ACCEPTED MANUSCRIPT

Molecular alterations during larval development of *Haemonchus contortus* in vitro are under tight post-transcriptional control[⋆]

Guangxu Ma [a], Tao Wang [a], Pasi K. Korhonen [a], Ching-Seng Ang [b], Nicholas A. Williamson [b], Neil D. Young [a], Andreas J. Stroehlein [a], Ross S. Hall [a], Anson V. Koehler [a], Andreas Hofmann [c], Robin B. Gasser [a,*]

[a] *Department of Veterinary Biosciences, Melbourne Veterinary School, The University of Melbourne, Parkville, Victoria 3010, Australia*
[b] *The Bio21 Institute of Molecular Science and Biotechnology, The University of Melbourne, Parkville, Victoria 3010, Australia*
[c] *Griffith Institute for Drug Discovery, Griffith University, Nathan, Queensland 4111, Australia*

ACCEPTED MANUSCRIPT

[⋆]*Note:* Nucleic acid sequence data from this study have been deposited in the National Center for Biotechnology Information (NCBI) sequence read archive (SRA) – www.ncbi.nlm.nih.gov/sra/ under accession number SRP136037.

* Corresponding author. Tel: +613-97312283; fax: +613-97312366.
E-mail address: robinbg@unimelb.edu.au (R.B. Gasser).

ACCEPTED MANUSCRIPT

ABSTRACT

In this study, we explored the molecular alterations in the developmental switch from the L3 to the exsheathed L3 (xL3) and to the L4 stage of *Haemonchus contortus* in vitro using an integrated transcriptomic, proteomic and bioinformatic approach. Totals of 9,754 mRNAs, 88 microRNAs (miRNAs) and 1,591 proteins were identified, and 6,686 miRNA-mRNA pairs inferred in all larval stages studied. Approximately 16% of transcripts in the combined transcriptome (representing all three larval stages) were expressed as proteins, and there were positive correlations (r = 0.39 to 0.44) between mRNA transcription and protein expression in the three distinct developmental stages of the parasite. Of the predicted targets, 1,019 (27.0%) mRNA transcripts were expressed as proteins, and there was a negative correlation (r = -0.60 to -0.50) in the differential mRNA transcription and protein expression between developmental stages upon pairwise comparison. The changes in transcription (mRNA and miRNA) and protein expression from the free-living to the parasitic life cycle phase of *H. contortus* related to enrichments in biological pathways associated with metabolism (e.g., carbohydrate and lipid degradation, and amino acid metabolism), environmental information processing (e.g., signal transduction, signalling molecules and interaction) and/or genetic information processing (e.g., transcription and translation). Specifically, fatty acid degradation, steroid hormone biosynthesis and the Rap1 signalling pathway were suppressed, whereas transcription, translation and protein processing in the endoplasmic reticulum were upregulated during the transition from the free-living L3 to the parasitic xL3 and L4 stages of the nematode in vitro. Dominant post-transcriptional regulation was inferred to elicit these changes, and particular miRNAs (e.g., *hco-miR-34* and *hco-miR-252*) appear to play roles in stress responses and/or environmental adaptations during developmental transitions of *H. contortus*. Taken together, these integrated results provide a comprehensive insight into the developmental biology of this important parasite at the molecular level in vitro. The approach applied here to *H. contortus* can be readily applied to other parasitic nematodes.

Keywords: Transcriptomics; microRNA; Proteomics; Developmental transition; *Haemonchus contortus*; Parasitic nematode

ACCEPTED MANUSCRIPT

1. Introduction

The free-living nematode *Caenorhabditis elegans* is one of the best-characterised metazoan organisms (Corsi et al., 2015). Integrated analyses of molecular data sets for this nematode have made it possible to gain deep, global insights into areas such as the developmental and reproductive biology, physiology, biochemistry, genetics and neurobiology of metazoans (Piano et al., 2006; van Assche et al., 2015). By comparison, little is known about these areas for parasitic nematodes. Since *C. elegans* is relatively closely related to the socioeconomically important parasitic nematodes of the Order Strongylida (Blaxter et al., 1998), the information and data sets available for *C. elegans* can significantly assist molecular investigations of strongylid nematodes (Gilleard, 2004; Howe et al., 2016a, b). There have been major advances in the genomics and transcriptomics of worms such as *Haemonchus contortus* (barber's pole worm), *Necator americanus* (hookworm) and *Oesophagostomum dentatum* (nodule worm) (see Laing et al., 2013; Schwarz et al., 2013; Tang et al., 2014; Tyagi et al., 2015), allowing novel insights into the systems biology of strongylid nematodes.

