocapsae neuropeptide gene orthologues were identified via reciprocal BLAST analysis. A list of available *C. elegans* FLP, NLP and INS pre-propeptide sequences were obtained from Wormbase [46] and used as BLASTp and tBLASTn search strings via the Wormbase ParaSite BLAST server [42] under default settings. The protein sequences for overlapping genes associated with each high scoring pair were then employed as BLASTp search strings against the available *C. elegans* protein dataset via the Wormbase BLAST server [46]. Where no overlapping gene annotation was available, novel predicted proteins were generated by concatenating high-scoring return sequences to facilitate reciprocation [47]. The top reciprocal BLAST hit from *C. elegans* was used to assign putative neuropeptide gene names, and subsequent manual comparison of *S. carpocapsae* neuropeptide primary sequences

to established nematode neuropeptide motifs [48] was used to confirm or reassign gene names where appropriate. GPCR, ion channel and innexin genes were exported from Wormbase parasite according to GO term, using the biomaRt function. All neurotransmitter genes were identified by reciprocal BLAST, as above. Additionally, the identity of each gene represented in heatmaps and tables of this manuscript was confirmed by reciprocal BLAST, as above. In a number of instances, different *S. carpocapsae* genes reciprocated to the same *C. elegans* gene. In any such case, a simple numbering system was applied to gene names in order to reflect clustered identity.

Results

Behavioural variation across *S. carpocapsae* strains

S. carpocapsae Breton demonstrated an increased chemotaxis index in response to *G. mellonella* volatiles, indicating that significantly more IJs migrated toward the larvae relative to other strains (Fig. 1a; 0.88 ± 0.009 , relative to 0.43 ± 0.06 for ALL, and 0.31 ± 0.08 for UK1, $P < 0.0001^{****}$). *S. carpocapsae* UK1 exhibited a reduced nictation phenotype relative to the other strains (Fig. 1b; $2.26\% \pm 0.99$ of UK1 IJs were nictating at the 5 min time point, relative to $26.5\% \pm 3.8$ of ALL strain IJs, and $27.7\% \pm 3.7$ of Breton strain IJs,

$P < 0.0001^{****}$). No statistically significant difference was observed in the dispersal behaviour of IJs from *S. carpocapsae* strains (Fig. 1c).

Transcriptional variation across strains

7494 (28%) and 3662 (13.7%) of *S. carpocapsae* UK1 genes were differentially expressed ($P < 0.0001^{****}$) relative to Breton and ALL strains, respectively (Fig. 2; Additional file 1). 4762 (17.7%) of *S. carpocapsae* Breton genes were differentially expressed ($P < 0.0001^{****}$) relative to the ALL strain (Fig. 2). GO term enrichment revealed that transmembrane transport was the most statistically enriched gene classification when considering both enhanced chemotaxis (Breton relative to All and UK1) and decreased nictation (UK1 relative to All and Breton). Several additional statistically significant enrichments suggested a contribution from neuronal gene families (Additional file 1).

22 (0.08%) of the 1509 (5.3%) most highly differentially expressed genes across all pairwise strain comparisons ($> 2 \log_2$ fold, $P < 0.0001^{****}$) were representative of neuronal gene families, based on GO term annotations (Additional file 1). Of these 22 neuronal genes, only eight were observed to share differential expression

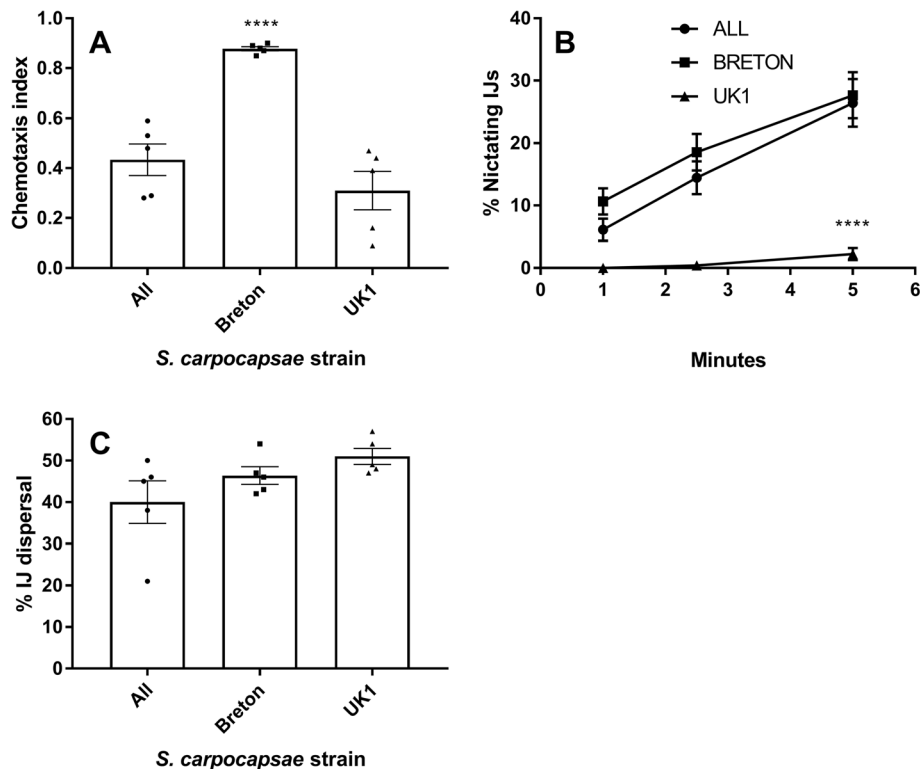


Fig. 1 *S. carpocapsae* strains display variant host-finding behaviours. **a** Mean chemotaxis index of *S. carpocapsae* strains in response to *Galleria mellonella* larvae; **b** Mean number of nictating *S. carpocapsae* IJs over a time-course; **c** Mean percentage dispersal of *S. carpocapsae* IJs into the peripheral assay zone (dispersed) after a 1 h timecourse. Data points represent mean \pm SEM; assessed by ANOVA and Tukey's multiple comparison test using Graphpad Prism 7.02; $P < 0.0001^{****}$

