Computational and Structural Biotechnology Journal 23 (2024) 1026–1035

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Analysis of *Haemonchus* embryos at single cell resolution identifies two eukaryotic elongation factors as intervention target candidates

Pasi K. Korhonen, Tao Wang, Neil D. Young, Joseph J. Byrne, Tulio L. Campos, Bill C.H. Chang, Aya C. Taki, Robin B. Gasser *

Department of Veterinary Biosciences, Melbourne Veterinary School, Faculty of Science, The University of Melbourne, Parkville, Victoria 3010, Australia

ARTICLE INFO

Keywords:
Caenorhabditis elegans
Single cell/nucleus RNA-sequencing
Cell atlas
Haemonchus contortus
Eggs during embryonation
Essential gene
Interaction network
Eukaryotic elongation factor
Protein biosynthesis
Intervention target candidate

ABSTRACT

Advances in single cell technologies are allowing investigations of a wide range of biological processes and pathways in animals, such as the multicellular model organism *Caenorhabditis elegans* – a free-living nematode. However, there has been limited application of such technology to related parasitic nematodes which cause major diseases of humans and animals worldwide. With no vaccines against the vast majority of parasitic nematodes and treatment failures due to drug resistance or inefficacy, new intervention targets are urgently needed, preferably informed by a deep understanding of these nematodes' cellular and molecular biology – which is presently lacking for most worms. Here, we created the first single cell atlas for an early developmental stage of *Haemonchus contortus* – a highly pathogenic, *C. elegans*-related parasitic nematode. We obtained and curated RNA sequence (snRNA-seq) data from single nuclei from embryonating eggs of *H. contortus* (150,000 droplets), and selected high-quality transcriptomic data for > 14,000 single nuclei for analysis, and identified 19 distinct clusters of cells. Guided by comparative analyses with *C. elegans*, we were able to reproducibly assign seven cell clusters to body wall muscle, hypodermis, neuronal, intestinal or seam cells, and identified eight genes that were transcribed in all cell clusters/types, three of which were inferred to be essential in *H. contortus*. Two of these genes (i.e. *Hc-eef-1A* and *Hc-eef1G*), coding for eukaryotic elongation factors (called *Hc*-eEF1A and *Hc*-eEF1G), were also demonstrated to be transcribed and expressed in all key developmental stages of *H. contortus.* Together with these findings, sequence- and structure-based comparative analyses indicated the potential of *Hc*-eEF1A and/or *Hc*-eEF1G as intervention targets within the protein biosynthesis machinery of *H. contortus*. Future work will focus on single cell studies of all key developmental stages and tissues of *H. contortus*, and on evaluating the suitability of the two elongation factor proteins as drug targets in *H. contortus* and related nematodes, with a view to finding new nematocidal drug candidates.

1. Introduction

Since the 1960s, the free-living roundworm (= nematode) *Caenorhabditis elegans* has been used as a model organism [1–3] to extensively study many aspects of the biology of multicellular organisms generally (https://www.wormatlas.org/; http://www.wormbook.org/). This soil nematode has offered major benefits as a model, particularly because of its short (3-day) life cycle, small size (1.5 mm), mode of reproduction (selfing), and relative ease of cultivation and 'manipulation' in laboratory. Other key features of this nematode are its anatomical simplicity (~ 1000 cells), including the 302-cell hermaphrodite nervous system, and its small genome (100 Mb). Complemented by the advent of cutting-edge genomic, biochemical, transcriptomic, proteomic, informatic and imaging tools, the sequencing and subsequent annotation of the *C. elegans* genome and availability of extensive web-based resources, such as WormBase [4,5], have accelerated fundamental explorations using this model organism. Some of latest technologies have revolutionised discovery, and allowed the deconvolution of biological processes and pathways to the tissue and single-cell levels (e.g., [6–8]). In particular, the application of single cell (sc) and/or single nucleus (sn) RNA-sequencing technologies have allowed the creation of single cell atlases for *C. elegans*, providing unprecedented, 'granular' insights into aspects such as embryogenesis, cell differentiation, types, fate and plasticity, and neurobiology [9–13], and also providing a solid foundation for comparative studies of the cellular and molecular biology of related organisms (e.g., [14–16]).

* Corresponding author.
E-mail address: robinbg@unimelb.edu.au (R.B. Gasser).

https://doi.org/10.1016/j.csbj.2024.01.008
Received 15 October 2023; Received in revised form 14 January 2024; Accepted 15 January 2024
Available online 17 January 2024

Caenorhabditis elegans is in the same evolutionary clade (V) as one of the largest groups of socioeconomically important parasitic nematodes of humans and animals, including the pathogenic, blood-feeding barber's pole worm – *Haemonchus contortus* [17]. Consequently, much molecular research of *H. contortus* has relied heavily on knowledge, data sets and information available for *C. elegans* [18–20]. Particularly over the past decade, intensive genomic, transcriptomic and proteomic investigations have enabled the creation of substantially improved resources and tools for *H. contortus* – which has now been elevated to model organism status [21,22].

Despite these advances, there are still major knowledge gaps in many fundamental areas of *H. contortus* anatomy, biology and development at the cellular level, limiting our understanding of this parasite and parasitism. As a first step to gaining insights into *H. contortus* at this level, we conducted the first single-nucleus RNA-seq analysis of *H. contortus* at the stage of embryonation, as a proof-of-principle, to infer essential genes based on comparative analysis with *C. elegans*, as a foundation for detailed molecular investigations of the life history of this and related parasitic nematodes at the cellular level in the future.

2. Materials and methods

2.1. Production and purification of embryonated Haemonchus contortus eggs

With animal ethics approval (no. 1714374) from the University of Melbourne, a parasite-free merino lamb (*Ovis aries*; male; 6 months of age) was inoculated orally with 8000 infective third-stage larvae (L3s) of *H. contortus* (Haecon 5 strain, Australia) as described previously [23]. One month after inoculation, faeces were collected for 1 h and processed. From 100 g of faeces from an infected sheep, *H. contortus* eggs (n = 200,000) were purified by sucrose flotation [24] at 24 °C within 4 h of collection. By this time, the majority (95%) of eggs were at various stages of embryonation – consistent with description by Veglia [25].

2.2. Isolation of nuclei

Subsequently, cell nuclei were isolated from these embryonated eggs using a modification of the Frankenstein protocol [26] In brief, 30,000 embryonated eggs were transferred to a 1.5 ml low-DNA binding Eppendorf tube containing 200 μl of ice-cold nuclei lysis buffer (Merck Millipore, USA), and homogenised with a pellet pestle (Sigma-Aldrich, USA; 20 back-and-forth rotations). The homogenate was then transferred to 2.0 ml low-DNA binding Eppendorf tube, after which 1.3 ml of ice-cold nuclei lysis buffer were added to the tube and incubated on ice for 5 min (gently mixing with a wide-bore 1 ml tip twice during the incubation). Following the incubation, the homogenate was filtered through a 40 μm nylon mesh sieve (Thomas Scientific, USA), and then centrifuged at 500 × *g* and 4°C for 5 min. The supernatant was aspirated and discarded, leaving 70 μl in the tube. The pellet of cell nuclei was resuspended in 1.5 ml of ice-cold nucleus lysis buffer and incubated on ice for 5 min. This homogenate was centrifuged at 500 × *g* and 4°C for 5 min and resuspended in 1.5 ml of ice-cold nuclei wash buffer (phosphate-buffered saline [pH 7.0] containing 1% bovine serum albumin [BSA] and 0.2 U/μl RNase inhibitor); this step was repeated once and the supernatant removed. Subsequently, the pellet of nuclei was resuspended in 500 μl of ice-cold nuclei wash buffer with 10 μg/ml DAPI (4, 6-diamidino-2-phenylindole; Sigma-Aldrich, USA) and an aliquot (5 μl) thereof examined at 300-times magnification using a fluorescence microscope (ECLIPSE Ti2 Nikon, Japan) to appraise the shape and size of the cell nuclei.

2.3. Purification of nuclei, library construction, sequencing and preprocessing of data

First, fluorescence-activated cell sorting (FACS) was used to separate nuclei from cell debris. Then, the nuclei were then encapsulated into droplets, cDNA libraries constructed using the standard 10x Genomics (Chromium) protocol, and sequencing (5'-end of mRNA) was conducted using the Illumina HiSeq 2500 system. Raw sequence data were assessed for quality, mapped against the reference genome of *H. contortus* (version WBPS17; ref. [21] and transcript-count matrix created using Cell Ranger v7.1.0 software [27]. Read counts in the matrix were processed further using the programs R v4.0.3 [28] and Seurat v3.9.9 [29]; genes with a transcript read-count of < 3 and cells with a gene count of < 50 were removed; data from droplets containing ≥ 2 nuclei were eliminated using the program scds v1.6.0 [30]; data from droplets with transcripts inferred to be 'background' were eliminated based on an analysis of 2000 genes with the most variable transcription levels among all cells using the program DIEM v2.4.1 [31]; only cells containing ≥ 200 transcribed genes were retained.

2.4. Data normalisation, clustering, annotation of cell types and other bioinformatic analyses

The curated snRNA-seq read counts obtained were normalised using the method SCTransform [32] and clustered using the Seurat software package. Marker-genes for individual clusters were identified based on their transcription profile employing the program Scran v1.18.7 [33]. Clusters were illustrated using the programs Seurat and UMAP [34,35]. Cell clusters were inferred based on a comparative analysis of transcription profiles obtained from a single cell data set (GSE126954) for *C. elegans* [10]. In short, orthologous genes were identified by reciprocal BlastP analysis of the proteomes of *C. elegans* (version WBcel235; NCBI) and *H. contortus* (version WBP17; WormBase). In the *C. elegans* dataset, cell types were identified from column "cell_type1" using the program SingleCellExperiment [36], and, using exclusively orthologous genes, cell-types were inferred for the *H. contortus* data set using the program clustifyr v1.2.0 [37]. This approach was taken because cell types have not yet been defined in *H. contortus*. In addition, we performed gene ontology (GO) enrichments for the genes linked to individual cell-type clusters using the program topGO v2.42.0 [38], in which a weighted Fisher's exact test is applied.