Haemonchus contortus is one of the most important parasitic nematodes of livestock animals worldwide (cf. Gasser and von Samson-Himmelstjerna, 2016). This parasite is transmitted orally from contaminated pasture to the host through a relatively complex, 3 week life cycle (Veglia, 1915): eggs are excreted in host faeces; the L1s develop inside the eggs and then hatch (within ~1 day) and moult to the L2 and L3 stages within 1 week; the infective L3s are then ingested by the host, exsheath and, following a histotropic phase, develop through the L4 stage to dioecious adults which both feed on blood from vessels in the stomach wall.

Although some studies have explored changes or differences in mRNAs or proteins in *H. contortus* (see Campbell et al., 2008; Cantacessi et al., 2010; Laing et al., 2013; Schwarz et al., 2013; Wang et al., 2016), no investigation has yet integrated or compared transcriptomic and proteomic profiles during developmental transitions in this nematode or explored regulatory molecular processes. For instance, while small RNAs, such as microRNAs (miRNAs) (Kim, 2005; Grosshans and Filipowicz, 2008), play key roles in regulating genes via mRNA silencing or chromatin modification (Kim, 2005; Grosshans and Filipowicz, 2008), limited information is available for *H. contortus* (see Winter et al., 2012; Gillan et al., 2017; Gu et al., 2017), particularly in relation to this parasite's developmental biology. Here, we explored mRNA, miRNA and protein alterations in *H. contortus* during the developmental transition from ensheathed L3 to xL3 and then

ACCEPTED MANUSCRIPT

to the L4 stage in vitro, utilising a combination of transcriptomic, proteomic and bioinformatic tools. The goal was to identify key molecular changes associated with this crucial phase of this nematode's development using a data integration approach.

2. Materials and methods

2.1. Procurement of parasite and storage

Haemonchus contortus (Haecon-5 strain) was maintained in experimental sheep (cf. Schwarz et al., 2013), in accordance with the institutional animal ethics guidelines (permit no. 1613878; The University of Melbourne, Australia). In brief, helminth-free Merino sheep (8 weeks of age) were inoculated with 7,000 L3s of *H. contortus*. Four weeks p.i., faecal samples were collected each day; L3s were produced by incubating faeces at 27 °C for 1 week (cf. Schwarz et al., 2013), purified through two layers of nylon mesh (pore size: 20 µM; Rowe Scientific, Australia) and stored at 10 °C until use.

2.2. Exsheathment of L3s and culturing to L4s in vitro

L3s were sterilised and exsheathed by incubation in 0.15% v/v sodium hypochlorite (NaClO) at 37 °C for 20 min (Preston et al., 2015). This method was optimised to consistently achieve an exsheathment rate of ≥ 80% (Preston et al., 2015). xL3s were rapidly washed thrice in sterile physiological saline and twice in sterile Luria Bertani (LB medium) by centrifugation at 600 *g* (5 min) supplemented with 100 IU/ml of penicillin, 100 µg/ml of streptomycin, 2.5 µg/ml of amphotericin (Fungizone®, antibiotic-antimycotic; cat no. 15240-062; Thermo Fisher Scientific, USA) and 200 µg/ml of gentamicin (Gibco®, cat no. 15710-064; Thermo Fisher Scientific, USA) - supplemented LB was designated LB*. After the last wash, xL3s were suspended in LB* (38 °C) at a density of 6,000 larvae per ml, transferred to six wells in a culture plate (cat no. 4430500N; Orange Scientific, Belgium) and immediately incubated at 38 °C in an atmosphere of 10% v/v CO_2 in a water-jacketed incubator (model no. 2406 Shel Lab, USA) for 8 h to recover, or 7 days to develop to L4s. Subsequently, all samples were washed thrice in sterile saline (38 °C), snap frozen in liquid nitrogen and stored at – 80 °C.