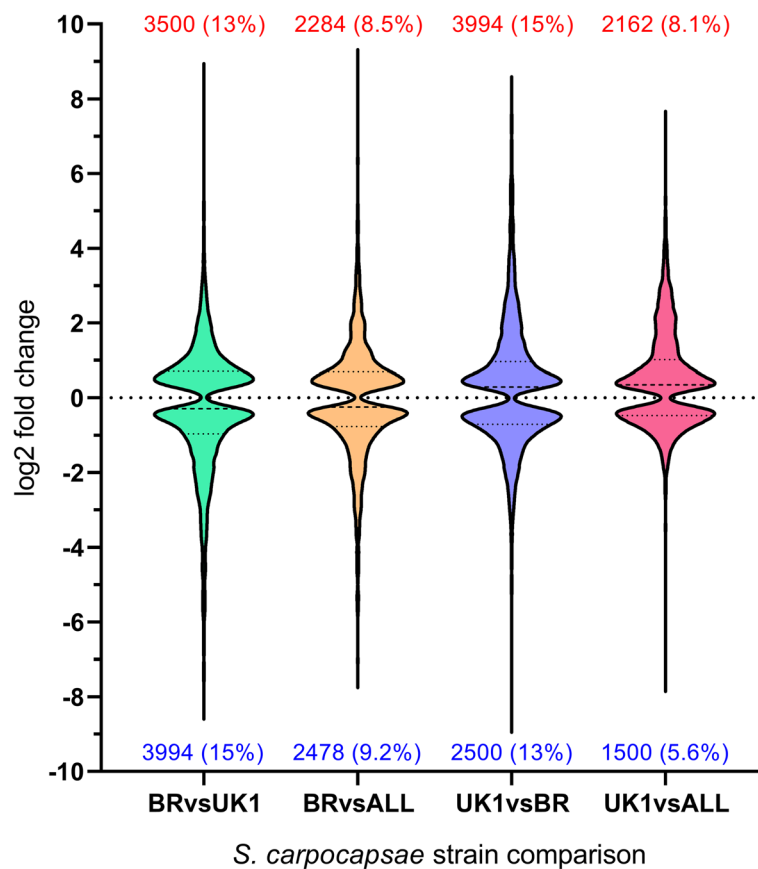


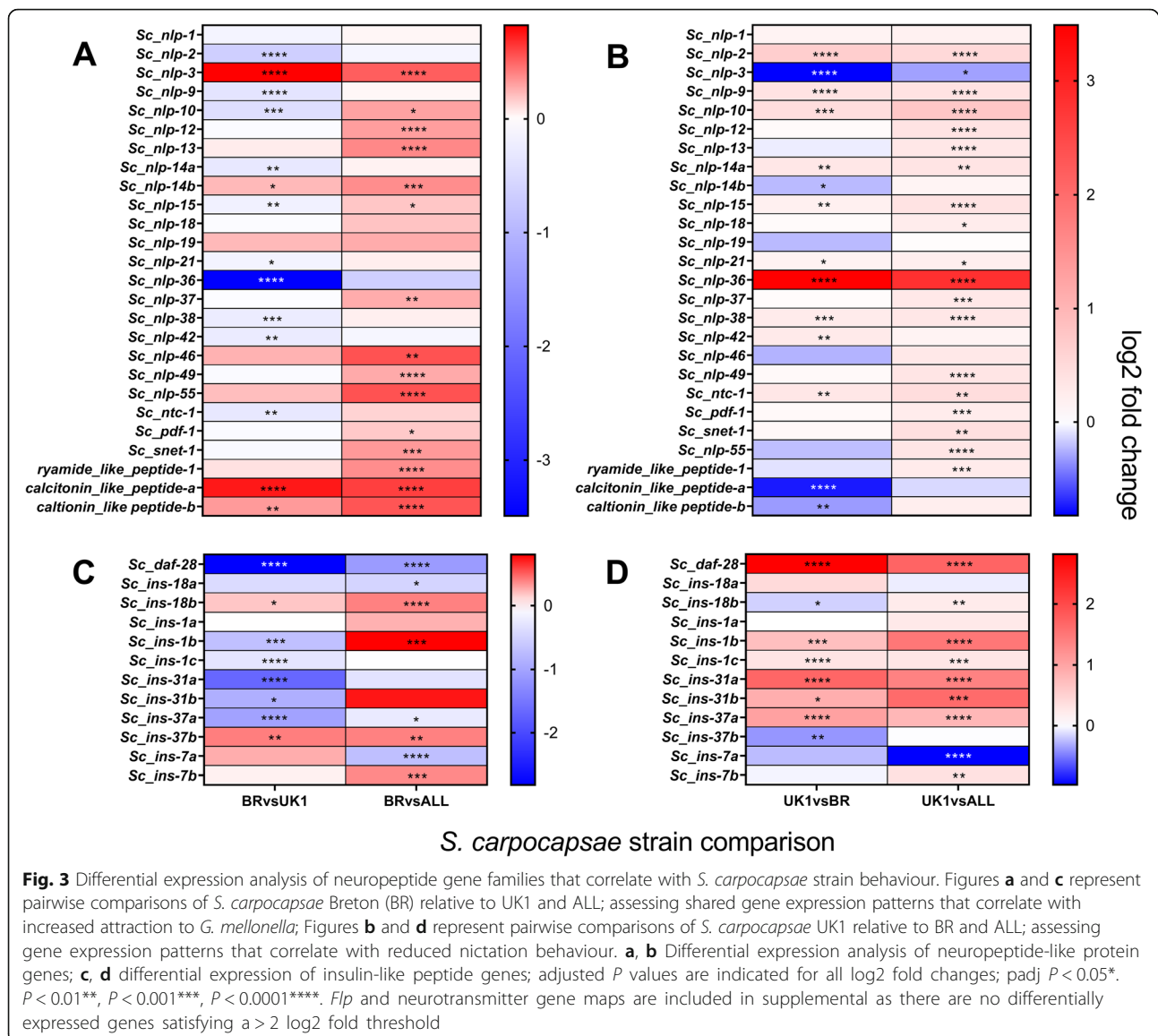
Fig. 2 Violin plot showing significantly up-regulated and down-regulated genes across pairwise *S. carpocapsae* strain comparisons. Statistically significant ($P < 0.0001^{****}$) differentially expressed genes are plotted irrespective of log₂ fold change. Total number and relative percentage of up-regulated and down-regulated genes are presented above and below each plot, respectively

patterns across pairwise comparisons that correlate with distinct IJ behaviours; either increased chemotaxis (*S. carpocapsae* Breton; Fig. 1a) or decreased nictation (*S. carpocapsae* UK1; Fig. 1b). Specifically, shared up or down-regulation in *S. carpocapsae* Breton relative to both ALL and UK1 strains would correlate gene differential expression with increased chemotaxis to *G. mellonella* (Fig. 1a). Conversely, shared up or down-regulation in *S. carpocapsae* UK1 relative to both Breton and ALL strains would correlate gene differential expression with reduced nictation behaviour (Fig. 1b). In order to explore the transcriptional regulation of neuronal gene families implicated in these behavioural differences, we assessed each gene family represented in the highly differentially expressed category, as defined above. This encompassed G-protein coupled receptors (GPCRs), ion channels, innexins, neuropeptides and neurotransmitter synthesis, degradation and transport genes (Additional file 1).

Neuropeptide and neurotransmitter genes

Nlp-36 was the only neuropeptide-like protein gene to demonstrate significant fold change differences (> 2 log₂ fold, $P < 0.0001^{****}$) that correlate with decreased

nictation behaviour in *S. carpocapsae* UK1 across pairwise comparisons (up-regulated 2.8 and 3.5 log₂ fold, $P < 0.0001^{****}$ relative to Breton and ALL strains, respectively) (Fig. 3a-b; Additional file 2). Though falling short of our shared > 2 log₂ fold change threshold, insulin-like peptide gene, *daf-28*, was the single most differentially regulated *ins* gene, exhibiting a polarised expression pattern that correlates inversely with both enhanced chemotaxis toward *G. mellonella* (*S. carpocapsae* Breton), and reduced nictation behaviour (*S. carpocapsae* UK1) (down-regulated 1.1 and 2.8 log₂ fold, $P < 0.0001^{****}$ in Breton relative to ALL and UK1, respectively; up-regulated 1.7 and 2.8 log₂ fold, $P < 0.0001^{****}$ in UK1 relative to ALL and Breton, respectively) (Fig. 2c-d; Additional file 2). *Flp* genes were comparatively less variant between strains, with *flp-34* exhibiting the largest pairwise expression change that correlates with increased chemotaxis behaviour in *S. carpocapsae* Breton (up-regulated 0.7 and 0.6 log₂ fold, $P < 0.0001^{****}$, relative to UK1 and ALL strains, respectively) (Additional file 2). The tyrosine decarboxylase gene, *Sc-tdc-1*, was the most differentially expressed neurotransmitter gene (down-regulated 1.4 and 1.9 log₂ fold, $P < 0.0001^{****}$ in *S.*



carpocapsae UK1 relative to ALL and Breton, respectively), correlating with decreased nictation behaviour, but falling short of a shared > 2 log₂ fold change threshold (Additional file 2).

GPCR, innexin and ion channel genes

Significant up-regulation of four *srsx* GPCR genes (*Sc-srsx-25v*, *Sc-srsx-3ii*, *Sc-srsx-22i*, and *Sc-srsx-24ii*) (> 2 log₂ fold, *P* < 0.0001****) correlates with increased chemotaxis of *S. carpocapsae* Breton to the insect host *G. mellonella*, relative to UK1 and ALL strains (Fig. 4a; Additional file 2). By way of contrast, the *Sc-srt-62* chemosensory GPCR gene was co-ordinately down-regulated in *S. carpocapsae* UK1, correlating with reduced nictation behaviour (down-regulated 2.6 and 2.5 log₂ fold, *P* < 0.0001**** relative to ALL and Breton strains, respectively) (Fig. 4b; Additional file 2).

Although falling below our threshold of shared > 2 log₂ fold change across both strain pairwise comparisons, *Sc-inx-7ii* demonstrated the largest expression change in the innexin / gap junction gene family, correlating with increased chemotaxis behaviour (down-regulated 1.03 and 2.65 log₂ fold in *S. carpocapsae* Breton relative to UK1 and ALL strains, respectively; Fig. 4c-d; Additional file 2). Two paralogous *asic-2* sodium channel genes were found to be differentially expressed, and inversely related to altered chemotaxis behaviour (*Sc-asic-2ii*; down-regulated 2.3 and 2.7 log₂ fold, *P* < 0.0001**** in *S. carpocapsae* Breton, relative to ALL and UK1 strains, respectively; Fig. 4e, Additional file 2), and reduced nictation behaviour (*Sc-asic-2i*; up-regulated 2.9 and 3 log₂ fold, *P* < 0.0001**** in *S. carpocapsae* UK1, relative to ALL and Breton strains, respectively; Fig. 4f, Additional file 2).