For individual genes transcribed at high levels in all cell clusters, genetic, physical, regulatory and/or predicted interaction networks, and their essentiality were explored with WormBase. The essentiality of genes was inferred based on RNAi-based gene knockdown in *C. elegans* (see WormBase) and transcription in all clusters (cf. [39]); essentiality was linked to an embryonic lethal (Emb) or larval lethal (Lvl) phenotype. Interaction networks and expression/transcription information were accessed from WormBase (15 March 2023). This approach was taken, because no large-scale functional genomic data set is available for *H. contortus* and RNAi does not always work reproducibly or effectively in this species [40], likely due to the absence of some components (i.e. RDE-2 or SID-2) of the RNAi machinery [21,23].

Transcription levels of RNA-seq data represent distinct developmental stages/sexes (i.e. egg, L1, L2, L3, L4 female, L4 male, adult female and adult male) were extracted from previously published data (ref. [23]; SRP026668). Sequence reads were cleaned and trimmed using the software Trimmomatic v0.39 [41] and curated reads were mapped to all coding sequences of *H. contortus* (ref. [42]; WormBase ParaSite version 18) using the program RSEM v1.3.3 [43]. Resultant transcripts per million (TPM) were then used to indicate the transcription level of essential genes in the distinct developmental stages of

Table 1
Nineteen cell clusters inferred based on the analysis of single nucleus (sn)RNA-sequence data from *Haemonchus contortus* embryos, annotatable cell types and respective marker genes (cf. Fig. 1).

Cluster number	Number of individual nuclei	Cell type (cluster annotation)	Marker genes (top-five; $p < 0.001$)
1	2071	-	-
2	1591	Neurons	HCON_00176790
3	1425	-	HCON_00176310; HCON_00173800; HCON_00173380; HCON_00177780 HCON_00176280
4	1256	-	HCON_00094700
5	1068	Body wall muscle	HCON_00181850; HCON_00174260; HCON_00173460; HCON_00013360 HCON_00176775
6	784	Hypodermis	HCON_00036640; HCON_00144180; HCON_00092510; HCON_00032040 HCON_00075270
7	724	-	HCON_00089225; HCON_00176775; HCON_00186780; HCON_00091470 HCON_00050940
8	692	-	HCON_00160300; HCON_00127710
9	689	-	HCON_00096110; HCON_00034760; HCON_00013670; HCON_00049570 HCON_00078730
10	617	-	HCON_00185590; HCON_00103410; HCON_00056730; HCON_00050260 HCON_00001240
11	557	-	HCON_00190170; HCON_00190180; HCON_00027070; HCON_00150565 HCON_00104060
12	508	Neurons	HCON_00179900; HCON_00009880; HCON_00164070; HCON_00115910 HCON_00172500
13	422	-	HCON_00023690; HCON_00061920; HCON_00067190; HCON_00108700 HCON_00099280
14	376	Neurons	HCON_00017720; HCON_00017700
15	322	Intestine	HCON_00178645; HCON_00160690; HCON_00073380; HCON_00107430 HCON_00148120
16	317	Seam cells	HCON_00143140; HCON_00187800; HCON_00089910; HCON_00187910 HCON_00187890
17	287	-	-
18	273	-	HCON_00176800; HCON_00091128; HCON_00091125; HCON_00036620 HCON_00172840
19	149	-	HCON_00190170

H. contortus.

The structures of orthologous proteins encoded by selected genes were predicted using the program AlphaFold2 [44], and the confidence of domains, regions and/or overall structures assessed using the predicted local distance difference test (pLDDT) values, with the per-residue estimate of confidence given on a scale from 0–100. Evolutionary conservation was assessed at each amino acid position in a multiple sequence alignment of orthologous proteins using the program Capra07 [45]; the conservation score represented the similarity between two probability distributions (Jensen–Shannon divergence). For the phylogenetic analysis, orthologous protein sequences were selected using the program OrthoMCL v2.0.4 [46], aligned using the program MAFFT v7.407 [47] and the aligned sequences subjected to separate analyses using the programs RAxML v8.2.9 [48] and MrBayes v3.2.6 [49], in which the LG amino acid substitution model (with gamma-distribution, G, and proportion of invariable sites, F; ref. [50]) was employed.

3. Results

3.1. Single cell atlas based on an analysis of snRNA-seq data

Raw snRNA-seq data (44.6 Gb) and read-index data (5.0 Gb) (Table S1) were transferred into a droplet × feature (gene) read-count matrix. Of the sequence reads in the 150,000 droplets selected, 87.1% mapped to the *H. contortus* genome (including introns), and 64.2% to gene regions including introns (version WBPS17); 96.5% of the reads were within droplets; the median number of genes per droplet was 280; in total, 17,101 genes were represented (Table S2). After the removal of data derived from droplets with multiple or no nuclei, the remaining data from 14,128 droplets (each containing one nucleus) and 13,687 genes that were transcribed in ≥ 3 nuclei, were subjected to further analysis. Using transcripts from individual nuclei (Table 1; Fig. 1a), we inferred 19 distinct cell clusters (Table S3), 17 of which had marker genes ($p < 0.001$) (Table 1 and S4). We inferred orthologous proteins (n = 7913) between the *H. contortus* and *C. elegans* proteomes, and used 6797 of these orthologs to annotate the clusters (Table S5). Guided by *C. elegans* single cell data (GSE126954), we were able to reproducibly assign seven individual cell clusters to body wall muscle (cluster no. 5), hypodermis (6), neurons (2, 12, and 14), intestine (15) or seam cells (16) (Table 1; Fig. 1a).

GO enrichments for genes linked to the 19 clusters inferred for *H. contortus* revealed (i) molecular functions including transcription, translation, regulation and binding activities; (ii) cellular processes including metabolism, RNA splicing, energy and transport; and (iii) cellular compartments such as nucleus, organelle, protein and complexes in most clusters, but no identifiable cluster-specific enrichment(s) (Tables S6, S7 and S8).

3.2. Analysis of genes transcribed among cell clusters reveals a subset of genes/proteins with key roles in the mitochondrion, RNA biology and protein synthesis

Subsequently, we enumerated and explored transcribed genes that were common among ≥ 3 or all 19 clusters. Represented in ≥ 3 clusters were 1703 *H. contortus* genes (including 229 orphans), 1132 of which had known (n = 889) or orphan (n = 243) orthologs in *C. elegans* (Table S9). Represented in all 19 clusters were eight *H. contortus* genes encoding proteins matching *C. elegans* homologs (Table 2 and S10): an ADP, ATP carrier protein encoded by *Hc-ant-1* (HCON_00082980); five RNA recognition motif (RRM) domain-containing protein homologs represented by *Hc-rbm-3* (HCON_00018960, HCON_00030820, HCON_00030825, HCON_00106740 and HCON_00133560); a eukaryotic elongation factor 1-alpha encoded by *Hc-eef-1A* (HCON_00071960); and a eukaryotic elongation factor 1-gamma encoded by *Hc-eef-1 G* (HCON_00147420). Three of these genes – *Hc-ant-1*, *Hc-eef-1A* and *Hc-*

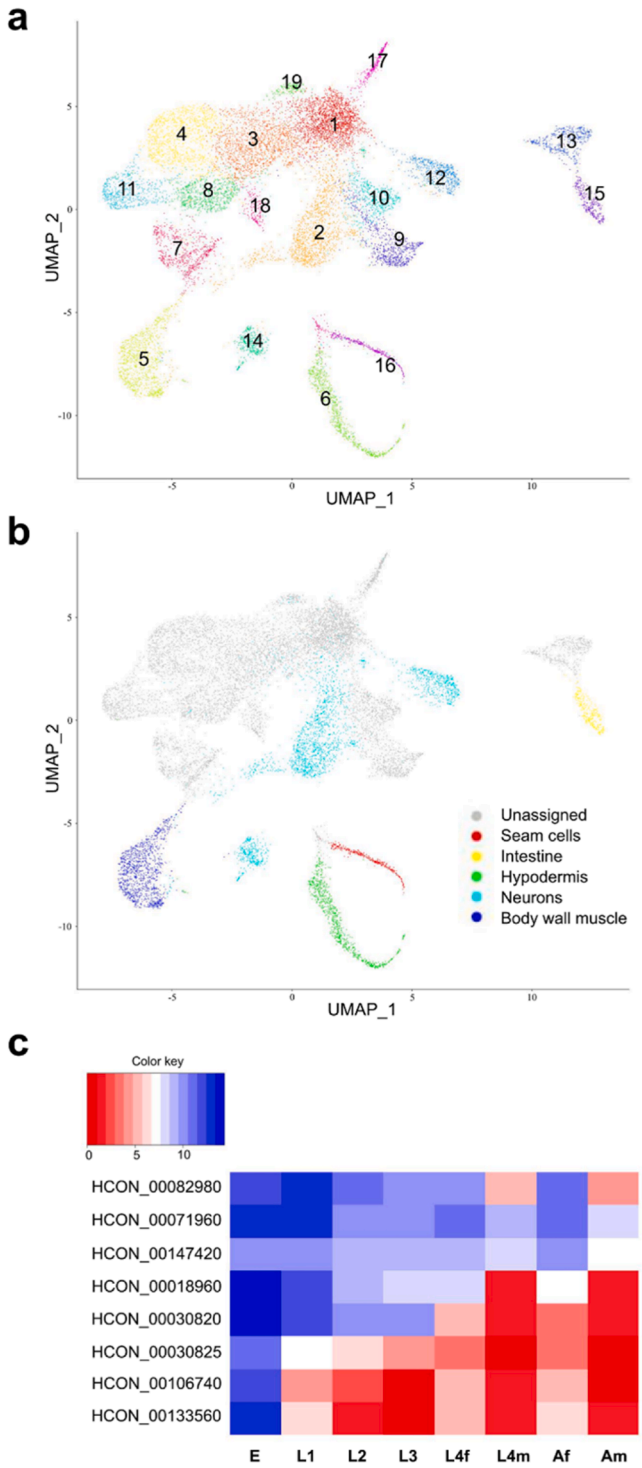

Fig. 1. Clusters of cell-types (n = 19; Table 1) based on the analysis of single nucleus (sn)RNA-sequence data from *Haemonchus contortus* eggs during embryonation (**a**) the annotation of five distinct cell types of *H. contortus* by comparative analysis with a reference single cell data set (GSE126954) for *Caenorhabditis elegans* [10] (**b**). Levels of transcription (given in transcripts per million, TPM) in eight distinct stages/sexes of *H. contortus* (i.e. E = egg; L1, L2, L3 and L4 = four distinct larval stages; L4f = female L4; L4m = male L4; Af = female adult; Am = male adult) of eight protein-coding genes represented in all 19 clusters of cell-types inferred for *H. contortus* embryos (cf. Table 2) (**c**).

eef-1 G – are classified as essential, with an Emb (lethal) phenotype resulting directly from RNAi knockdown in *C. elegans* [51–57].