ACCEPTED MANUSCRIPT

2.3. RNA isolation, RNA-seq and RNA data processing

Total RNA was extracted from five replicates of each L3s, xL3s and L4s (30,000 larvae per replicate) of *H. contortus* using TriPure isolation reagent (cat no. 11667157001; Roche, Germany), and treated with RNase-free TURBO DNase (Ambion®, cat no. AM1907; Thermo Fisher Scientific, USA). Libraries for sequencing were constructed from mRNA and small RNA from L3s, xL3s and L4s using TruSeq Stranded Total RNA Library Prep Kit (Illumina) and TruSeq Small RNA Library Prep Kit (Illumina), respectively. In brief, mRNA was isolated using magnetic beads with Oligo (dT), cleaved into short fragments, transcribed into complementary DNA by reverse synthesis, connected with adaptors and amplified by PCR for mRNA-enriched libraries. For small RNA libraries, RNAs of 20-30 nucleotides (nt) in length were selected by electrophoresis, ligated to adaptors and amplified. All libraries were evaluated using the Agilent 2100 Bioanalyzer and ABI StepOnePlus real-time PCR system, and then sequenced using the Illumina HiSeq™ 4000 platform.

Strand-specific, raw mRNA reads (150 nt in length) were processed to remove adaptors and low quality reads using Trimmomatic (Bolger et al., 2014). High quality mRNA reads were used for de novo assembly and genomic-guided assembly of transcripts employing the programs Trinity (Grabherr et al., 2011), TopHat2 (Kim et al., 2013) and Cufflinks2 (Trapnell et al., 2010). Genome-guided assembly was assisted using the publicly available genome assembly of *H. contortus* (MHco3 ISE strain) (Doyle et al., 2017). The Illumina reads were mapped to predicted genes using the program RSEM (Li and Dewey, 2011) to calculate expected read counts. The threshold for mRNA transcription to be recorded was > 1 read counts per million (CPM) in at least five samples.

Strand-specific, high quality short reads (50 nt in length) produced from small RNA sequencing were processed and analysed using a miARma-Seq package (Andrés-León et al., 2016). In brief, the short reads were processed to remove adaptors, reads of low quality and reads of < 18 nt in length; clean reads were mapped to the genome assembly of *H. contortus* (perfect match) to identify known miRNAs and to predict novel miRNAs. Precursor and mature miRNA sequences representing *H. contortus* as well as *Ascaris suum*, *Brugia malayi*, *Caenorhabditis brenneri*, *Caenorhabditis briggsae*, *C. elegans*, *Caenorhabditis remanei*, *Pristionchus pacificus*, *Panagrellus redivivus* and *Strongyloides ratti* were downloaded from miRBase

ACCEPTED MANUSCRIPT

(release 21; Kozomara and Griffiths-Jones, 2014) and used to identify miRNAs in *H. contortus*. The reads that mapped to the genome were summarised using featureCounts within the miARma-Seq package. The threshold for miRNA transcription to be recorded was >1 CPM in at least three samples. Targets of miRNAs were predicted using TargetScan (Lewis et al., 2005) and PITA (Kertesz et al., 2007), and the consensus target was recorded.

2.4. Protein extraction, LC-MS/MS and proteomic analysis

Proteins were extracted from each of four replicates of L3s, xL3s and L4s (each replicate containing 30,000 individuals). In brief, each larval sample was sonicated (20 kHz) in lysis buffer (8 M urea containing 100 mM triethyl ammonium bicarbonate, pH 8.5) on ice for 20 min, with a 30 s pulse every 30 s. Each sample was supplemented with protease inhibitor cocktail set I (cat no. 539131; Merck, Denmark) and incubated for 1 h at 22 °C. Insoluble debris was removed by centrifugation at 10,000 *g* for 20 min. Protein concentrations were measured using a BCA Protein Assay Kit (cat no. 23227; Thermo Fisher Scientific, USA). Then, samples were subjected to in-solution reduction, alkylation and digestion. In brief, protein aliquots (200 µg) from each sample were reduced with 10 mM tris(2-carboxyethyl)phosphine (TCEP), alkylated with 55 mM iodoacetamide, followed by a double digestion with Lys-C/trypsin Mix (cat no. V5072; Promega, USA) at 37 °C for 16 h. The tryptic samples were acidified with 1.0% (v/v) formic acid and purified using Oasis HLB cartridges (cat no. 186000383; Waters, USA). The processed samples were spiked with calibration peptides (iRT kit, cat no. ki-3002-1; Biognosys, Switzerland) and subjected to LC-MS/MS using a QExactive plus Orbitrap mass spectrometer (Thermo Fisher Scientific, USA) with a nanoESI interface in conjunction with an Ultimate 3000 RSLC nanoHPLC (Dionex Ultimate 3000).