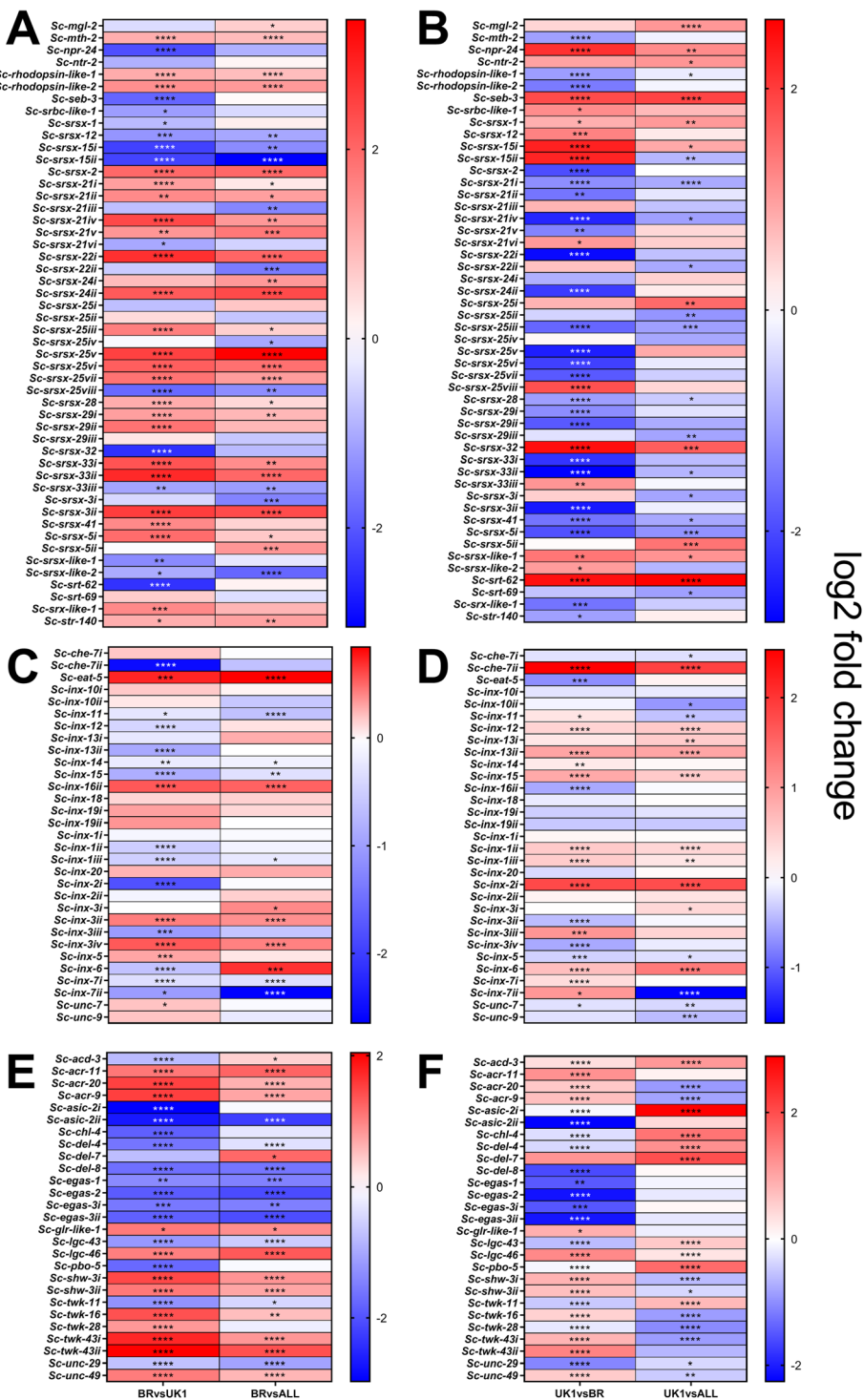


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Differential expression analysis of GPCR, innexin and ion channel gene families that correlate with *S. carpocapsae* strain behaviour. Figures **a**, **c** and **e** represent pairwise comparisons of *S. carpocapsae* Breton (BR) relative to UK1 and All; assessing shared gene expression patterns that correlate with increased chemotaxis to *G. mellonella*; Figures **b**, **d** and **f** represent pairwise comparisons of *S. carpocapsae* UK1 relative to BR and All; assessing gene expression patterns that correlate with reduced nictation behaviour. **a**, **b** putative GPCR genes with > 1 log₂ fold difference between at least one of the pairwise comparisons; **c** and **d** putative innexin genes across all log₂ fold changes; **e** and **f** putative ion channel genes with > 1 log₂ fold difference between at least one of the pairwise comparisons. *P* values are indicated for all log₂ fold changes; padj $P < 0.05^*$, $P < 0.01^{**}$, $P < 0.001^{***}$, $P < 0.0001^{****}$

MicroRNA prediction and quantification

Following miRDeep2 identification, quality control and manual curation, a total of 283 high confidence microRNA genes and 321 predicted mature microRNAs (306 of which are unique within the genome) were identified across a deep analysis ($n = 18$ libraries) of *S. carpocapsae* strains (Additional file 3). One hundred fifty of the unique mature microRNA sequences were previously identified in *S. carpocapsae* Breton [49], with an additional 156 unique mature microRNAs identified within this study. 102 (36%) and 103 (36%) *S. carpocapsae* Breton microRNAs were differentially expressed ($P < 0.0001^{****}$) relative to UK1 and ALL strains, respectively. 50 (17.6%) *S. carpocapsae* UK1 microRNAs were differentially expressed ($P < 0.0001^{****}$) relative to *S. carpocapsae* ALL (Fig. 5; Additional file 3).

Three microRNAs (*Sc-mir-117*, *Sc-mir-27*, and *Sc-mir-774*) were highly differentially expressed, correlating with enhanced chemotaxis behaviour in *S. carpocapsae* Breton (> 6 log₂ fold, $P < 0.0001^{****}$ relative to UK1 and All strains). A further 18 microRNAs were likewise differentially expressed and correlated with enhanced chemotaxis behaviour (> 2 log₂ fold, $P < 0.0001^{****}$) (Fig. 6a; Additional files 3 and 4). Comparatively fewer microRNAs were differentially regulated and correlated with reduced nictation behaviour in *S. carpocapsae* UK1, with *Sc-mir-772* (up-regulated 12.9 and 12.7 log₂ fold, $P < 0.0001^{****}$ relative to Breton and All strains, respectively) and *Sc-mir-773* (up-regulated 8.7 and 8.5 log₂ fold, $P < 0.0001^{****}$ relative to Breton and All strains, respectively) representing notable exceptions. A further five microRNAs (*Sc-mir-754*, *Sc-mir-756*, *Sc-mir-760*, *Sc-let-7*

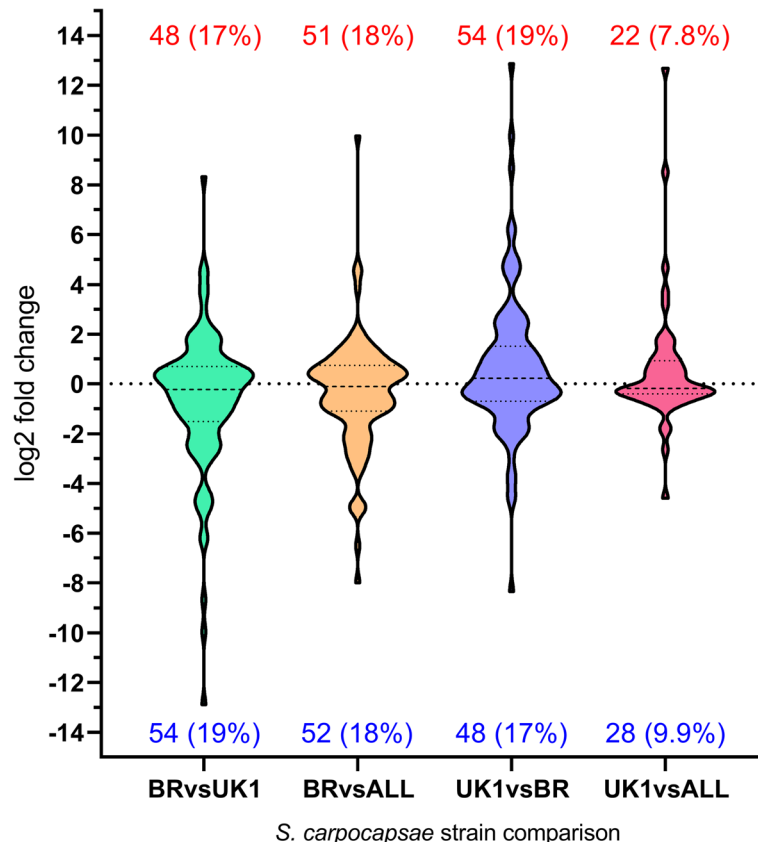
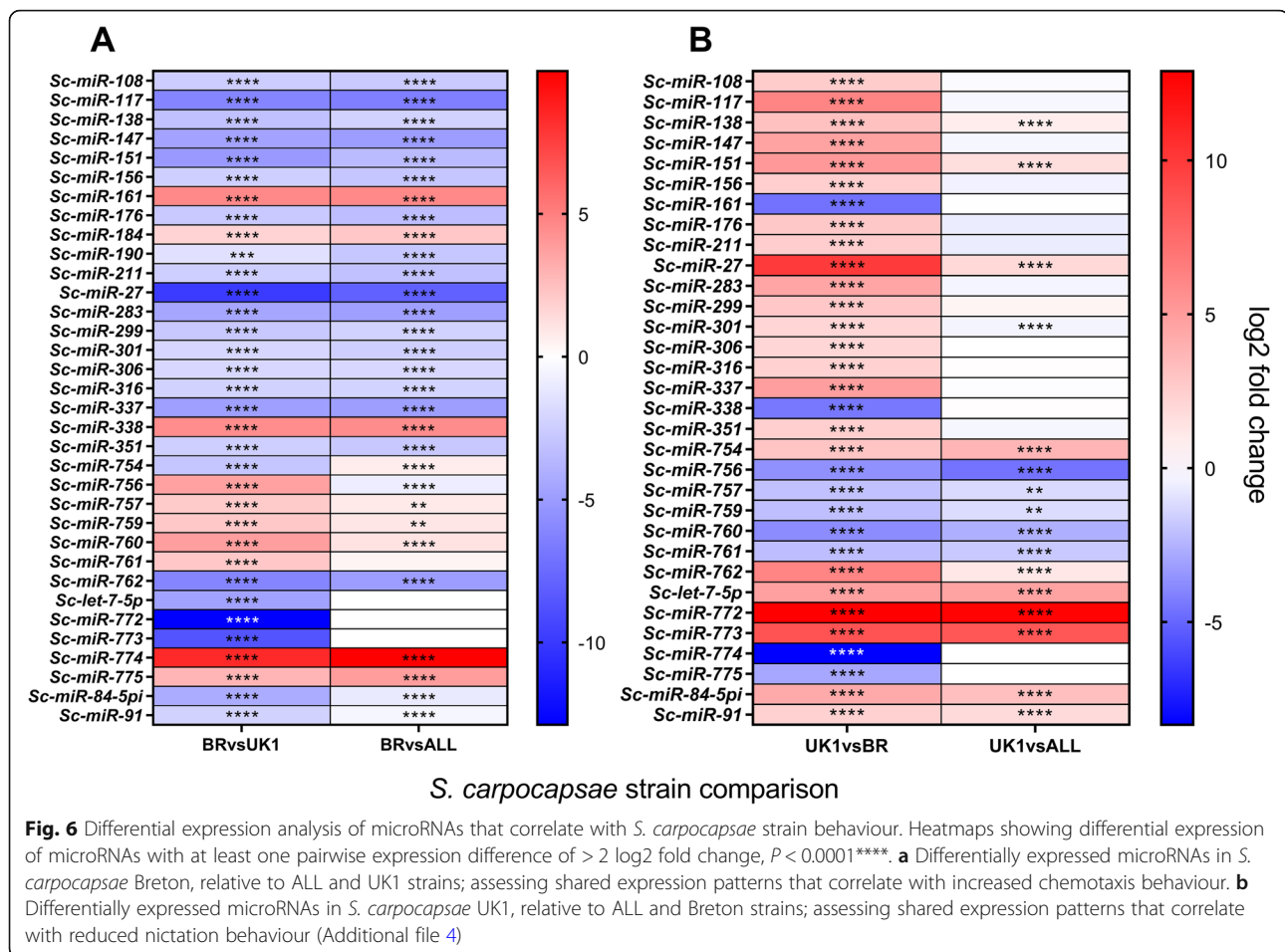


Fig. 5 Violin plot showing significantly up-regulated and down-regulated microRNA genes across pairwise strain comparisons. Statistically significant ($P < 0.0001^{****}$) differentially expressed genes are plotted irrespective of log₂ fold change. Total number and relative percentage of up-regulated and down-regulated microRNA genes are presented above and below each plot, respectively



and *Sc-mir-84-5pi*) were likewise differentially expressed and correlated with reduced nictation behaviour (> 2 log₂ fold, $P < 0.0001$ ***) (Fig. 6b; Additional files 3 and 4).