Hc-ant-1 encodes an ANT-1 which has a higher amino acid sequence identity (88.6%) to *C. elegans* ANT-1.1 than to ANT-1.2 to ANT-1.4 (80.3–84.3%). The transcription profile seen here for *Hc-ant-1* (Fig. 1) is consistent with that observed previously [58], with expression present in all key developmental stages, and being higher in adult females than males of *H. contortus* (Fig. 1c and Table S10; cf. [58]). Current evidence indicates that ANT-1.1 is located in the mitochondrion and inferred to enable ATP:ADP antiporter activity (ADP/ATP exchange between cytosol and mitochondria); it is critical for mitochondrion organisation, is required for embryogenesis and normal postembryonic body size and adult life-span, and positively regulates the apoptotic process. Previous work on transgenic lines of *C. elegans* containing GFP-constructs (5′-UTRs plus a region of exon 1 representing each *Hc-ant-1* and *C. elegans ant-1.1* [58] showed consistent expression in pharynx, body wall, tail muscle, body muscle, hypodermis and intestine, with slight variation in expression among individual worms and tissues for *Hc-ant-1*. The differential transcription seen between adult female and male worms of *H. contortus* (Fig. 1; cf. [58]) supports a key role in oogenesis, in addition to embryogenesis.

Current evidence also shows that the *ant-1.1* gene has a vast network of at least 95 interactors – 41 of which are physical (protein-protein), 8 genetic (suppressive), 4 regulatory and 42 predicted – and 372 co-transcribed molecules [59–80] (via WormBase – 15 March 2023). Interestingly, Shen et al. [62] identified ANT-1.1 (previously referred to as 'WAN-1') as an important cell death regulator. Genetic inactivation of *ant-1.1* significantly suppressed both somatic and germ line cell deaths in *C. elegans*, the chemical inhibition of ANT-1.1 activity caused major reduction in germ line apoptosis. In the mitochondria, ANT-1.1 can form a complex with both CED-4 and CED-9, the cell death initiator EGL-1 can disrupt its interaction with CED-9 [62]. These findings suggest that ANT-1.1 (also as a dimer or oligomer) is involved in the cell-killing process, and works with the core cell death machinery to promote programmed cell death in *C. elegans*.

Based on evidence for *C. elegans*, the five RRM domain-containing proteins identified are proposed to be part of the mRNA cleavage and polyadenylation specificity factor (CPSF) complex [81,82], which enable mRNA binding activity and pre-mRNA processing required for polyadenylation. The *rbm-3.2* gene, for example, has genetic interactions with *spn-4* (predicted to be involved in nervous system development and the regulation of alternative mRNA splicing, via the spliceosome) and *dhc-1* (predicted to enable dynein intermediate chain-binding activity) (cf. WormBase). Collectively, these proteins are part of the cleavage–polyadenylation step required for the maturation of primary protein-encoding transcripts into functional mRNAs, which can be exported from the nucleus to the cytoplasm for translation; the 3′-end processing depends on the binding of multi-protein complexes to specific elements at the 3′-end of the pre-mRNA [83].

An analysis of function in *C. elegans* indicated that both proteins *Hc*-eEF1A and *Hc*-eEF1G are linked to a key step of protein synthesis (i.e. translation elongation) in *H. contortus* (cf. Fig. 2), with *eef-1A.1* and *eef1G* being essential and linked to knockdown phenotypes including Emb and/or Lvl upon RNAi [51,52,54,56]. In the free-living nematode, eEF1A.1 is located in the cytoplasm, is expressed predominantly in hypodermal, intestinal, muscle and neuronal cells, and promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis. We observed that *eef1A.1* has an extensive network of at least 119 interactions – 59 of which are regulatory, 31 physical (protein-protein), one genetic and 28 predicted – and associates with hundreds of co-transcribed genes [60,61,69,70,71,75,84–98] (via WormBase – 15 March 2023). Subsequently, we predicted the structure

Table 2
Eight protein-coding genes transcribed in all 19 cell clusters representing *Haemonchus contortus* eggs during embryonation; their names, annotations, essentiality status based on information for *Caenorhabditis elegans*, and WormBase accession codes for orthologous genes in *C. elegans*.

Protein-coding gene in *H. contortus*	Assigned gene name	Annotation of encoded protein	Essential	WormBase accession (*C. elegans* gene homolog)
HCON_00082980	*Hc-ant-1*	ADP, ATP carrier protein	Yes	WBGene00006439 (*ant-1.1*), WBGene00007057 (*ant-1.2*), WBGene00010485 (*ant-1.3*), WBGene00020140 (*ant-1.4*)
HCON_00071960	*Hc-eef-1A*	Translation elongation factor 1-alpha	Yes	WBGene00001168 (*eef-1A.1*), WBGene00001169 (*eef-1A.2*)
HCON_00147420	*Hc-eef-1G*	Translation elongation factor 1-gamma	Yes	WBGene00008920 (*eef-1G*; three isoforms in *C. elegans*)
HCON_00018960	*Hc-rbm-3*	RRM domain-containing protein[a]	No	WBGene00011155 (*rbm-3.1*), WBGene00011156 (*rbm-3.2*), WBGene00011059 (R06C1.4)
HCON_00030820	*Hc-rbm-3*	RRM domain-containing protein	No	WBGene00011155 (*rbm-3.1*), WBGene00011156 (*rbm-3.2*), WBGene00011059 (R06C1.4)
HCON_00030825	*Hc-rbm-3*	RRM domain-containing protein	No	WBGene00011155 (*rbm-3.1*), WBGene00011156 (*rbm-3.2*), WBGene00011059 (R06C1.4)
HCON_00106740	*Hc-rbm-3*	RRM domain-containing protein	No	WBGene00011155 (*rbm-3.1*), WBGene00011156 (*rbm-3.2*), WBGene00011059 (R06C1.4)
HCON_00133560	*Hc-rbm-3*	RRM domain-containing protein	No	WBGene00011155 (*rbm-3.1*), WBGene00011156 (*rbm-3.2*), WBGene00011059 (R06C1.4)

[a] RRM = RNA recognition motif.

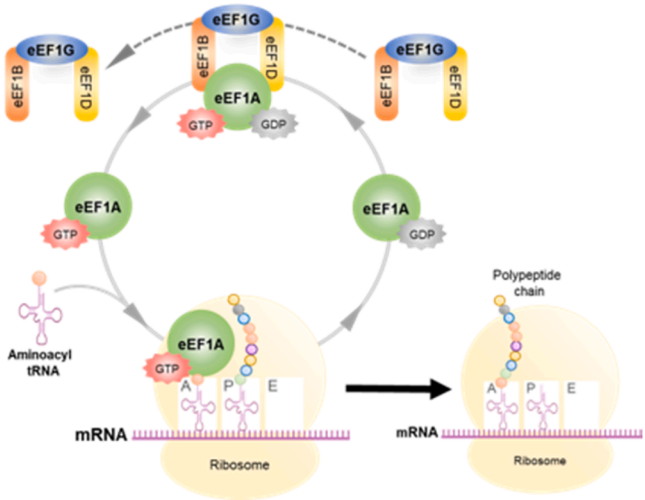

Fig. 2. A schematic representation of the translation elongation process in metazoans – indicating the involvement of eukaryotic elongation factors eEF1A and eEF1G. The active form of eEF1A complexed with GTP delivers an aminoacylated tRNA to the A-site of the ribosome. Upon codon-anticodon recognition, the GDP is hydrolysed, and the inactive eEF1A-GDP is released from the ribosome and is then bound by the eEF1B2GD complex, which forms a macromolecular protein aggregate, designated eEF1H (comprising the subunits eEF1B2, eEF1G and eEF1D). This complex promotes the exchange between GDP and GTP to regenerate the active eEF1A.
Image adapted from Li et al. [116].

of the protein encoded by HCON_00071960 using AlphaFold2, with high confidence (pLDDT > 70) for most of this protein, including tr-type G domain (IPR031157), and limited confidence (pLDDT≤70) for the very N- and C-termini and the 'disordered' region (see Fig. 3). A comparison of this structure with that of an ortholog from a representative host animal, *Ovis aries* (GenBank identifier KAG5196448.1), revealed minor structural variation (TM score: 0.92) (Fig. 3).

In *C. elegans*, eEF1G is located within both nucleus and cytoplasm, and *eef-1 G* has a network of at least 75 interactions – 24 of which are regulatory (positive or negative), 27 are physical (protein-protein) 2 genetic (suppressive) and 22 predicted – and 41 co-transcribed genes [60,61,63,69,73,75,77,79,87,95,99–107]. Although *eef-1 G* is recognised as an essential gene in many organisms, including mammals, meaning that its function is crucial for the survival of an organism [20], the sequence and structural conservation of its gene product (eEF1G) had not been studied. Here, we investigated sequence and structural variation in eEF1G between *H. contortus* and representative invertebrate and vertebrate animals, to explore whether HCON_00147420, or parts thereof, might have potential as an intervention target.