MS data were analysed by employing Spectronaut software (v.11, Biognosys) using default settings. The spectral library used for 'searching' contained 3350 protein groups (22,721 peptides), generated from in-gel-digested and basic reverse phase-separated peptides from key developmental (i.e., egg, L3, xL3, L4 and adult) stages of *H. contortus* (unpublished data). Peptides were accepted based on a false discovery rate (FDR) of < 0.01 at both the peptide and protein levels. The threshold for protein expression to be recorded was ≥ 4 matches in at least four samples.

ACCEPTED MANUSCRIPT

2.5. Comparative analysis

Pairwise comparative analyses were performed based on the identified mRNAs, miRNAs and proteins. To assess the coverage of transcription and expression, the mRNAs and proteins identified in L3, xL3 and L4 stages were compared. To test the differences among these stages, pairwise comparative analyses of transcription and expression were performed. The mRNAs, miRNAs and proteins identified in each stage were compared and displayed using the R package VennDiagram. The transcription or expression levels of the mRNAs, miRNAs and proteins in these stages were subjected to hierarchical cluster analysis using hclust (stat.ethz.ch/R-manual/R-devel/library/stats/html/hclust.html) and illustrated using heatmap2 in the R language (rdocumentation.org/packages/gplots/versions/3.0.1/topics/heatmap.2).

Differentially transcribed mRNAs and miRNAs were identified by pairwise comparison among developmental stages using the limma (Ritchie et al., 2015), Glimma (Su et al., 2017) and edgeR (Robinson et al., 2010) packages (Law et al., 2016) employing a fold change (FC) threshold of > 2, and an FDR of < 0.05. Differentially expressed proteins were identified using the program Spectronaut (v.11, Biognosys) using a threshold of FC of > 2 and FDR of < 0.05. To assess the concordance between mRNA and protein levels, Spearman's correlation analyses of log2-transformed data (CPM for mRNAs; relative abundance for proteins), and log2-transformed FC were performed or calculated using Prism (v.7.0, GraphPad). To evaluate the relationships between miRNAs and their predicted targets, Spearman's correlation and Hoeffding's D measure analyses of log2-transformed data (CPM for mRNAs and miRNAs; relative abundance for proteins) were conducted using miRLAB (in an R language environment) (Le et al., 2015). The statistical significance of correlation analyses was set at $P < 0.05$.

2.6. Functional annotation and integration

Biological functions were assigned to the differentially transcribed mRNAs and differentially expressed proteins as well as targets of differentially transcribed miRNAs based on the Reactome (Vastrik et al., 2007) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (Huntley et al., 2016; Kanehisa et al., 2016). Reactome pathway annotation was performed by BLAST (E-value of < 10^{-5}) using *C. elegans* orthologs, which were subsequently assigned to pathways using ReactomePA (in an R language environment)

ACCEPTED MANUSCRIPT

(Yu and He, 2016). The enriched Reactome pathway annotations were defined using a threshold of FDR of < 0.05 (Beyer-Hardwick method). By contrast, KEGG pathway annotation was conducted based on KEGG BLAST hits (E-value: $< 10^{-5}$) and corresponding KEGG Orthology (KO) terms (Mao et al. 2005); KO terms were then assigned to KEGG pathways and KEGG BRITE orthologous protein families by mapping KO terms to KEGG Orthology Based Annotation System (KOBAS) database (Xie et al. 2011). Enriched KEGG pathways were identified using a cut-off of $P < 0.01$ (Fisher's Exact test). KEGG functional enrichments of differentially transcribed mRNAs and miRNAs as well as differentially expressed proteins were integrated and displayed using the program FuncTree (Uchiyama et al., 2015).