MicroRNA target analysis

In silico microRNA target prediction through miRanda suggests a substantial bias towards interactions with gene 5'UTRs. A total of 231,120 microRNA interactions, representing 247,237 binding sites are predicted for the 5'UTR of *S. carpocapsae* genes, relative to 127,796 microRNA interactions, representing 132,309 binding site predictions within the 3'UTR of *S. carpocapsae* genes. Through screening each of the most differentially expressed (> 2 log₂ fold change, $P < 0.0001$ ****) microRNAs that correlate with behavioural variants across pairwise comparisons (Fig. 6; Additional file 4), we identify a substantial number of predicted interactions with neuronal gene families considered in this study (Table 1). These datasets demonstrate potential cooperation between microRNAs that are predicted to interact with shared target genes. For example, the ion channel gene *Sc-asic-2ii* is a predicted target for *Sc-mir-147*, *Sc-mir-301-3p*, and *Sc-mir-316*, all highly differentially expressed, and correlated with increased chemotaxis

behaviour (Table 1; Fig. 5a). Similarly, the *let-7* family members, *Sc-mir-84-5pi* and *Sc-let-7* co-ordinately target two *ins-1* paralogues; both microRNAs are highly differentially expressed, and correlated with reduced nictation behaviour in *S. carpocapsae* UK1 (Fig. 5b). Our data also suggest that different microRNAs are predicted to interact with and converge on a number of shared neuronal gene targets considered here (Table 1). For example, *Sc-mir-301-3p* is predicted to simultaneously target the *twk-12*, *egas-1* and *asic-2* ion channel genes, alongside the *nlp-39* neuropeptide-like protein gene, correlating with increased chemotaxis behaviour (Table 1; Figs. 5a and 1a). MiRanda microRNA target predictions across both strict and default settings for global 5' and 3'UTRs are presented in Additional files 5, 6, 7 and 8.

Differential isoform usage between strains

Two genes implicated in the synthesis and transport of classical neurotransmitters were found to exhibit statistically significant differences in isoform preference across *S. carpocapsae* strains following RSEM analysis (Figs. 6, 7 and 8; Additional file 9). Interestingly, these isoform variants exhibit altered UTR sequences in

Table 1 Predicted gene targets for differentially expressed microRNAs (> 2 log₂ fold, $P < 0.0001$) that correlate with *S. carpocapsae* strain behaviour. Differentially expressed *S. carpocapsae* microRNAs were assessed for binding sites across 5' and 3'UTRs of all ion channel, innexin, GPCR, neurotransmitter and neuropeptide genes. *Ce* target refers to direct 1 to 1 gene orthologues that are biochemically confirmed microRNA targets in *C. elegans* [20]. U refers to the number of predicted interacting microRNAs following default microRNA target prediction; S refers to the number of predicted interacting microRNAs following strict microRNA target prediction. Within the implicated behaviour column, "Both (inverse int.*)" refers to involvement in both behaviours, through polarised regulation states (significantly upregulated for one, and significantly downregulated for the other)

	microRNA	Gene	Gene Name	Gene family	<i>Ce</i> target	D	S	Implicated behaviour
3'UTR	<i>sc-miR-301-3p</i>	L596_g12890.t2	<i>twk-12</i>	Ion Channel	y	2	1	Chemotaxis
	<i>sc-miR-301-3p</i>	L596_g12890.t3	<i>twk-12</i>	Ion Channel	Y	2	1	Chemotaxis
	<i>sc-miR-301-3p</i>	L596_g12890.t1	<i>twk-12</i>	Ion Channel	y	2	1	Chemotaxis
	<i>sc-miR-759</i>	L596_g2695.t1	<i>acr-21</i>	Ion Channel		2	1	Nictation
	<i>sc-miR-138</i>	L596_g3511.t1	<i>eat-5</i>	Innexin		2	0	Chemotaxis
	<i>sc-miR-117</i>	L596_g24535.t1	<i>ador-1</i>	GPCR		2	0	Chemotaxis
	<i>sc-miR-283</i>	L596_g24535.t1	<i>ador-1</i>	GPCR		2	0	Chemotaxis
	<i>sc-miR-84-5pi</i>	L596_g1430.t1	<i>ins-1b</i>	Neuropeptide	y	2	1	Nictation
5'UTR	<i>sc-miR-147</i>	L596_g23064.t1	<i>asic-2ii</i>	Ion Channel		2	2	Chemotaxis
	<i>sc-miR-147</i>	L596_g14933.t1	<i>unc-36</i>	Ion Channel		2	2	Chemotaxis
	<i>sc-miR-156</i>	L596_g11170.t1	<i>egas-1</i>	Ion Channel		2	2	Chemotaxis
	<i>sc-miR-176</i>	L596_g11280.t1	<i>lgc-22</i>	Ion Channel		2	2	Chemotaxis
	<i>sc-miR-184</i>	L596_g11170.t1	<i>egas-1</i>	Ion Channel		3	0	Chemotaxis
	<i>sc-miR-184</i>	L596_g13756.t1	<i>acr-25</i>	Ion Channel		2	2	Chemotaxis
	<i>sc-miR-190</i>	L596_g21438.t2	<i>unc-2</i>	Ion Channel	y	3	3	Chemotaxis
	<i>sc-miR-190</i>	L596_g21438.t1	<i>unc-2</i>	Ion Channel	y	3	3	Chemotaxis
	<i>sc-miR-190</i>	L596_g25880.t2	<i>slo-1</i>	Ion Channel	y	2	2	Chemotaxis
	<i>sc-miR-190</i>	L596_g25880.t1	<i>slo-1</i>	Ion Channel	y	2	2	Chemotaxis
	<i>sc-miR-211</i>	L596_g19116.t1	<i>unc-49</i>	Ion Channel	y	5	5	Chemotaxis
	<i>sc-miR-299</i>	L596_g11170.t1	<i>egas-1</i>	Ion Channel		2	1	Chemotaxis
	<i>sc-miR-301-3p</i>	L596_g11170.t1	<i>egas-1</i>	Ion Channel		3	1	Chemotaxis
	<i>sc-miR-301-3p</i>	L596_g23064.t1	<i>asic-2ii</i>	Ion Channel		2	1	Chemotaxis
	<i>sc-miR-301-5pi</i>	L596_g9175.t1	<i>clhm-1</i>	Ion Channel		2	2	Chemotaxis
	<i>sc-miR-301-5pi</i>	L596_g13756.t1	<i>acr-25</i>	Ion Channel		2	1	Chemotaxis
	<i>sc-miR-301-5pii</i>	L596_g13756.t1	<i>acr-25</i>	Ion Channel		2	1	Chemotaxis
	<i>sc-miR-301-5pii</i>	L596_g11170.t1	<i>egas-1</i>	Ion Channel		2	2	Chemotaxis
	<i>sc-miR-301-5pii</i>	L596_g9175.t1	<i>clhm-1</i>	Ion Channel		2	0	Chemotaxis
	<i>sc-miR-306</i>	L596_g21215.t1	<i>twk-46</i>	Ion Channel		2	1	Chemotaxis
	<i>sc-miR-316</i>	L596_g14933.t1	<i>unc-36</i>	Ion Channel		2	2	Chemotaxis
	<i>sc-miR-316</i>	L596_g23064.t1	<i>asic-2ii</i>	Ion Channel		2	2	Chemotaxis
	<i>sc-miR-337</i>	L596_g9175.t1	<i>clhm-1</i>	Ion Channel		2	0	Chemotaxis
	<i>sc-miR-338</i>	L596_g14875.t1	<i>kvs-5</i>	Ion Channel		2	1	Chemotaxis
	<i>sc-miR-351</i>	L596_g15440.t1	<i>twk-43</i>	Ion Channel		2	1	Chemotaxis
	<i>sc-miR-756</i>	L596_g9186.t2	<i>twk-2</i>	Ion Channel		2	2	Nictation
	<i>sc-miR-756</i>	L596_g9186.t1	<i>twk-2</i>	Ion Channel		2	2	Nictation
	<i>sc-miR-759</i>	L596_g15570.t1	<i>shk-1</i>	Ion Channel	y	2	1	Nictation
	<i>sc-miR-772</i>	L596_g15357.t1	<i>clh-5</i>	Ion Channel		2	1	Nictation
	<i>sc-miR-772</i>	L596_g15357.t2	<i>clh-5</i>	Ion Channel		2	1	Nictation
	<i>sc-miR-772</i>	L596_g11576.t1	<i>avr-14</i>	Ion Channel	y	2	0	Nictation

Table 1 Predicted gene targets for differentially expressed microRNAs ($> 2 \log_2$ fold, $P < 0.0001$) that correlate with *S. carpocapsae* strain behaviour. Differentially expressed *S. carpocapsae* microRNAs were assessed for binding sites across 5' and 3'UTRs of all ion channel, innexin, GPCR, neurotransmitter and neuropeptide genes. *Ce* target refers to direct 1 to 1 gene orthologues that are biochemically confirmed microRNA targets in *C. elegans* [20]. U refers to the number of predicted interacting microRNAs following default microRNA target prediction; S refers to the number of predicted interacting microRNAs following strict microRNA target prediction. Within the implicated behaviour column, "Both (inverse int.)" refers to involvement in both behaviours, through polarised regulation states (significantly upregulated for one, and significantly downregulated for the other) (Continued)

microRNA	Gene	Gene Name	Gene family	<i>Ce</i> target	D	S	Implicated behaviour
<i>sc-miR-775</i>	L596_g20944.t1	<i>twk-29</i>	Ion Channel		2	2	Chemotaxis
<i>sc-miR-91</i>	L596_g12530.t1	<i>unc-58</i>	Ion Channel		2	1	Nictation
<i>sc-miR-759</i>	L596_g25584.t1	<i>unc-7</i>	Innexin	Y	2	2	Nictation
<i>sc-miR-138</i>	L596_g9922.t1	<i>gar-2</i>	GPCR	y	2	0	Chemotaxis
<i>sc-miR-151</i>	L596_g14843.t1	<i>npr-26</i>	GPCR		2	2	Both (inverse int.)
<i>sc-miR-184</i>	L596_g17321.t1	<i>npr-23</i>	GPCR	y	2	0	Chemotaxis
<i>sc-miR-27</i>	L596_g22293.t1	<i>C30A5.10</i>	GPCR		2	1	Both (inverse int.)
<i>sc-miR-301-5pi</i>	L596_g14261.t2	<i>npr-11</i>	GPCR		2	1	Chemotaxis
<i>sc-miR-351</i>	L596_g16439.t1	<i>srsx-24</i>	GPCR		2	2	Chemotaxis
<i>sc-miR-756</i>	L596_g17321.t1	<i>npr-23</i>	GPCR	y	2	1	Nictation
<i>sc-miR-761</i>	L596_g25676.t1	<i>gar-1</i>	GPCR		2	1	Nictation
<i>sc-miR-773</i>	L596_g25676.t1	<i>gar-1</i>	GPCR		2	0	Nictation
<i>Sc-let-7</i>	L596_g13647.t1	<i>ins-1a</i>	Neuropeptide	y	2	1	Nictation
<i>sc-miR-301-3p</i>	L596_g14298.t1	<i>nlp-49</i>	Neuropeptide		2	1	Chemotaxis
<i>sc-miR-84-5pi</i>	L596_g5500.t2	<i>vglu-2</i>	Neurotransmission		2	2	Both (inverse int.)