Hc-eEF1G, encoded by HCON_00147420, has distinct GST C-terminal (IPR010987; amino acid positions 73–205) and an elongation factor 1B-gamma C-terminal (IPR001662; positions 253–413) domains as well as an intervening, 'disordered' region (with compositional bias; positions 206–252). The AlphaFold2-predicted structure of the N-terminal domain (positions 1–205) was markedly more confident ($p = 0$; Mann-Whitney one-tailed U-test) than the C-terminal domain (positions 253–413), and the predicted structure of the intervening region (positions 206–252) had low confidence (Fig. 3; Table S11 – with a median pLDDT value of 33). A comparison of the *Hc*-eEF1G sequence with an ortholog from a representative host animal, *Ovis aries* (GenBank accession KAG5196448.1), revealed more variation in the N-terminal ($p = 0$; Mann Whitney one-tailed U-test) than the C-terminal domain (Fig. 3; conservation scores given in Table S11). A subsequent phylogenetic analysis of aligned protein sequences of eEF1G orthologs from selected parasitic nematodes of the order Strongylida (including haematophagous worms – *Haemonchus contortus*, *H. placei*, *Ancylostoma caninum* and *A. ceylanicum*); free-living nematodes (including *C. elegans*, *C. brenneri* and *Pristionchus pacificus*); arthropods (including *Drosophila melanogaster*); and mammal host species (*Ovis aries*, *Bos taurus*, *Homo sapiens* and *Canis lupus familaris*) revealed a close relationship of *Hc*-eEF1G with full-length orthologs from (mainly clade V) nematodes (nodal support – pp = 1.00), to the exclusion of those from arthropods (pp = 1.00) and mammals (pp = 1.00) (Fig. 4; Table S11).

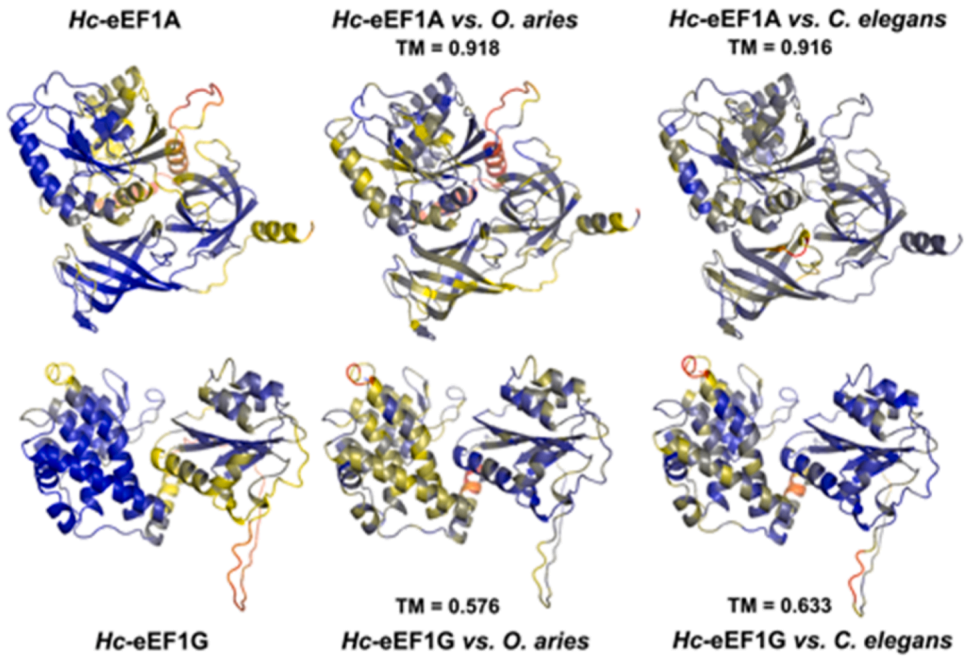

Fig. 3. Three-dimensional structure models for eukaryotic elongation factor proteins *Hc*-eEF1A and *Hc*-eEF1G, inferred from respective *Haemonchus contortus* genes HCON_00071960 (*Hc-eef-1A*) and HCON_00147420 (*Hc-eef-1 G*) (cf. Table 2), and a comparison of each of these proteins with their respective homolog/ortholog from a representative host animal – *Ovis aries* (accession nos. XP_042085813.1 and XP_027815168.1), or the free-living nematode – *Caenorhabditis elegans* (accession nos. F17C11.9 and R03G5.1). Blue (high), yellow (moderate) and red (low) indicate the confidence of the structure prediction (left), or the conservation (Capra07) upon comparison (mid and right). Template modelling (TM) scores for structural comparisons are indicated on the image.

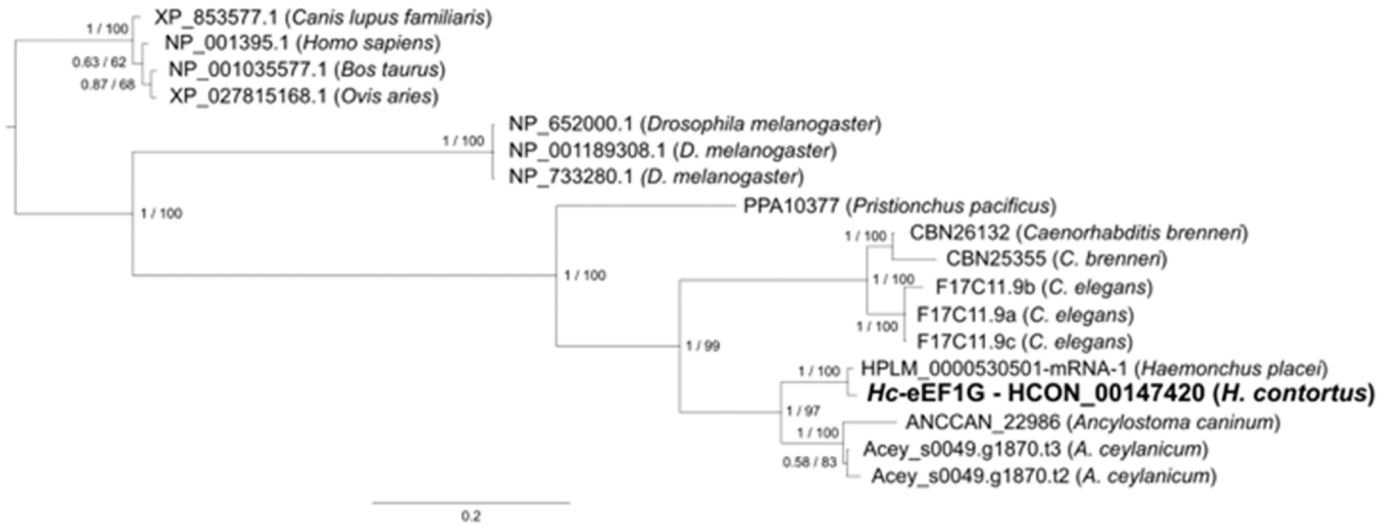

Fig. 4. Phylogenetic tree showing the relationship of protein *Hc*-eEF1G, inferred from the *Haemonchus contortus* gene (HCON_00147420), with orthologous proteins from parasitic nematodes of the order Strongylida, including *Ancylostoma caninum* (WormBase ParaSite accession ANCCAN_22986), *A. ceylanicum* (WormBase ParaSite accession Acey_s0033.g1870.t2 and Acey_s0033.g1870.t3), *Haemonchus placei* (WormBase ParaSite accession HPLM_000530501-mRNA-1); free-living nematodes including *C. elegans* (WormBase accessions F17C11.9 – isoforms a to c), *Caenorhabditis brenneri* (WormBase accession CBN026132) and *Pristionchus pacificus* (WormBase ParaSite accession PPA10377); the model insect *Drosophila melanogaster* (NCBI accessions NP_652000.1, NP_001189308.1 and NP_733280.1); and mammalian host species of the strongylid nematodes, including *Ovis aries* (NCBI accession XP_027815168.1), *Bos taurus* (NCBI accession XP_001035577.1), *Canis lupus familiaris* (NCBI accession XP_853577.1) and *Homo sapiens* (NCBI accession NP_001395.1). Nodal support values given as posterior probability (MrBayes) and bootstrap percentage (RAxML).

4. Discussion

Parasitic nematodes cause some of the world's most neglected diseases, affecting hundreds of millions of people and animals worldwide (e.g., [108–111]). Despite efforts to control these nematodes, no highly effective, commercial vaccines are available for the majority of nematodes of animals, and treatment relies on only a small number of drug classes [112]. Even when using the native protein vaccine (Barbervax®), additional anthelmintic treatment is required to reduce infection intensity or achieve control [113]. Widespread resistance to most of these compounds seriously compromises the control of parasitic worms, particularly in livestock animals [114]. Consequently, there is an urgent need to develop new interventions against these parasites, built on a deep understanding of their biology.

The mechanisms of drug resistance in parasitic nematodes remain poorly understood at the molecular and biochemical levels, such that there is a major need for profound fundamental investigations at the cellular level. In spite of recent progress, there are distinct gaps in our knowledge and understanding of these nematodes' biology, and precisely how they cause disease, preventing the discovery of new methods of control. *Haemonchus contortus* is arguably the most important and pathogenic nematodes of ruminant animals worldwide [17], and causes substantial disease as well as production losses to the livestock and associated industries and financial losses due to the costs of treatment and control [110]. For these reasons, much research effort has been focused on establishing essential genomic, transcriptomic and proteomic infrastructure for *H. contortus* to underpin detailed molecular and cellular biology investigations [19,22,115], with a view to identifying new intervention targets in this and related parasitic nematode species, particularly those of the order Strongylida.