3. Results

3.1. Transcriptome, miRNAome and proteome of L3s, xL3s and L4s of H. contortus

Three datasets (i.e. transcriptome, miRNAome and proteome) were produced for each L3, xL3 and L4 of *H. contortus*. For the transcriptome, a total of 9,754 mRNA transcripts were identified amongst > 700 million pair-end reads for all three developmental stages (Supplementary Table S1). The numbers of mRNA transcripts detected in the L3, xL3 and L4 stages were 9,492, 9,581 and 9,487, respectively (Supplementary Table S1). Small RNA libraries yielded > 400 million reads, with 88 miRNAs identified (Supplementary Table S2), of which *hco-miR-34* and *hco-miR-100* were recorded for the first time in *H. contortus*. Similar numbers of miRNAs were identified in L3 (n = 86), xL3 (n = 87) and L4 (n = 84) (Supplementary Table S2). Based on the prediction of 11,783 peptides, the total proteome representing L3, xL3 and L4 was inferred to contain 1,591 proteins (Supplementary Table S3). The numbers of proteins identified in each of the three stages were 1,302, 1,520 and 1,578, respectively (Supplementary Table S3).

3.2. Transcription and expression of genes among developmental stages

Hierarchical cluster analyses resulted in a division of each transcriptomic (mRNA and miRNA) and proteomic dataset into three distinct groups corresponding to developmental stage (L3, xL3 or L4) of *H. contortus* (Fig. 1A-C). The numbers of mRNAs, miRNAs and proteins transcribed/expressed in the three

ACCEPTED MANUSCRIPT

larval stages (L3, xL3 and L4) are shown in Venn diagrams (Fig. 1D-F). Here, most mRNAs (n = 9,092), miRNAs (n = 82) or proteins (n = 1,278) were shared by all larval stages studied, and small numbers (≤ 59) of transcripts or proteins were unique to particular stages.

Specifically, 35 mRNAs were transcribed exclusively in L3s, including those of G protein-coupled receptor (GPCR) domain-containing, zinc finger domain-containing, sperm coating protein (SCP) extracellular-domain or collagen triple helix repeat domain-containing protein-coding genes (Fig. 1D); one miRNA (i.e., *hco-miR-34*) was transcribed exclusively in xL3s (Fig. 1E); and 56 proteins, including heat shock proteins, peptidases, intestinal prolyl carboxypeptidases, cuticle collagen triple helix repeat domain-containing protein, SCP extracellular domain-containing protein and galectin, were expressed exclusively in L4s (Fig. 1F). Pairwise comparisons identified mRNAs that were differentially transcribed between L3 and xL3 (n = 1,151), between xL3 and L4 (n = 1,428) and between L3 and L4 (n = 2,756) stages (Supplementary Table S4). Pairwise comparative analysis identified miRNAs that were differentially transcribed between L3 and xL3 (i.e. *hco-miR-34* and *-5976*), between xL3 and L4 (i.e. *hco-miR-34*, *-252*, *-5888*, *-5901*, *-5939* and *-5976*), and between L3 and L4 (i.e. *hco-miR-252*, *-5888*, *-5901*, *-5932*, *-5939* and *-5976*) stages (Supplementary Table S5). Pairwise comparative analysis also identified proteins that were differentially expressed between L3 and xL3 (n = 62), between xL3 and L4 (n = 126), and between L3 and L4 (n = 238) stages (Supplementary Table S6).

3.3. Relationship between transcription and protein expression

In total, 16.1% of transcripts within the combined transcriptome of L3, xL3 and L4 stages of *H. contortus* represented expressed proteins. Specifically, proteins were identified for 1,276, 1,483 and 1,531 mRNA transcripts in L3, xL3 and L4, respectively (Supplementary Table S7), although there was no transcriptomic evidence for 18 proteins (Supplementary Table S7).