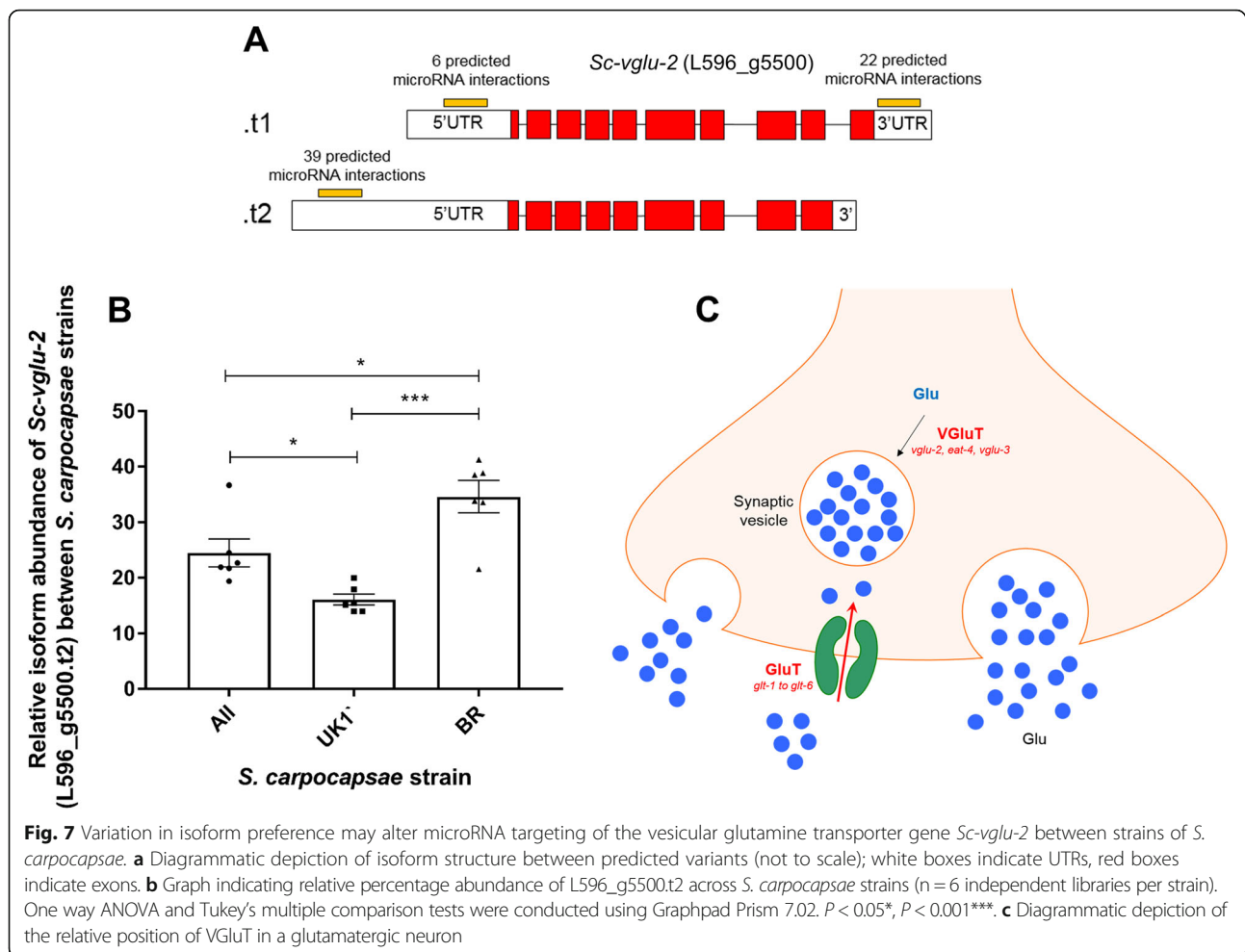
additional to altered exon usage. In silico microRNA target predictions were conducted for all predicted microRNAs relative to the respective isoform UTRs, identifying quantitative differences in predicted microRNA interactions with the various gene isoforms (Table 2; Additional file 9).

Discussion

This study has generated high quality transcriptomes ($n = 18$) and small RNA libraries ($n = 18$) for the IJ stage of three *S. carpocapsae* strains, and has catalogued a number of protein-coding and non-coding RNAs that are differentially expressed between strains with distinct host-finding behaviours. We have applied an arbitrary threshold of $> 2 \log_2$ fold change ($P < 0.0001^{****}$) across both pairwise strain comparisons in order to link gene expression markers to behavioural differences. Many of these genes have already been shown to regulate behaviour in the model nematode *C. elegans*, however none have been implicated specifically in host-finding behaviours of parasitic nematodes. These data support the potential benefit of establishing and investing in natural diversity resources for parasitic nematodes, alongside more expansive studies in the model nematode *C. elegans* [50, 51].

GO term enrichment revealed that a range of gene families and biological processes were enriched in differentially expressed gene sets that correspond with behavioural variation in *S. carpocapsae* strains (Additional file 1).

Transmembrane transport was the single most enriched biological process when considering either nictation or chemotaxis behaviours, following by a series of metabolic and biosynthetic processes. Of particular interest, several neuronal functions were also enriched. Given that we know much of how neuronal genes regulate behaviour in the model nematode *C. elegans*, neuronal gene families became the focus of our analysis. Only one neuropeptide gene, *nlp-36*, was found to exhibit differential expression ($> 2 \log_2$ -fold, $P < 0.0001^{****}$) that correlates with distinct strain behaviours. NLP-36 is positively regulated by the cyclic nucleotide-gated channel subunit TAX-2 in *C. elegans*, which is implicated in the regulation of olfaction, chemosensation, thermosensation and axon guidance for a number of sensory neurons [52, 53]. In *S. carpocapsae* UK1, significant up-regulation of *nlp-36* correlates with reduced nictation behaviour (Figs. 1b and 3b). Down-regulation of insulin signalling during *C. elegans* development has been found to influence nictation behaviour in dauer juveniles, indicating a tuning of behaviour to environmental conditions [10]. Gruner et al. [54] point to a potential neuronal programming circuit for behavioural adjustment in *C. elegans* through the combined influence of insulin signalling, TRPV signalling and agonism of NPR-1, a known receptor of neuropeptides, FLP-21 and FLP-18. Although the differential expression of *daf-28* does not meet the shared $> 2 \log_2$ fold change threshold we have applied to sequencing datasets, it is nonetheless down-regulated in *S. carpocapsae*



Breton, relative to the other strains, correlating with enhanced chemotaxis toward *G. mellonella*. Conversely, *daf-28* was up-regulated in UK1 relative to both Breton and All strains (Fig. 3a, b), correlating with reduced nictation (Fig. 1b). As was found for the two *asic-2* sodium channel paralogues, *daf-28* and other genes could function as part of a neurobiological switch, tuning behavioural strategy toward either migratory or stationary host-finding modes. It may be informative to track the dynamic abundance of such genes in IJs naturally enacting and transitioning between these behaviours, or in conditions that are known to enhance or suppress these behaviours to further strengthen the correlation of expression with behaviour [55, 56].

The up-regulation of *srsx* GPCR genes in *S. carpocapsae* Breton correlates with enhanced chemotaxis to the lab host insect *G. mellonella* (Fig. 4a). One of these genes is orthologous to a cluster of *srsx* genes (*srsx-22* and *srsx-24*) that are enriched in dauer stage *C. elegans*. It has been shown that dauer stage *Caenorhabditis* spp. are attracted to certain insect species, increasing the opportunity to engage in phoresis [57]. This cluster of shared *srsx* GPCRs may therefore mediate this attraction, in isolation, or in

synergy with other such receptors. Two *asic-2* sodium channel paralogues (denoted here as *Sc-asic-2i* and *Sc-asic-2ii*) are also notable as being differentially expressed and inversely correlated with both increased chemotaxis and reduced nictation behaviours (Fig. 4e-f). ASIC-2 is known to regulate aspects of nematode body posture and mechanosensation in *C. elegans* [58].