Here, we created the first single cell atlas for *H. contortus* eggs undergoing embryonation through snRNA-seq and comparative analyses with *C. elegans*, employing advanced informatic tools. By focusing on genes that were transcribed in all clusters and types of cells in *H. contortus* embryos and recognised to be essential, we inferred a small subset of essential genes – *Hc-ant-1*, *eef-1A* or *eef-1 G* – the latter two of which are likely crucially involved in protein biosynthesis (Fig. 2; cf. [116]). Guided by information from *C. elegans*, we infer that *Hc-eef-1A* (HCON_00071960) and *Hc-eef-1 G* (HCON_00147420) are co-transcribed and/or interact (genetically, physically or in a regulatory manner) with large numbers (hundreds) of other genes involved in networks of critical biological processes/pathways, and show that both of these genes are highly transcribed and expressed in all key stages of *H. contortus*. Structure modelling and comparisons of orthologous proteins revealed relative conservation for eEF1A orthologs, but marked variation in the N-terminus of *Hc*-eEF1G (HCON_00147420) with respect to the host animal (sheep) and free-living nematodes, indicating that the latter protein may be more specific to *H. contortus* and related nematodes and, thus, might be amenable as a drug target. Nonetheless, both of these elongation factors have potential as targets, and should be considered in future evaluations.

Thus, this fundamental investigation of a first cell atlas of *H. contortus* has guided us toward eukaryotic protein synthesis machinery (specifically, elongation; Fig. 2) as a potential target for nematocidal drugs. An appraisal of published literature reveals that a range of natural product-derived compounds have been studied and shown to interfere with the elongation phase of eukaryotic protein biosynthesis (e.g., [117,118]); some of these compounds have been explored and assessed as anti-cancer or ant-infective agents. Although cycloheximide, originating from *Streptomyces griseus*, has long been a prototype inhibitor of eukaryotic translation elongation, other natural compounds ($n > 23$) of various origins have been explored and shown to target the elongation step in different ways, including the inference with elongation factors eEF1 and eEF2, or binding to a particular site (A, P or E) in the ribosome [118]. Of the many compounds that have been studied, plitidepsin (Aplidin®) and homoharringtonine (Synribo®) have reached market approval. Other translation elongation inhibitors investigated to tackle infectious diseases include cycloheximide, lactimidomycin and homoharringtonine for select viruses [118] and, for example, compound DDD107498 for the malaria parasite, *Plasmodium* [119], which targets eEF2 in the GTP-dependent translocation of the ribosome along messenger RNA, and is essential for protein synthesis (cf. Fig. 2). DDD107498 was obtained through a phenotypic screen of a protein kinase scaffold library of 4731 compounds [120] against the blood stage of multi-drug-sensitive *Plasmodium falciparum* 3D7 strain; a compound series from this screen, based on a 2,6-disubstituted quinoline-4-carboxamide scaffold, had sub-micromolar potency against *Plasmodium* parasites, but suffered from poor physicochemical properties. Subsequent chemical optimisation resulted in DDD107498 with improved physicochemical properties and a 100-fold increase in potency [119]. This example indicates that medicinal chemistry approaches might be applicable to improve the selectivity of the effects of modified compounds to the parasite. While the possibility of "rational" design of compounds (cf. [121]) to eEF1A and eEF1G seems appealing, from an applied perspective, we would initially propose to screen known elongation inhibitors and then assess their potency on parasitic stages of *H. contortus* and related nematodes, and evaluate toxicity/mitochondrial toxicity in mammalian cells (representing the host animal) and in free-living nematodes (including *C. elegans*) to establish their selectivity/suitability for parasitic nematodes. If promising small molecules were identified that might be synthesized, we would propose medicinal chemistry studies to maximise potency and minimise toxicity. Such work would benefit from the use of a recently-established thermal proteome profiling assay [122] to identify or confirm compound-target interaction(s) as well as complementary snRNA-seq studies of different developmental stages of *H. contortus* to evaluate mechanism(s) of action.

5. Conclusion

Through the present investigation, our ability to identify cell populations or lineages of cells undergoing embryonation from snRNA-seq data using the present informatic workflow creates a novel paradigm for investigating *H. contortus* and related parasites by transferring the focus from research on whole organisms and tissues to the fine detail of cell populations within individual worms and tissues. Our plan now is to extend single nuclei analyses to other developmental stages and complement these analyses with single-cell analyses. This should allow the construction of an atlas of cell types and their differentiation for throughout *H. contortus* development, from embryos through to adult, blood-feeding females and males, which live within the host animal, to fill knowledge gaps in the complex biology of this worm and create critical data sets and capacity to explore related eukaryotic parasites, and to support the discovery of novel and improved interventions to combat this and related parasitic worms. These advances are also expected to open up completely new avenues to study drug resistance; drug discovery; vaccine- and drug-target discovery; host-parasite interactions and parasitism, thereby advancing the field of molecular helminthology.

Data and code availability

The project data is available through NCBI BioProject PRJNA957936 in SRA database (accession identifiers SRR24235293–6), and the scripts developed for the analysis of this snRNA data are available at https://gitlab.unimelb.edu.au/bioscience/haemonchus-contortus-snrna (version 0.01-beta).

CRediT authorship contribution statement

Pasi K. Korhonen: Curated RNA sequence data, Conducted all bioinformatic analyses using software described, Curation, Prepared results for tables and figures, Contributed to the drafting respective sections in the manuscript. **Tao Wang:** Collected and purified eggs of *H. contortus*, Isolated nuclei, Contributed to FACS and RNA sequencing, The drafting of sub-sections in the manuscript. **Neil D. Young:** Bioinformatics support, Discussion and comments on draft manuscript. **Joe J. Byrne:** Maintenance of the life cycle of *H. contortus* (Haecon5). **Tulio L. Campos:** Provision of gene essentiality data for comparison and comments on manuscript. **Aya C. Taki:** Maintenance of the life cycle of *H. contortus* (Haecon5) and contribution to figure preparation. **Robin B. Gasser:** Conceptualization, Supervision, Investigation and drafting of the manuscript, Editing and revision, Research funding.

Declaration of Competing Interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors sincerely thank Dr Daniela Zalcenstein and Casey Anttila, Walter and Eliza Hall Institute of Medical Research, for excellent service and support, and Associate Professor Luciano G. Martelotto, University of Adelaide, for support with RNA-seq analysis in the early stage of this project. This research was funded through grants from the Australian Research Council (ARC; LP180101085, LP180101334 and LP190101209). We are grateful to Jody Zawadzki and the late Paul Presidente for originally providing us with the Haecon-5 strain of *H. contortus*.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2024.01.008.