Levels of mRNA transcription and protein expression were positively correlated in each L3, xL3 and L4 (r = 0.39, 0.44 and 0.43), with a statistical significance of $P < 0.0001$ (Fig. 2A-C). Changes in mRNA transcription and protein expression between L3 and xL3 were negatively correlated (n = 1,098, r = -0.03, $P > 0.05$), whereas changes between xL3 and L4 (n = 1,363, r = 0.20, $P < 0.0001$) and between L3 and L4 (n = 1015, r = 0.30, $P < 0.0001$) stages were positively correlated (Fig. 2D-F). Specifically, the correlations in the

ACCEPTED MANUSCRIPT

differential transcription and expression (FC > 2 and FDR < 0.05) between L3 and xL3 (n = 9), between xL3 and L4 (n = 33) and between L3 and L4 (n = 91) stages were r = -0.55 (P > 0.05), 0.30 (P > 0.05) and 0.31 (P < 0.01), respectively (Fig. 2D-F).

3.4. Association between miRNAs and their gene targets

A total of 86 miRNAs were predicted to target 3,773 genes based on both seed pairing and site accessibility, and 6,686 mRNA-miRNA pairs were inferred (Supplementary Table S8). Of the predicted miRNA-mRNA pairs, 45.4% (L3), 45.3% (xL3) and 40.0% (L4) had negative Spearman's correlation coefficients, and 79.3% (L3), 85.5% (xL3) and 88.1% (L4) had negative Hoeffding's D measure scores, respectively (Supplementary Table S8). Specifically, 1,019 (27.0%) of the predicted targets were expressed as proteins, and similar percentages of proteins had a negative correlation with miRNAs in L3s (44.7%/83.7%), xL3 (44.6%/96.6%) and L4s (46.8%/99.6%) (Supplementary Table S9).

For predicted mRNA-miRNA pairs, changes in mRNA transcription and protein expression were compared in a pairwise manner between larval stages; there was a positive correlation in the level of transcription or protein expression between L3 and xL3 (n = 411, r = 0.03, P > 0.05), between xL3 and L4 (n = 538, r = 0.21, P < 0.0001), and between L3 and L4 (n = 430, r = 0.29, P < 0.0001). By contrast, there was a negative correlation in differential transcription and expression (FC > 2 and FDR < 0.05) between L3 and xL3 (n = 14, r = -0.60, P < 0.05), between xL3 and L4 (n = 16, -0.50, P >0.05), and between L3 and L4 (n = 16, r = -0.54, P < 0.05) stages (Fig. 2G-I).

3.5. Biological pathway analysis in different developmental stages

Two approaches were employed to explore the enriched biological functions of the differentially transcribed mRNAs and miRNAs as well as differentially expressed proteins.

First, the Reactome pathway annotation linked differentially transcribed mRNAs and proteins to multiple biological pathways. In L3s, enriched pathways related to the metabolism of fatty acid, triacylglycerol, ketone body, amino acids and derivatives as well as the citric acid (TCA) cycle (Fig. 3A); in xL3s, they related to the metabolism of amino acids and derivatives, the TCA cycle and HSF1-dependent

ACCEPTED MANUSCRIPT

transactivation (Fig. 3B); in L4s, they related exclusively to the metabolism of amino acids and derivatives (Fig. 3C). No functional enrichment was identified for the predicted targets of differentially transcribed miRNAs.