Small non-coding RNAs, including microRNAs, are increasingly implicated in complex aspects of biology and behaviour [51, 59–61]. We conducted a deep analysis of small RNA profiles across *S. carpocapsae* strains, and found a substantial degree of variation in relative abundance (Figs. 5 and 6, Additional file 3). Numerous microRNAs are differentially expressed, and correlate with behavioural differences across pairwise comparisons (Fig. 6, Additional file 4). Key to extrapolating biologically relevant information from microRNA networks is the identification of gene targets. Although there are many in silico tools that predict microRNA-mRNA transcript interactions, false positives are likely to be common [62]. Whilst many factors influence the reality and significance of predicted interactions, bioavailability of target

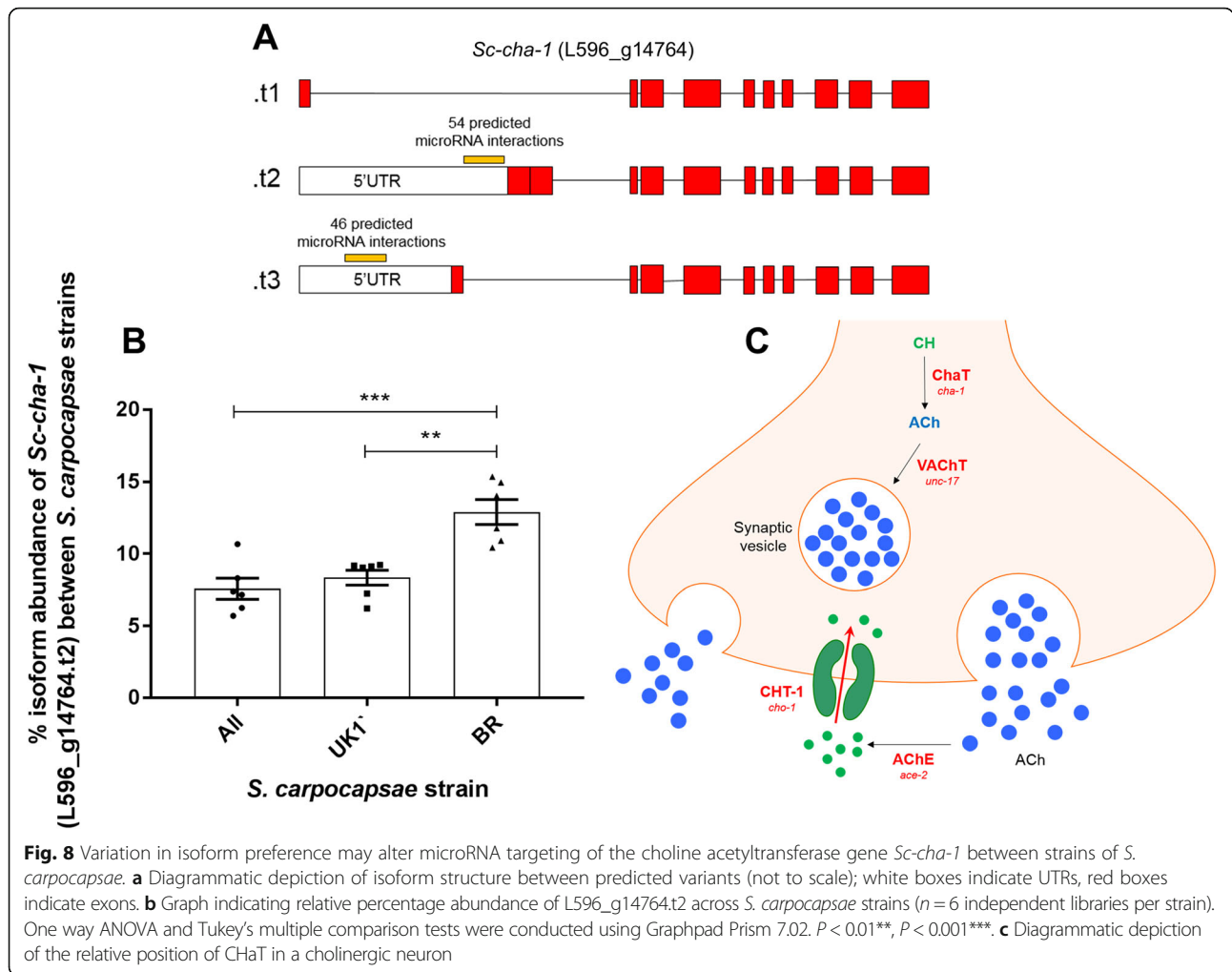


Fig. 8 Variation in isoform preference may alter microRNA targeting of the choline acetyltransferase gene *Sc-cha-1* between strains of *S. carpocapsae*. **a** Diagrammatic depiction of isoform structure between predicted variants (not to scale); white boxes indicate UTRs, red boxes indicate exons. **b** Graph indicating relative percentage abundance of L596_g14764.t2 across *S. carpocapsae* strains ($n = 6$ independent libraries per strain). One way ANOVA and Tukey's multiple comparison tests were conducted using Graphpad Prism 7.02. $P < 0.01^{**}$, $P < 0.001^{***}$. **c** Diagrammatic depiction of the relative position of ChAT in a cholinergic neuron

gene and microRNA in terms of spatial and temporal expression patterns will be key, along with the number of available microRNA binding sites, the relative enthalpy of binding interactions, and local competition for available microRNAs. In order to build on the basic knowledge presented in this study, it will be necessary to biochemically validate gene transcripts as microRNA targets through Argonaute CLIP-seq [45], and further, to demonstrate co-localisation of microRNA and target mRNA transcript to

confirm interaction of discrete partners. This represents a substantial, but necessary task if we are to unravel the biological significance of microRNA regulation in the context of complex phenotypes and behaviours. Previous publications have employed a hierarchical and cooperative in silico target prediction approach using several programs simultaneously to arrive at an agreed set of targets [59]. Whilst this will certainly reduce the complexity of any target gene set, it will also constrain the output according

Table 2 Differential microRNA targeting across *Sc-cha-1* and *Sc-vglu-2* gene isoform 5' and 3'UTRs. D denotes number of predicted microRNA interactions following default target prediction in miRanda; S denotes number of predicted microRNA interactions following strict target prediction in miRanda. Numbers in brackets indicate the relative number of binding sites across both D and S settings

Transcript ID	Mean %Isoform abundance			N° microRNA interactions in 3'UTR (binding sites)		N° microRNA interactions in 5'UTR (binding sites)	
	ALL	UK1	BR	D	S	D	S
L596_g14764.t1	90.9%	90.1%	85.8%	–	–	–	–
L596_g14764.t2	7.6%	8.3%	12.9%	–	–	54 (57)	36 (36)
L596_g14764.t3	1.5%	1.6%	1.3%	–	–	46 (49)	31 (31)
L596_g5500.t1	75.5%	83.9%	65.4%	22 (22)	5 (5)	6	4 (4)
L596_g5500.t2	24.5%	16.1%	34.6%	–	–	39 (49)	25 (27)

to the most stringent program, leading inevitably to false negatives when perfect seed site complementarity is required by one or more programs. Here we have employed a dual analysis strategy within the miRanda discovery environment [43], using both default and strict discovery modes; the latter requires perfect seed site complementarity between microRNA and target mRNA transcript. Collectively, this strategy maps well to biochemically-validated microRNA target interactions, including those that do not require perfect seed site binding [45]. However, as with any in silico prediction approach, biological validation is still required to corroborate interactions.

Our datasets suggest a strong microRNA target site enrichment within predicted 5'UTRs of *S. carpocapsae*. This may be biologically significant, or could represent an artefact of computational UTR prediction. Ultimately, higher confidence interactions could be established by sequencing full length transcripts and confirming UTR identity on a transcriptome wide scale. In silico target prediction for differentially expressed ($> 2 \log_2$ fold, $P < 0.0001^{****}$) and behaviourally correlated microRNAs points to instances of potential microRNA co-operation. For example, the ion channel genes *Sc-asic-2ii* and *Sc-egas-1* are predicted targets for three and five individual differentially expressed microRNAs each (Table 1). The neuropeptide GPCR gene, *Sc-npr-23* is likewise targeted independently by two differentially expressed microRNAs. The differential expression pattern of each targeting microRNA correlates with altered chemotaxis behaviour, suggesting that microRNAs may also drive behavioural variation (Table 1).

The predicted targeting of *Sc-npr-11* by *sc-miR-301-5pi* reveals that UTR sequence variation between the two annotated *npr-11* isoforms allows the dominant isoform (representing ~ 75% of all *npr-11* transcript across strains) to escape *mir-301-5pi* interaction (Table 1; Additional file 9). Likewise, variation in the 5'UTR and 3'UTRs of *Sc-vglu-2* and *Sc-cha-1* (Figs. 6 and 7, Table 2) reveal a potential quantitative difference in microRNA interaction events across different gene isoforms. In the case of *Sc-cha-1*, gene-level regulation of microRNA target site visibility could represent a strain specific mechanism to dampen global cholinergic signalling, or could allow selective dampening in a subset of cholinergic neurons. Ascertaining the relevance of isoform variance as it relates to microRNA interactions will require detailed gene localisation and functional studies. Our datasets highlight 5' and 3'UTR variation as a factor in differential microRNA target site visibility, although it seems especially likely that alternative polyadenylation signals in the 3'UTR of protein-coding genes will exert substantial control over microRNA target site visibility, especially given the evident pervasiveness of 3'UTR variation in nematodes [63, 64]. Any such regulation could modify the relative percentage of gene isoform

variants accessible to microRNAs, which could allow for cell, or tissue-specific regulatory events that may be difficult to ascertain from organism-wide datasets. Collectively, our data point to a complex co-regulatory environment involving gene-specific isoform variation, and microRNA transcriptional regulation that is likely to influence various aspects of biology, including behaviour.

In the same way that “the best model for a cat is several cats” [65], the best model for a nematode parasite of vertebrates, is a sustainable population of the very same. However, this might not always be preferable in terms of ethics, available tools, sustainability or reproducibility. ‘Model’ status for any organism is contingent on high quality genomic and transcriptomic resources, easily amenable research tools, both in terms of genetic and molecular manipulation, alongside robust behavioural and phenotypic assays. Ease of culture, handling, and short generation times should also be major considerations. A model organism and the data generated from it must also be sufficiently relevant to trigger near-term impact on other species that have implications for health and economy. The datasets presented in this study build upon a growing catalogue of tools and resources for *Steinernema* spp. entomopathogenic nematodes [9, 21, 30, 49, 65–67]. The close phylogenetic relationship between *Steinernema* spp. and economically important mammalian parasites [68], coupled with striking behavioural similarities, and a pathogenic life style that can be fully recapitulated in a Petri dish suggests that these entomopathogenic nematodes could represent attractive and convenient new surrogate models for parasite biology and behaviour. In addition to their potential worth as model organisms, *Steinernema* spp. have attracted considerable attention as bioinsecticidal agents [69]. The interest in *Steinernema* spp. as biological models, and as economically relevant end-point organisms in their own right represents a unique proposition for researchers interested in comparative biology, behaviour and host-parasite interactions.