References

[1] Brenner S. In the beginning was the worm. Genetics 2009;182:413–5.
[2] Nigon VM, Félix M-A. History of research on *C. elegans* and other free-living nematodes as model organisms. (7 Sept. 2017). In: WormBook, editor. The *C. elegans* Research Community. WormBook; 2017. doi/10.1895/wormbook.1.181.1, ⟨http://www.wormbook.org⟩.
[3] White J, Bretscher MS. Sydney Brenner. 13 January 1927—5 April 2019. Biogr Mems Fell R Soc 2020;69(6).
[4] Davis P, Zarowiecki M, Arnaboldi V, Becerra A, Cain S, Chan J, Chen WJ, Cho J, da Veiga Beltrame E, Diamantakis S, Gao S, Grigoriadis D, Grove CA, Harris TW, Kishore R, Le T, Lee RYN, Luypaert M, Müller HM, Nakamura C, Nuin P, Paulini M, Quinton-Tulloch M, Raciti D, Rodgers FH, Russell M, Schindelman G, Singh A, Stickland T, Van Auken K, Wang Q, Williams G, Wright AJ, Yook K, Berriman M, Howe KL, Schedl T, Stein L, Sternberg PW. WormBase in 2022-data, processes, and tools for analyzing *Caenorhabditis elegans*. Genetics 2022;220:iyac003.
[5] da Veiga Beltrame E, Arnaboldi V, Sternberg PW. WormBase single-cell tools. Bioinform Adv 2022;2:vbac018.
[6] Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. Nat Rev Genet 2019;20:631–56.
[7] Longo SK, Guo MG, Ji AL, Khavari PA. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. Nat Rev Genet 2021;22:627–44.
[8] Martin BK, Qiu C, Nichols E, Phung M, Green-Gladden R, Srivatsan S, Blecher-Gonen R, Beliveau BJ, Trapnell C, Cao J, Shendure J. Optimized single-nucleus transcriptional profiling by combinatorial indexing. Nat Protoc 2023;18:188–207.
[9] Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ, Adey A, Waterston RH, Trapnell C, Shendure J. Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 2017;357:661–7.
[10] Packer JS, Zhu Q, Huynh C, Sivaramakrishnan P, Preston E, Dueck H, Stefanik D, Tan K, Trapnell C, Kim J, Waterston RH, Murray JI. A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. Science 20 2019;365(6459):eaax1971.
[11] Lorenzo R, Onizuka M, Defrance M, Laurent P. Combining single-cell RNA-sequencing with a molecular atlas unveils new markers for *Caenorhabditis elegans* neuron classes. Nucleic Acids Res 2020;48:7119–34.
[12] Lambert J, Lloret-Fernández C, Laplane L, Poole RJ, Jarriault S. On the origins and conceptual frameworks of natural plasticity - lessons from single-cell models in *C. elegans*. Curr Top Dev Biol 2021;144:111–59.
[13] Ma X, Zhao Z, Xiao L, Xu W, Kou Y, Zhang Y, Wu G, Wang Y, Du Z. A 4D single-cell protein atlas of transcription factors delineates spatiotemporal patterning during embryogenesis. Nat Methods 2021;18:893–902.
[14] Soria CLD, Lee J, Chong T, Coghlan A, Tracey A, Young MD, et al. Single-cell atlas of the first intra-mammalian developmental stage of the human parasite *Schistosoma mansoni*. Nat Commun 2020;11:6411.
[15] Wang J, Sun H, Jiang M, Li J, Zhang P, Chen H, Mei Y, Fei L, Lai S, Han X, Song X, Xu S, Chen M, Ouyang H, Zhang D, Yuan GC, Guo G. Tracing cell-type evolution by cross-species comparison of cell atlases. Cell Rep 2021;34:108803.
[16] Wang R, Zhang P, Wang J, Ma L, Weigao E, Suo S, Jiang M, Li J, Chen H, Sun H, Fei L, Zhou Z, Zhou Y, Chen Y, Zhang W, Wang X, Mei Y, Sun Z, Yu C, Shao J, Fu Y, Xiao Y, Ye F, Fang X, Wu H, Guo Q, Fang X, Li X, Gao X, Wang D, Xu PF, Zeng R, Xu G, Zhu L, Wang L, Qu J, Zhang D, Ouyang H, Huang H, Chen M, Ng SC, Liu GH, Yuan GC, Guo G, Han X. Construction of a cross-species cell landscape at single-cell level. Nucleic Acids Res 2023;51:501–16.
[17] Gasser RB, Samson-Himmelstjerna GV. *Haemonchus contortus* and Haemonchosis – Past, Present and Future Trends. Advances in Parasitology. Academic Press,; 2016. ISBN 9780128103951.
[18] Stroehlein AJ, Young ND, Gasser RB. Advances in kinome research of parasitic worms – implications for fundamental research and applied biotechnological outcomes. Biotechnol Adv 2018;36:915–34.
[19] Ma G, Wang T, Korhonen PK, Hofmann A, Sternberg PW, Young ND, Gasser RB. Elucidating the molecular and developmental biology of parasitic nematodes: moving to a multiomics paradigm. Adv Parasitol 2020;108:175–229.
[20] Campos TL, Korhonen PK, Hofmann A, Gasser RB, Young ND. Machine learning for the prediction and prioritisation of essential genes in eukaryotes – challenges and prospects. Biotechnol Adv 2022;54:107822.
[21] Doyle SR, Tracey A, Laing R, Holroyd N, Bartley D, Bazant W, Beasley H, Beech R, Britton C, Brooks K, Chaudhry U, Maitland K, Martinelli A, Noonan JD, Paulini M, Quail MA, Redman E, Rodgers FH, Sallé G, Shabbir MZ, Sankaranarayanan G, Wit J, Howe KL, Sargison N, Devaney E, Berriman M, Gilleard JS, Cotton JA. Genomic and transcriptomic variation defines the chromosome-scale assembly of *Haemonchus contortus*, a model gastrointestinal worm. Commun Biol 2020;3:656.
[22] Doyle SR. Improving helminth genome resources in the post-genomic era. Trends Parasitol 2022;38:831–40.
[23] Schwarz EM, Korhonen PK, Campbell BE, Young ND, Jex AR, Jabbar A, Hall RS, Mondal A, Howe AC, Pell J, Hofmann A, Boag PR, Zhu XQ, Gregory TR, Loukas A, Williams BA, Antoshechkin I, Brown CT, Sternberg PW, Gasser RB. The genome and developmental transcriptome of the strongylid nematode *Haemonchus contortus*. Genome Biol 2013;14:R89.
[24] Mes TH, Eysker M, Ploeger HW. A simple, robust and semi-automated parasite egg isolation protocol. Nat Protoc 2007;2:486–9.
[25] Veglia F. The Anatomy and Life-History of *Haemonchus contortus* (Rud.). Rep Dir Vet Res (Union South Afr) 1915;3-4:347–500.
[26] Martelotto, L.G., 2020. 'Frankenstein' protocol for nuclei isolation from fresh and frozen tissue for snRNAseq V.3. dx.doi.org/10.17504/protocols.io.bqxymxpw.
[27] Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, Gregory MT, Shuga J, Montesclaros L, Underwood JG, Masquelier DA, Nishimura SY, Schnall-Levin M, Wyatt PW, Hindson CM, Bharadwaj R, Wong A, Ness KD, Beppu LW, Deeg HJ, McFarland C, Loeb KR, Valente WJ, Ericson NG, Stevens EA, Radich JP, Mikkelsen TS, Hindson BJ, Bielas JH. Massively parallel digital transcriptional profiling of single cells. Nat Commun 2017;8:14049.
[28] Team RDC. R: A Language and Environment for Statistical Computing. In: Book R, editor. A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing; 2011. ISBN: 3-900051-07-0. Available online at http://www.R-project.org/. Vienna, Austria.
[29] Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck 3rd WM, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive integration of single-cell data. Cell 2019;177:1888–1902.e21.
[30] Bais AS, Kostka D. scds: computational annotation of doublets in single-cell RNA sequencing data. Bioinformatics 2020;36:1150–8.
[31] Alvarez M, Rahmani E, Jew B, Garske KM, Miao Z, Benhammou JN, Ye CJ, Pisegna JR, Pietiläinen KH, Halperin E. Enhancing droplet-based single-nucleus RNA-seq resolution using the semi-supervised machine learning classifier DIEM. Sci Rep 2020;10:1–16.
[32] Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol 2019;20:1–15.
[33] Lun AT, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Res 2016;5:2122.
[34] McInnes, L., Healy, J., Melville, J., 2018. UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:180203426.
[35] Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, Ginhoux F, Newell EW. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol 2019;37:38–44.
[36] Amezquita RA, Lun AT, Becht E, Carey VJ, Carpp LN, Geistlinger L, Marini F, Rue-Albrecht K, Risso D, Soneson C. Orchestrating single-cell analysis with Bioconductor. Nat Methods 2020;17:137–45.
[37] Fu R, Gillen AE, Sheridan RM, Tian C, Daya M, Hao Y, Hesselberth JR, Riemondy KA. clustifyr: an R package for automated single-cell RNA sequencing cluster classification. F1000Research 2020;9:223.
[38] Alexa, A., Rahnenführer, J., 2023. topGO: enrichment analysis for gene ontology. ⟨doi:10.18129/B9.bioc.topGO⟩.
[39] Campos TL, Korhonen PK, Sternberg PW, Gasser RB, Young ND. Predicting gene essentiality in *Caenorhabditis elegans* by feature engineering and machine-learning. Comp Struct Biotechnol J 2020;18:1093–102.
[40] Geldhof P, Visser A, Clark D, Saunders G, Britton C, Gilleard J, Berriman M, Knox D. RNA interference in parasitic helminths: current situation, potential pitfalls and future prospects. Parasitology 2007;134:609–19.
[41] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–20.
[42] Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite− a comprehensive resource for helminth genomics. Mol Biochem Parasitol 2017;215:2–10.
[43] Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC Bioinforma 2011;12:1–16.
[44] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A. Highly accurate protein structure prediction with AlphaFold. Nature 2021;565:1–11.