Second, KEGG pathway enrichment analysis revealed that the differentially transcribed and/or expressed genes were associated with biological categories including environmental information processing, genetic information processing, organismal systems and metabolism as well as diseases (Supplementary Tables S10-S12). In L3s, highly transcribed (mRNA) and expressed genes related to carbohydrate metabolism (TCA cycle), lipid metabolism (fatty acid degradation), xenobiotic biodegradation and metabolism, and energy metabolism (oxidation phosphorylation) and diseases (e.g., neurodegenerative, endocrine and/or metabolic processes) (Fig. 4A); highly transcribed protein-coding (exclusively) genes related to processes including nucleotide (purine) metabolism, steroid hormone biosynthesis, signalling molecules and interactions (neuroactive ligand-receptor interaction), and digestive and nervous systems. In xL3s, highly transcribed protein-coding genes were predominantly associated with carbohydrate metabolism (TCA cycle), lipid metabolism (fatty acid degradation and steroid hormone biosynthesis), and xenobiotic biodegradation and metabolism, whereas high miRNA transcription was linked to signal transduction (Rap1 signalling pathway) and endocrine system (peroxisome proliferator-activated receptors [PPAR] signalling) (Fig. 4B). In L4s, highly mRNA-transcribed and expressed genes were linked predominantly to transcription, translation, protein folding, sorting and degradation (protein processing in endoplasmic reticulum), amino acid metabolism and the digestive system (protein digestion and absorption) (Fig. 4C). In addition, highly transcribed miRNAs (n = 6) related to genes enriched for the Rap1 signalling pathway and the immune system, but the mRNA transcription and protein expression of these genes was low (Fig. 4B). Similarly, highly transcribed miRNAs (n = 6) related to genes involved in signalling molecules and interactions, but this was not the case for mRNAs (Fig. 4C).

ACCEPTED MANUSCRIPT

4. Discussion

Using transcriptomic, miRNA and proteomic datasets, we explored molecular alterations in the developmental switch from L3 to L4 of *H. contortus* in vitro. Key changes were seen in mRNA, miRNA and protein profiles, and plausible relationships were established between the transcription (mRNA and/or miRNA) and protein expression.

Qualitative and quantitative analyses performed in this study identified substantial molecular changes in the developmental switch of *H. contortus*. The molecular changes were reflected by the specific transcription and expression between developmental stages. Although most mRNAs, miRNAs and proteins detected were shared by all three larval stages, exclusive transcription or expression indicated specific biological features for individual stages. In particular, the specific transcription of GPCR domain-containing and zinc finger domain-containing protein-coding genes suggested important chemosensory activities in the L3 stage, as the roles of these molecules in signal transduction and transcriptional regulation have been reported in *C. elegans* (see Clarke and Berg, 1998; Bargmann et al., 2006; Kim et al., 2009; Procko et al., 2012). Similarly, the exclusive expression in L4 of peptidases, SCP and galectin, which associate with protein digestion and immune responses, suggest features characteristic of a parasitic stage, such as blood feeding and parasite-host interactions (Laing et al., 2013; Schwarz et al., 2013). Molecular changes were also reflected in different miRNA, mRNA and protein levels in the developmental stages of *H. contortus*. Although only seven miRNAs with significant transcriptional changes among the three developmental stages were identified, substantial changes in transcription and expression were detected for thousands of mRNAs and hundreds of proteins, respectively. In particular, the marked differential transcription and expression between the L3 and L4 stages of *H. contortus* indicated significant alterations during the in vitro developmental transition from the free-living to the parasitic stage of this parasite.

To further understand the molecular alterations, we tested the concordance between transcriptome and proteome. For 16.1% of total mRNA transcription, a positive correlation (r = 0.39 to 0.44) with protein expression in all three developmental stages of *H. contortus* was observed. In contrast, negative and positive correlations (r = -0.55 to 0.31) were seen for differential mRNA transcripts and proteins between developmental stages. These results are consistent with previous studies of eukaryotes (Csárdi et al., 2015; Edfors et al., 2016; Lahtvee et al., 2017), and might relate to technical aspects in protein discovery, but are

ACCEPTED MANUSCRIPT

likely explained by a predominance of post-transcriptional regulation (Csárdi et al., 2015). For example, the most studied post-transcriptional regulators, miRNAs, play key roles in translational repression and mRNA degradation (Kim, 2005; Filipowicz et al., 2008; Bartel, 2009; Morris and Mattick, 2014) and would explain the tightly controlled expression seen here in *H. contortus.*