Conclusions

This study is the first to identify differentially expressed genes that correlate with behavioural variation across three *S. carpocapsae* strains. We have identified four *srsx* GPCR genes and a sodium transporter gene that correlate with enhanced chemotaxis of *S. carpocapsae* IJs. A neuropeptide-like protein gene, *srt* GPCR gene and sodium transporter gene are found to correlate with reduced nictation behaviour. Numerous microRNAs are predicted to target neuronal genes, and display expression profiles that correlate with behavioural variation. These data provide important insight to the transcriptional basis of host-finding behaviours in this species, highlighting a potential role for microRNAs and gene isoform variation.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-019-6179-y>.

Additional file 1. List of all differentially expressed genes (> 2 log₂ fold, $P < 0.0001$) across all pairwise strain comparisons.

Additional file 2. List of all differentially expressed genes represented in heatmaps, including annotated gene names, fold changes across pairwise comparisons, and adjusted P values.

Additional file 3. List of *S. carpocapsae* microRNAs, sequences and read counts across strains.

Additional file 4. List of differentially expressed microRNAs (> 2 log₂ fold, $P < 0.0001$), including pairwise fold changes, and adjusted P values.

Additional file 5. MicroRNA target predictions 1: Miranda output file for microRNA targets in 5'UTRs using default setting.

Additional file 6. MicroRNA target predictions 2: Miranda output file for microRNA targets in 5'UTRs using strict setting.

Additional file 7. MicroRNA target predictions 3: Miranda output file for microRNA targets in 3'UTRs using default setting.

Additional file 8. MicroRNA target predictions 4: Miranda output file for microRNA targets in 3'UTRs using strict setting.

Additional file 9. RSEM isoform quantification file for *S. carpocapsae* strains.

Abbreviations

FLP: FMRFamide-Like Peptide; GPCR: G-Protein Coupled Receptor; IJ: Infective Juvenile; INS: INSulin-like peptide; NLP: Neuropeptide-Like Protein; UTR: UnTranslated Region

Acknowledgements

Waxworms infected with *S. carpocapsae* (Breton), were a gift from Prof Nelson Simões, University of the Azores, Portugal. *S. carpocapsae* (UK1) IJs were a gift from BASF. Microdirt chip PDMS moulds were a gift from Prof Junho Lee, Seoul National University, Korea. The authors would like to thank Matthew Sturrock and Steven Dyer for critical reading of the manuscript.

Authors' contributions

NDW, DC, CM and RM generated data and wrote the manuscript. JJD conceived the project, and wrote the manuscript. All authors read and approved the final manuscript.

Funding

NDW was funded by the Bill and Melinda Gates Foundation, DC was funded by a GCRF pilot grant from the NI Department for the Economy, CMCc was funded by a proof of concept grant from Invest NI, RM was funded by a PhD studentship from the QUB Business Alliance Fund. The funders had no role in experimental design or preparation of this manuscript. All authors have read and approved this manuscript.

Availability of data and materials

The *S. carpocapsae* strains used in this study are maintained within the Dalzell lab, and are freely available upon request. Illumina HiSeq reads for transcriptome and small RNAs across each *S. carpocapsae* strain are available within the SRA library under project code: PRJNA436466.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 28 November 2018 Accepted: 10 October 2019

Published online: 21 November 2019

References

- Spencer SG, Hallem EA. Mechanisms of host seeking by parasitic nematodes. *Mol Biochem Parasitol.* 2016;208(1):23–32. <https://doi.org/10.1016/j.molbiopara.2016.05.007>.
- Castelletto ML, Gang SS, Okubo RP, Tselikova AA, Nolan TJ, Platzer EG, et al. Diverse host-seeking behaviors of skin-penetrating nematodes. *PLoS Pathog.* 2014;10(8):e1004305. <https://doi.org/10.1371/journal.ppat.1004305>.
- Chaisson KE, Hallem EA. Chemosensory behaviors of parasites. *Trends Parasitol.* 2012;28(10):427–36. <https://doi.org/10.1016/j.pt.2012.07.004>.
- Dillman AR, Guillermin ML, Lee JH, Kim B, Sternberg PW, Hallem EA. Olfaction shapes host–parasite interactions in parasitic nematodes. *PNAS.* 2012;109(35):E2324–E33. <https://doi.org/10.1073/pnas.1211436109>.
- Ruiz F, Gang SS, Castelletto ML, Hallem EA. Experience-dependent olfactory behaviors of the parasitic nematode *Heligmosomoides polygyrus*. *PLoS Pathog.* 2017;13(11):e1006709. <https://doi.org/10.1371/journal.ppat.1006709>.
- Mousley A, McVeigh P, Dalzell JJ, Maule AG. In: Kennedy MW, Harnett W, editors. ISBN: 9781845937591 Nematode neuropeptide communication systems. In, parasitic nematodes: molecular biology, biochemistry and immunology: CAB International; 2013. <https://doi.org/10.1079/9781845937591.0279>.
- Peymen K, Watteyne J, Frooninckx L, Schoofs L, Beets I. The FMRFamide-like peptide family in nematodes. *Front Endocrinol.* 2014;5:90. <https://doi.org/10.3389/fendo.2014.00090>.
- Lee JS, Shih PY, Schaedel ON, Quintero-Cadena P, Rogers AK, Sternberg PW. FMRF amide-like peptides expand the behavioural repertoire of a densely connected nervous system. *PNAS.* 2017;114(50):E10726–35. <https://doi.org/10.1073/pnas.1710374114>.
- Morris R, Wilson L, Sturrock M, Warnock ND, Carrizo D, Cox D, et al. A neuropeptide modulates sensory perception in the entomopathogenic nematode *Steinernema carpocapsae*. *PLoS Pathog.* 2017;13(3):e1006185. <https://doi.org/10.1371/journal.ppat.1006185>.
- Lee D, Lee H, Kim N, Lim DS, Lee J. Regulation of a hitchhiking behavior by neuronal insulin and TGF- β signaling in the nematode *Caenorhabditis elegans*. *Biochem Biophys Res Commun.* 2017;484(2):323–30. <https://doi.org/10.1016/j.bbrc.2017.01.113>.
- Warnock ND, Wilson L, Patten C, Fleming CC, Maule AG, Dalzell JJ. Nematode neuropeptides as transgenic nematocides. *PLoS Pathog.* 2017; 13(2):e1006237. <https://doi.org/10.1371/journal.ppat.1006237>.
- Hobert O. The neuronal genome of *Caenorhabditis elegans*. *WormBook: The C. elegans Research Community, WormBook;* 2013. <https://doi.org/10.1895/wormbook.1.161.1>.
- Jovelin R, Cutter AD. Microevolution of nematode miRNAs reveals diverse modes of selection. *Genome Biol Evol.* 2014;6(11):3049–63. <https://doi.org/10.1093/gbe/evu239>.
- Heimberg AM, Sempere LF, Moy VN, Donoghue PC, Peterson KJ. MicroRNAs and the advent of vertebrate morphological complexity. *PNAS.* 2008;105: 2946–50. <https://doi.org/10.1073/pnas.0712259105>.
- Iwama H, Kato K, Imachi H, Murao K, Masaki T. Human microRNAs originated from two periods at accelerated rates in mammalian evolution. *Mol Biol Evol.* 2013;30:613–26. <https://doi.org/10.1093/molbev/mss262>.
- Meunier J, Lemoine F, Soumillon M, Liechti A, Guschanski K, Hu H, Khaitovich P, Kaessmann H. Birth and expression evolution of mammalian microRNA genes. *Genome Res.* 2013;23(1):34–45. <https://doi.org/10.1101/gr.140269>.
- Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell.* 1993; 75(5):843–54 DOI not available.
- Konopka W, Kiryk A, Novak M, Herwerth M, Parkitna JR, et al. MicroRNA loss enhances learning and memory in mice. *J Neurosci.* 2010;30(44):14835–42. <https://doi.org/10.1523/JNEUROSCI.3030-10.2010>.
- Ashraf SI, McLoon AL, Sclarsic SM, Kunes S. Synaptic protein synthesis associated with memory is regulated by the RISC pathway in *Drosophila*. *Cell.* 2006;124(1):191–205. <https://doi.org/10.1016/j.cell.2005.12.017>.
- Than MT, Kudlow BA, Han M. Functional analysis of neuronal MicroRNAs in *Caenorhabditis elegans* Dauer formation by combinational genetics and neuronal miRISC Immunoprecipitation. *PLoS Genet.* 2013;9(6):e1003592. <https://doi.org/10.1371/journal.pgen.1003592>.