[45] Capra JA, Singh M. Predicting functionally important residues from sequence conservation. Bioinformatics 2007;23:1875–82.
[46] Li L, Stoeckert Jr CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 2003;13:2178–89.
[47] Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013;30:772–80.
[48] Stamatakis A, Ludwig T, Meier H. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics 2005;21:456–63.
[49] Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 2003;19:1572–4.
[50] Le SQ, Gascuel O. An improved general amino acid replacement matrix. Mol Biol Evol 2008;25:1307–20.
[51] Gonczy P, Echeverri C, Oegema K, Coulson AR, Jones SJM, Copley RR, Duperon J, Oegema J, Brehm M, Cassin E, Hannak E, Kirkham M, Pichler SC, Flohrs K, Goessen A, Leidel S, Alleaume AM, Martin C, Ozlu N, Bork P, Hymann AA. Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. Nature 2000;408:331–6.
[52] Piano F, Schetter AJ, Mangone M, Stein LD, Kemphues KJ. RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans*. Curr Biol 2000;10:1619–22.
[53] Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, Kanapin A, Le Bot N, Moreno S, Sohrmann M, Welchman DP, Ziperlen P, Ahringer J. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 2003;421:231–7.
[54] Rual JF, Ceron J, Koreth J, Hao T, Nicot AS, Hirozane-Kishikawa T, Vandenhaute J, Orkin SH, Hill DE, van den Heuvel S, Vidal M. Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. Genome Res 2004;14:2162–8.
[55] Fernandez AG, Gunsalus KC, Huang J, Chuang LS, Ying N, Liang HL, Tang C, Schetter AJ, Zegar C, Rual JF, Hill DE, Reinke V, Vidal M, Piano F. New genes with roles in the *C. elegans* embryo revealed using RNAi of ovary-enriched ORFeome clones. Genome Res 2005;15:250–9.
[56] Sönnichsen B, Koski LB, Walsh A, Marschall P, Neumann B, Brehm M, Alleaume AM, Artelt J, Bettencourt P, Cassin E, Hewitson M, Holz C, Khan M, Lazik S, Martin C, Nitzsche B, Ruer M, Stamford J, Winzi M, Heinkel R, Röder M, Finell J, Häntsch H, Jones SJ, Jones M, Piano F, Gunsalus KC, Oegema K, Gönczy P, Coulson A, Hyman AA, Echeverri CJ. Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. Nature 2005;434:462–9.
[57] Ceron J, Rual JF, Chandra A, Dupuy D, Vidal M, van den Heuvel S. Large-scale RNAi screens identify novel genes that interact with the C. elegans retinoblastoma pathway as well as splicing-related components with synMuv B activity. BMC Dev Biol 2007;7:30.
[58] Hu M, Zhong W, Campbell BE, Sternberg PW, Pellegrino MW, Gasser RB. Elucidating ANTs in worms using genomic and bioinformatic tools – Biotechnological prospects? Biotechnol Adv 2010;28:49–60.
[59] Tewari M, Hu PJ, Ahn JS, Ayivi-Guedehoussou N, Vidalain PO, Li S, Milstein S, Armstrong CM, Boxem M, Butler MD, Busiguina S, Rual JF, Ibarrola N, Chaklos ST, Bertin N, Vaglio P, Edgley ML, King KV, Albert PS, Vandenhaute J, Pandey A, Riddle DL, Ruvkun G, Vidal M. Systematic interactome mapping and genetic perturbation analysis of a *C. elegans* TGF-beta signaling network. Mol Cell 2004;13:469–82.
[60] Zhong W, Sternberg PW. Genome-wide Predict *C elegans* Genet Interact 2006;311:1481–4.
[61] Lee I, Lehner B, Crombie C, Wong W, Fraser AG, Marcotte EM. A single gene network accurately predicts phenotypic effects of gene perturbation in *Caenorhabditis elegans*. Nat Genet 2008;40:181–8.
[62] Shen Q, Qin F, Gao Z, Cui J, Xiao H, Xu Z, Yang C. Adenine nucleotide translocator cooperates with core cell death machinery to promote apoptosis in *Caenorhabditis elegans*. Mol Cell Biol 2009;29:3881–93.
[63] Lenfant N, Polanowska J, Bamps S, Omi S, Borg JP, Reboul J. A genome-wide study of PDZ-domain interactions in *C. elegans* reveals a high frequency of non-canonical binding. BMC Genom 2010;11:671.
[64] Berends CW, Munoz J, Portegijs V, Schmidt R, Grigoriev I, Boxem M, Akhmanova A, Heck AJ, van den Heuvel S. F-actin asymmetry and the endoplasmic reticulum-associated TCC-1 protein contribute to stereotypic spindle movements in the *Caenorhabditis elegans* embryo. Mol Biol Cell 2013;24:2201–15.
[65] Cecere G, Hoersch S, Jensen MB, Dixit S, Grishok A. The ZFP-1(AF10)/DOT-1 complex opposes H2B ubiquitination to reduce Pol II transcription. Mol Cell 2013;50:894–907.
[66] Bennett CF, Vander Wende H, Simko M, Klum S, Barfield S, Choi H, Pineda VV, Kaeberlein M. Activation of the mitochondrial unfolded protein response does not predict longevity in *Caenorhabditis elegans*. Nat Commun 2014;5:3483.
[67] Chen JX, Cipriani PG, Mecenas D, Polanowska J, Piano F, Gunsalus KC, Selbach M. In vivo interaction proteomics in *C. elegans* embryos provides new insights into P granule dynamics. Mol Cell Proteom 2016;15:1642–57.
[68] Tan D, Li Q, Zhang MJ, Liu C, Ma C, Zhang P, Ding YH, Fan SB, Tao L, Yang B, Li X, Ma S, Liu J, Feng B, Liu X, Wang HW, He SM, Gao N, Ye K, Dong MQ, Lei X. Trifunctional cross-linker for mapping protein-protein interaction networks and comparing protein conformational states. eLife 2016;5:12509.
[69] Waaijers S, Munoz J, Berends C, Ramalho JJ, Goerdayal SS, Low TY, Zoumaro-Djayoon AD, Hoffmann M, Koorman T, Tas RP, Harterink M, Seelk S, Kerver J, Hoogenraad CC, Bossinger O, Tursun B, van den Heuvel S, Heck AJ, Boxem M. A tissue-specific protein purification approach in *Caenorhabditis elegans* identifies novel interaction partners of DLG-1/Discs large. BMC Biol 2016;14:66.
[70] Wang W, Perens EA, Oikonomou G, Lu Y, Shaham S. IGDB-2, an Ig/FNIII protein, binds the ion channel LGC-34 and controls sensory compartment morphogenesis in *C. elegans*. Dev Biol 2017;430:105–12.
[71] Marnik EA, Fuqua JH, Sharp CS, Rochester JD, Xu EL, Holbrook SE, Updike DL. Germline maintenance through the multifaceted activities of GLH/Vasa in *Caenorhabditis elegans* P granules. Genetics 2019;213:923–39.
[72] Wang SY, Mao H, Shibuya H, Uzawa S, O'Brown ZK, Wesenberg S, Shin N, Saito TT, Gao J, Meyer BJ, Colaiacovo MP, Greer EL. The demethylase NMAD-1 regulates DNA replication and repair in the *Caenorhabditis elegans* germline. PLoS Genet 2019;15:e1008252.
[73] Zeng C, Weng C, Wang X, Yan YH, Li WJ, Xu D, Hong M, Liao S, Dong MQ, Feng X, Xu C, Guang S. Functional proteomics identifies a PICS complex required for piRNA maturation and chromosome segregation. Cell Rep 2019;27:3561–72.
[74] Flynn SM, Chen C, Artan M, Barratt S, Crisp A, Nelson GM, Peak-Chew SY, Begum F, Skehel M, De Bono M. MALT-1 mediates IL-17 neural signaling to regulate *C. elegans* behavior, immunity and longevity. Nat Commun 2020;11:2099.
[75] Jia R, Chai Y, Xie C, Liu G, Zhu Z, Huang K, Li W, Ou G. Spectrin-based membrane skeleton is asymmetric and remodels during neural development. J Cell Sci 2020;133:jcs248583.
[76] Sen I, Zhou X, Chernobrovkin A, Puerta-Cavanzo N, Kanno T, Salignon J, Stoehr A, Lin XX, Baskaner B, Brandenburg S, Bjorkegren C, Zubarev RA, Riedel CG. DAF-16/FOXO requires protein phosphatase 4 to initiate transcription of stress resistance and longevity promoting genes. Nat Commun 2020;11:138.
[77] Solinger JA, Rashid HO, Prescianotto-Baschong C, Spang A. FERARI is required for Rab11-dependent endocytic recycling. Nat Cell Biol 2020;22:213–24.
[78] Wan G, Yan J, Fei Y, Pagano DJ, Kennedy S. A conserved NRDE-2/MTR-4 complex mediates nuclear RNAi in *Caenorhabditis elegans*. Genetics 2020;216:1071–85.
[79] Kim H, Ding YH, Lu S, Zuo MQ, Tan W, Conte D, Dong MQ, Mello CC. PIE-1 SUMOylation promotes germline fates and piRNA-dependent silencing in *C. elegans*. eLife 2021;10:63300.
[80] Xia SL, Li M, Chen B, Wang C, Yan YH, Dong MQ, Qi YB. The LRR-TM protein PAN-1 interacts with MYRF to promote its nuclear translocation in synaptic remodeling. eLife 2021;10:e67628.
[81] Cui M, Allen MA, Larsen A, Macmorris M, Han M, Blumenthal T. Genes involved in pre-mRNA 3'-end formation and transcription termination revealed by a lin-15 operon Muv suppressor screen. Proc Natl Acad Sci USA 2008;105:16665–70.
[82] Sun Y, Hamilton K, Tong L. Recent molecular insights into canonical pre-mRNA 3'-end processing. Transcription 2020;11:83–96.
[83] Neve J, Patel R, Wang Z, Louey A, Furger AM. Cleavage and polyadenylation: Ending the message expands gene regulation. RNA Biol 2017;14:865–90.
[84] Davy A, Bello P, Thierry-Mieg N, Vaglio P, Hitti J, Doucette-Stamm L, Thierry-Mieg D, Reboul J, Boulton S, Walhout AJM, Coux O, Vidal M. A protein-protein interaction map of the *Caenorhabditis elegans* 26S proteasome. EMBO Rep 2001;2:821–8.
[85] Simonis N, Rual JF, Carvunis AR, Tasan M, Lemmens I, Hirozane-Kishikawa T, Hao T, Sahalie JM, Venkatesan K, Gebreab F, Cevik S, Klitgord N, Fan C, Braun P, Li N, Ayivi-Guedehoussou N, Dann E, Bertin N, Szeto D, Dricot A, Yildirim MA, Lin C, de Smet AS, Kao HL, Simon C, Smolyar A, Ahn JS, Tewari M, Boxem M, Milstein S, Yu H, Dreze M, Vandenhaute J, Gunsalus KC, Cusick ME, Hill DE, Tavernier J, Roth FP, Vidal M. Empirically controlled mapping of the *Caenorhabditis elegans* protein-protein interactome network. Nat Methods 2009;6:47–54.
[86] Havrylenko S, Legouis R, Negrutskii B, Mirande M. *Caenorhabditis elegans* evolves a new architecture for the multi-aminoacyl-tRNA synthetase complex. J Biol Chem 2011;286:28476–87.
[87] Li X, Matilainen O, Jin C, Glover-Cutter KM, Holmberg CI, Blackwell TK. Specific SKN-1/Nrf stress responses to perturbations in translation elongation and proteasome activity. PLoS Genet 2011;7:e1002119.
[88] Friend K, Campbell ZT, Cooke A, Kroll-Conner P, Wickens MP, Kimble J. A conserved PUF-Ago-eEF1A complex attenuates translation elongation. Nat Struct Mol Biol 2012;19:176–83.
[89] Kawasaki I, Jeong MH, Yun YJ, Shin YK, Shim YH. Cholesterol-responsive metabolic proteins are required for larval development in *Caenorhabditis elegans*. Mol Cells 2013;36:410–6.
[90] Kirstein-Miles J, Scior A, Deuerling E, Morimoto RI. The nascent polypeptide-associated complex is a key regulator of proteostasis. EMBO J 2013;32:1451–68.
[91] Xin X, Gfeller D, Cheng J, Tonikian R, Sun L, Guo A, Lopez L, Pavlenco A, Akintobi A, Zhang Y, Rual JF, Currell B, Seshagiri S, Hao T, Yang X, Shen YA, Salehi-Ashtiani K, Li J, Cheng AT, Bouamalay D, Lugari A, Hill DE, Grimes ML, Drubin DG, Grant BD, Vidal M, Boone C, Sidhu SS, Bader GD. SH3 interactome conserves general function over specific form. Mol Syst Biol 2013;9:652.
[92] Govindan JA, Jayamani E, Zhang X, Breen P, Larkins-Ford J, Mylonakis E, Ruvkun G. Lipid signalling couples translational surveillance to systemic detoxification in *Caenorhabditis elegans*. Nat Cell Biol 2015;17:1294–303.
[93] MacNeil LT, Pons C, Arda HE, Giese GE, Myers CL, Walhout AJ. Transcription factor activity mapping of a tissue-specific in vivo gene regulatory network. Cell Syst 2015;1:152–62.
[94] Kolundzic E, Ofenbauer A, Bulut SI, Uyar B, Baytek G, Sommermeier A, Seelk S, He M, Hirsekorn A, Vucicevic D, Akalin A, Diecke S, Lacadie SA, Tursun B. FACT sets a barrier for cell fate reprogramming in *Caenorhabditis elegans* and human cells. Dev Cell 2018;46:611–26.
[95] Mutlu B, Chen HM, Moresco JJ, Orelo BD, Yang B, Gaspar JM, Keppler-Ross S, Yates JR, Hall DH, Maine EM, Mango SE. Reulated nuclear accumulation of a

histone methyltransferase times the onset of heterochromatin formation in *C. elegans* embryos. Sci Adv 2018;4:eaat6224.
[96] Bailly AP, Perrin A, Serrano-Macia M, Maghames C, Leidecker O, Trauchessec H, Martinez-Chantar ML, Gartner A, Xirodimas DP. The balance between mono- and NEDD8-chains controlled by NEDP1 upon DNA damage is a regulatory module of the HSP70 ATPase activity. Cell Rep 2019;29:212–24.
[97] Gao K, Li Y, Hu S, Liu Y. SUMO peptidase ULP-4 regulates mitochondrial UPR-mediated innate immunity and lifespan extension. Elife 2019;8:e41792.
[98] Hammerquist AM, Curran SP. Roles for the RNA polymerase III regulator MAFR-1 in regulating sperm quality in *Caenorhabditis elegans*. Sci Rep 2020;10:19367.
[99] Li S, Armstrong CM, Bertin N, Ge H, Milstein S, Boxem M, Vidalain PO, Han JD, Chesneau A, Hao T, Goldberg DS, Li N, Martinez M, Rual JF, Lamesch P, Xu L, Tewari M, Wong SL, Zhang LV, Berriz GF, Jacotot L, Vaglio P, Reboul J, Hirozane-Kishikawa T, Li Q, Gabel HW, Elewa A, Baumgartner B, Rose DJ, Yu H, Bosak S, Sequerra R, Fraser A, Mango SE, Saxton WM, Strome S, Van Den Heuvel S, Piano F, Vandenhaute J, Sardet C, Gerstein M, Doucette-Stamm L, Gunsalus KC, Harper JW, Cusick ME, Roth FP, Hill DE, Vidal M. A map of the interactome network of the metazoan *C. elegans*. Science 2004;303:540–3.
[100] Boxem M, Maliga Z, Klitgord N, Li N, Lemmens I, Mana M, de Lichtervelde L, Mul JD, van de Peut D, Devos M, Simonis N, Yildirim MA, Cokol M, Kao HL, de Smet AS, Wang H, Schlaitz AL, Hao T, Milstein S, Fan C, Tipsword M, Drew K, Galli M, Rhrissorrakrai K, Drechsel D, Koller D, Roth FP, Iakoucheva LM, Dunker AK, Bonneau R, Gunsalus KC, Hill DE, Piano F, Tavernier J, van den Heuvel S, Hyman AA, Vidal M. A protein domain-based interactome network for *C. elegans* early embryogenesis. Cell 2008;134:534–45.
[101] Riedel CG, Dowen RH, Lourenco GF, Kirienko NV, Heimbucher T, West JA, Bowman SK, Kingston RE, Dillin A, Asara JM, Ruvkun G. DAF-16 employs the chromatin remodeller SWI/SNF to promote stress resistance and longevity. Nat Cell Biol 2013;15:491–501.
[102] Fleckenstein T, Kastenmüller A, Stein ML, Peters C, Daake M, Krause M, Weinfurtner D, Haslbeck M, Weinkauf S, Groll M, Buchner J. The chaperone activity of the developmental small heat shock protein Sip1 is regulated by pH-dependent conformational changes. Mol Cell 2015;58:1067–78.
[103] Truttmann MC, Cruz VE, Guo X, Engert C, Schwartz TU, Ploegh HL. The *Caenorhabditis elegans* protein FIC-1 is an AMPylase that covalently modifies heat-shock 70 family proteins, translation elongation factors and histones. PLoS Genet 2016;12:e1006023.
[104] Yi Z, Manil-Segalen M, Sago L, Glatigny A, Redeker V, Legouis R, Mucchielli-Giorgi MH. SAFER, an analysis method of quantitative proteomic data, reveals new interactors of the *C. elegans* autophagic protein LGG-1. J Proteome Res 2016;15:1515–23.
[105] Rehain-Bell K, Love A, Werner ME, Macleod I, Yates JR, Maddox AS. A sterile 20 family kinase and its co-factor CCM-3 regulate contractile ring proteins on germline intercellular bridges. Curr Biol 2017;27:860–70.
[106] Kaminsky R, Denison C, Bening-Abu-Shach U, Chisholm AD, Gygi SP, Broday L. SUMO regulates the assembly and function of a cytoplasmic intermediate filament protein in *C. elegans*. Dev Cell 2009;17:724–35.
[107] Zhu D, Wu X, Zhou J, Li X, Huang X, Li J, Wu J, Bian Q, Wang Y, Tian Y. NuRD mediates mitochondrial stress-induced longevity via chromatin remodeling in response to acetyl-CoA level. Sci Adv 2020;6:eabb2529.
[108] Kiontke K, Fitch DH. Nematodes. Curr Biol 2013;23:R862–4.
[109] Gordon CA, McManus DP, Jones MK, Gray DJ, Gobert GN. The increase of exotic zoonotic helminth infections: the impact of urbanization, climate change and globalization. Adv Parasitol 2016;91:311–97.
[110] Gilleard JS, Kotze AC, Leathwick D, Nisbet AJ, McNeilly TN, Besier B. A journey through 50 years of research relevant to the control of gastrointestinal nematodes in ruminant livestock and thoughts on future directions. Int J Parasitol 2021;51:1133–51.
[111] Loukas A, Maizels RM, Hotez PJ. The yin and yang of human soil-transmitted helminth infections. Int J Parasitol 2021;51:1243–53.
[112] Jiao Y, Preston S, Hofmann A, Taki AC, Baell JB, Chang BCH, Jabbar A, Gasser RB. A perspective on the discovery of selected compounds with anthelmintic activity against the barber's pole worm – where to from here? Adv Parasitol 2020;108:1–45.
[113] Kebeta, M.M., Doyle, E., Walkden-Brown, S.W., Kahn, L.P. and Hine, B., 2022. Performance of Barbervax® vaccine, with focus on lactating ewes and weaner merino Sheep. ⟨https://hdl.handle.net/1959.11/56653⟩.
[114] Sangster NC, Cowling A, Woodgate RG. Ten events that defined anthelmintic resistance research. Trends Parasitol 2018;34:553–63.
[115] Gasser RB, Schwarz EM, Korhonen PK, Young ND. Understanding *Haemonchus contortus* better through genomics and transcriptomics. Adv Parasitol 2016;93:519–67.
[116] Li D, Wei T, Abbott CM, Harrich D. The unexpected roles of eukaryotic translation elongation factors in RNA virus replication and pathogenesis. Microbiol Mol 2013;77(2):253–66.
[117] de Loubresse NG, Prokhorova I, Holtkamp W, Rodnina MV, Yusupova G, Yusupov M. Structural basis for the inhibition of the eukaryotic ribosome. Nature 2014;513:517–22.
[118] Brönstrup M, Sasse F. Natural products targeting the elongation phase of eukaryotic protein biosynthesis. Nat Prod Rep 2020;37:752–62.
[119] Baragaña B, Hallyburton I, Lee MC, Norcross NR, Grimaldi R, Otto TD, Proto WR, Blagborough AM, Meister S, Wirjanata G, Ruecker A, Upton LM, Abraham TS, Almeida MJ, Pradhan A, Porzelle A, Luksch T, Martinez MS, Luksch T, Bolscher JM, Woodland A, Norval S, Zuccotto F, Thomas J, Simeons F, Stojanovski L, Osuna-Cabello M, Brock PM, Churcher TS, Sala KA, Zakutansky SE, Jimenez-Diaz MB, Sanz LM, Riley J, Basak R, Campbell M, Avery VM, Sauerwein RW, Dechering KJ, Noviyanti R, Campo B, Frearson JA, Angulo-Barturen I, Ferrer-Bazaga S, Gamo FJ, Wyatt PG, Leroy D, Siegl P, Delves MJ, Kyle DE, Wittlin S, Marfurt J, Price RN, Sinden RE, Winzeler EA, Charman SA, Bebrevska L, Gray DW, Campbell S, Fairlamb AH, Willis PA, Rayner JC, Fidock DA, Read KD, Gilbert IH. A novel multiple-stage antimalarial agent that inhibits protein synthesis. Nature 2015;522:315–20.
[120] Brenk R, Schipani A, James D, Krasowski A, Gilbert IH, Frearson J, Wyatt PG. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. ChemMedChem 2008;3:435–44.
[121] Mandal S, Moudgil M, Mandal SK. Rational drug design. Eur J Pharmacol 2009;625:90–100.
[122] Taki AC, Wang T, Nguyen NN, Ang CS, Leeming MG, Nie S, Byrne JJ, Young ND, Zheng Y, Ma G, Korhonen PK, Koehler A, Williamson NA, Hofmann A, Chang BCH, Häberli C, Keiser J, Jabbar A, Sleebs BE, Gasser RB. Thermal proteome profiling reveals *Haemonchus* orphan protein HCO_011565 as a target of the nematocidal small molecule UMW-868. Front Pharmacol 2022;13:1014804.