21. Lu D, Macchietto M, Chang D, Barros MM, Baldwin J, Mortazavi A, Dillman AR. Activated entomopathogenic nematode infective juveniles release lethal venom proteins. *PLoS Pathog.* 2017;13(4):31006302. <https://doi.org/10.1371/journal.ppat.1006302>.
22. Spence KO, Lewis EE, Pery RN. Host-finding and invasion by entomopathogenic and plant-parasitic nematodes: evaluating the ability of laboratory bioassays to predict field results. *J Nematol.* 2008;40(2):93–8 DOI not available.
23. Campbell JF, Kaya HK. How and why a parasitic nematode jumps. *Nature.* 1999;397:485–6. <https://doi.org/10.1038/17254>.
24. Lee H, Choi MK, Lee D, Kim HS, Hwang H, Kim H, Park S, Paik YK, Lee J. Nictation, a dispersal behavior of the nematode *Caenorhabditis elegans*, is regulated by IL2 neurons. *Nat Neurosci.* 2011;15(1):107–12. <https://doi.org/10.1038/nn.2975>.
25. White G. A method for obtaining infective nematode larvae from cultures. *Am Assoc Adv Sci.* 1927;66(1709):302–3. <https://doi.org/10.1126/science.66.1709.302-a>.
26. Lee D, Lee H, Choi M-k, Park S, Lee J. Nictation assays for *Caenorhabditis* and other nematodes. *Bioprotocol.* 2015;5:e1433. <https://doi.org/10.21769/BioProtoc.1433>.
27. Andrew, S. FastQC: a quality control tool for high throughput sequence data, 2010; <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
28. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumine sequence data. *Bioinformatics.* 2014;30(15):2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
29. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, et al. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010;28(5):511–5. <https://doi.org/10.1038/nbt.1621>.
30. Dillman AR, Macchietto M, Porter CF, Rogers A, Williams B, Antoshechkin I, et al. Comparative genomics of *Steinernema* reveals deeply conserved gene regulatory networks. *Genome Biol.* 2015;16:200. <https://doi.org/10.1186/s13059-015-0746-6>.
31. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
32. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011;12(1):323. <https://doi.org/10.1186/1471-2105-12-323>.
33. Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BM, Haag JD, Gould MN, Stewart RM, Kendziorski C. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics.* 2013;29(8):1035–43. <https://doi.org/10.1093/bioinformatics/btt087>.
34. Team RDC. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2008. ISBN 3–900051–07–0, URL <http://www.R-project.org>
35. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol.* 2014;15(12):550. <https://doi.org/10.1186/s13059-014-0550-8>.
36. Neuwirth E. RColorBrewer: ColorBrewer Palettes. R package version 1.1–2, 2014; <https://CRAN.R-project.org/package=RColorBrewer>
37. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W et al. gplots: Various R Programming Tools for Plotting Data. R package version 3.0.1, 2016; <http://CRAN.R-project.org/package=gplots>
38. Gentleman R, Biocore. Geneploater: graphics related functions for Bioconductor, 2016; R package version 1.52.0.
39. Kolde R. pheatmap: Pretty Heatmaps. R package version 1.0.8, 2015; <http://CRAN.R-project.org/package=pheatmap>.
40. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 2011;17(1):10–2. <https://doi.org/10.14806/ej.17.1.200>.
41. Friedländer MR, Mackowiak SD, Li N, Chen W, Rajewsky N. MiRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 2012;40(1):37–52. <https://doi.org/10.1093/nar/gkr688>.
42. Howe KL, Bolt BJ, Cain S, Chan J, Chen WJ, Davis P, et al. WormBase 2016: expanding to enable helminth genomic research. *Nucleic Acids Res.* 2015; 1217. <https://doi.org/10.1093/nar/gkv1217>.
43. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in *Drosophila*. *Genome Biol.* 2003;5(1):R1. <https://doi.org/10.1186/gb-2003-5-1-r1>.
44. Thomas M, Lieberman J, Lal A. Desperately seeking microRNA targets. *Nat Struct Mol Biol.* 2010;17:1169–74. <https://doi.org/10.1038/nsmb.1921>.
45. Zisoulis DG, Lovci MT, Wilbert ML, Hutt KR, Liang TY, et al. Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat Struct Mol Biol.* 2010;17:173–9. <https://doi.org/10.1038/nsmb.1745>.
46. Chen N, Harris TW, Antoshechkin I, Bastiani C, Bieri T, et al. WormBase: a comprehensive data resource for *Caenorhabditis* biology and genomics. *Nucleic Acids Res.* 2005;33:D383–9. <https://doi.org/10.1093/nar/gki066>.
47. Dalzell JJ, McVeigh P, Warnock ND, Mitreva M, Bird DM, Abad P, et al. RNAi effector diversity in nematodes. *PLoS Negl Trop Dis.* 2011;5(6):e1176. <https://doi.org/10.1371/journal.pntd.0001176>.
48. McVeigh P, Kimber MJ, Novozhilova E, Day TA. Neuropeptide signalling systems in flatworms. *Parasitology.* 2005;131:S41–55. <https://doi.org/10.1017/S0031182005008851>.
49. Rougon-Cardoso A, Flores-Ponce M, Ramos-Aboites HE, Martinez-Guerrero CE, Hao Y-J, et al. The genome, transcriptome, and proteome of the nematode *Steinernema carpocapsae*: evolutionary signatures of a pathogenic lifestyle. *Sci Rep.* 2016;6:37536. <https://doi.org/10.1038/srep37536>.
50. Cook DE, Zdravljec S, Roberts JP, Adersen EC. CeNDR, the *Caenorhabditis elegans* natural diversity resource. *Nucleic Acids Res.* 2017;45(D1):D650–7. <https://doi.org/10.1093/nar/gkw893>.
51. Lee D, Yang H, Kim J, Brady S, Zdravljec S, Zamanian M, Kim H, Paik YK, Kruglyak L, Andersen EC, Lee J. The genetic basis of natural variation in a phoretic behaviour. *Nat Commun.* 2017;8(1):273. <https://doi.org/10.1038/s41467-017-00386-x>.
52. Coburn CM, Bargman CI. A putative cyclic nucleotide-gated channel is required for sensory development and function in *C. elegans*. *Neuron.* 1996; 17:695–706. [https://doi.org/10.1016/S0896-6273\(00\)80201-9](https://doi.org/10.1016/S0896-6273(00)80201-9).
53. Coburn CM, Mori I, Ohshima Y, Bargmann CI. A cyclic nucleotide-gated channel inhibits sensory axon outgrowth in larval and adult *Caenorhabditis elegans*: a distinct pathway for maintenance of sensory axon structure. *Development.* 1998;125:249–58 No doi available.
54. Gruner M, Nelson D, Winbush A, Hintz R, Ryu L, Chung SH, Kim K, Gabel CV, van der Linden AM. Feeding state, insulin and NPR-1 modulate chemoreceptor gene expression via integration of sensory and circuit inputs. *PLoS Genet.* 2014;10(10):e1004707. <https://doi.org/10.1371/journal.pgen.1004707>.
55. Wang XD, Ishibashi N. Infection of the entomopathogenic nematode, *Steinernema carpocapsae*, as affected by the presence of *Steinernema glaseri*. *J Nematol.* 1999;31(2):207–11 DOI not available.
56. Lee JH, Dillman AR, Hallem EA. Temperature-dependent changes in the host-seeking behaviors of parasitic nematodes. *BMC Biol.* 2016;14(1):1. <https://doi.org/10.1186/s12915-016-0259-0>.
57. Okumura E, Yoshiga T. Host orientation using volatiles in the phoretic nematode *Caenorhabditis japonica*. *J Exp Biol.* 2014;217(18):3197–9. <https://doi.org/10.1242/jeb.105353>.
58. Yemini EI, Jucikas T, Grundy LJ, Brown AE, Schafer WR. A database of *Caenorhabditis elegans* behavioural phenotypes. *Nat Methods.* 2013;10:877–9. <https://doi.org/10.1038/nmeth.2560>.
59. Gillan V, Maitland K, Laing R, Gu H, Marks ND, et al. Increased expression of a microRNA correlates with anthelmintic resistance in parasitic nematodes. *Front Cell Infect Microbiol.* 2017. <https://doi.org/10.3389/fcimb.2017.00452>.
60. Gu HY, Marks ND, Winter AD, Weir W, Tzelos T, McNeilly TN, Britton C, Devaney E. Conservation of a microRNA cluster in parasitic nematodes and profiling of miRNAs in excretory-secretory products and microvesicles of *Haemonchus contortus*. *PLoS Negl Trop Dis.* 2017;11(11):e0006056. <https://doi.org/10.1371/journal.pntd.0006056>.
61. Zamanian M, Cook DE, Zdravljec S, Brady SC, Lee D, Lee J. Discovery of unique loci that underlie nematode responses to benzimidazoles. *BioRxiv.* 2017. <https://doi.org/10.1101/116970>.
62. Pinzón N, Blaise L, Martinez L, Sergeeva A, Presumev J, Apparailly F, Seitz H. MicroRNA target prediction programs predict many false positives. *Genome Res.* 2017;27(2):234–45. <https://doi.org/10.1101/gr.205146.116>.
63. Mangone M, Manoharan AP, Thierry-Mieg D, Thierry-Mieg J, Han T, et al. The landscape of *C. elegans* 3'UTRs. *Science.* 2010;329(5990):432–5. <https://doi.org/10.1126/science.1191244>.
64. Blazie SM, Geissel HC, Wilky H, Joshi R, Newbern J, Mangone M. Alternative polyadenylation directs tissue-specific miRNA targeting in *Caenorhabditis elegans* somatic tissues. *Genetics.* 2017;206(2):757–74. <https://doi.org/10.1534/genetics.116.196774>.
65. Gustafsson C, Vallverdu J. The best model of a cat is several cats. *Trends Biotechnol.* 2016;34(3):207–13. <https://doi.org/10.1016/j.tibtech.2015.12.006>.
66. Macchietto M, Angdemby D, Heidarpour N, Serra L, Rodriguez B, El-Alli N, Mortazavi A. Comparative transcriptomics of *Steinernema* and *Caenorhabditis*

single embryos reveals orthologous gene expression convergence during late embryogenesis. *Genome Biol Evol.* 2017;9(10):2681–96. <https://doi.org/10.1093/gbe/evx195>.

67. Choe A, von Reuss SH, Kogan D, Gasser RB, Platzer EG, Schroeder FC, Sternberg PW. Ascaroside signalling is widely conserved among nematodes. *Curr Biol.* 2012;22(9):772–80. <https://doi.org/10.1016/j.cub.2012.03.024>.
68. Blaxter M, Koutsovoulos G. The evolution of parasitism in Nematoda. *Parasitology.* 2015;142(Suppl 1):S26–39. <https://doi.org/10.1017/S0031182014000791>.
69. Lu D, Baiocchi T, Dillman AR. Genomics of entomopathogenic nematodes and implications for pest control. *Trends Parasitol.* 2016;32(8):588–98. <https://doi.org/10.1016/j.pt.2016.04.008>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