Therefore, to understand whether miRNAs associate with the discrepancy between transcriptome and proteome, we explored the relationships among miRNAs, mRNAs and proteins. First, targets of miRNAs were predicted using two different algorithms, considering both base-pairing (TargetScan) and site accessibility (PITA), to increase the accuracy of prediction (Lewis et al., 2005; Kertesz et al., 2007). These prediction tools have been applied to other nematodes such as *C. elegans* and *Trichuris suis*, and have produced convincing results (Jan et al., 2011; Jex et al., 2014). Then, we used two independent parameters (using Spearman's correlation and Hoeffding's D measure; Fujita et al., 2009) to assess the reliability of target prediction. The reliability of the predictions was indicated by a strong negative correlation between levels of miRNAs and mRNAs (40.0%-88.1%), and between levels of miRNAs and proteins (44.6%-99.6%). The high correlation between miRNAs and proteins indicate substantial translational repression. This proposal is supported by the target predictions in this study, which were performed against 3'-untranslated region (UTR) sequences, and by limited protein expression of the predicted targets (27.0%). Furthermore, the consistent negative correlation (r = -0.60 to -0.50) between the differential mRNA transcription and differential protein expression indicates substantial post-transcriptional regulation.

Since there were substantial differences in the transcription of mRNA and miRNA as well as protein expression between the developmental stages of *H. contortus*, the next step was to clarify the biological meaning of these differences. Based on Reactome pathway analysis (for *C. elegans*), the enriched pathways inferred related to carbohydrate and fatty acid metabolism in the L3 stage, and the metabolism of amino acids and derivatives in the L4 stage. However, although *H. contortus* is closely related to *C. elegans* (see Blaxter et al., 1998), and the latter has been widely used as a model organism for parasitic nematodes (Politz and Philipp, 1992; Bürglin et al., 1998; Gilleard, 2004), *C. elegans* is a free-living nematode, restricting our ability to annotate molecules linked to parasitic biological processes (Geary and Thompson, 2001; Datu et al., 2008; Viney, 2017). This aspect was reflected in the limited number of Reactome pathway annotations, particularly in the parasitic L4 stage of *H. contortus*. Nevertheless, to address this issue, we also employed the KEGG database to assign biological meaning to the molecular changes, and, indeed, more enriched

ACCEPTED MANUSCRIPT

biological categories were predicted for differential mRNAs, miRNAs and proteins using KEGG than for the Reactome pathway analysis.

Based on the Reactome and KEGG pathway analyses, enriched fatty acid and carbohydrate degradation, TCA cycle, oxidation phosphorylation and neuroactive ligand-receptor interaction in L3s indicated the specific features (e.g., inability to feed, aerobic metabolism and active chemosensation) for this free-living stage (Nikolaou and Gasser, 2006; Cantacessi et al., 2010; Wang et al., 2016). These features also linked to the annotations of neural, endocrine and metabolic diseases, suggesting a dysregulated physiological and biochemical status, which accords with dauer in *C. elegans* (cf. Hu et al., 2007). The similarity between the L3 stage of some parasitic nematodes and the dauer stage of *C. elegans* has been extensively discussed (Beall and Pearce, 2002; Ogawa et al., 2009; Crook, 2014; Lok, 2016), and it has been proposed that the dauer stage is a pre-adaptation to parasitism (cf. Crook, 2014).

By contrast, down-regulated or suppressed pathways in the xL3 stage, such as fatty acid degradation, steroid hormone biosynthesis, PPAR signalling, hormone signal transduction and neuroactive ligand-receptor interaction, indicated switches in metabolism and changes in signalling in the developmental transition from L3 to L4 in vitro (cf. Bento et al., 2010; Lee and Schroeder, 2012; Lok, 2016; Butcher, 2017; Watts and Ristow, 2017). Additionally, the enriched pathways in the L4 stage, linked to amino acid metabolism, transcription, translation, protein processing in endoplasmic reticulum and pathogenesis, indicate biological processes for rapid development and parasitic features (e.g., blood feeding in vivo) (Nikolaou and Gasser, 2006; Laing et al., 2013; Schwarz et al., 2013). Notably, the present, integrated analysis showed that not all transcriptional enrichments translated into functional (protein) up-regulations. While such a discrepancy might be explained by limited coverage of the proteome, this is highly unlikely to be the case here. The most plausible explanation is a tight post-transcriptional regulation of gene expression. Taken together, the present findings reveal that the significant biological changes observed during the development from the L3 to the L4 stage of *H. contortus* in vitro relate to: (i) environmental information processing converting into genetic information processing; (ii) energy production switching from carbohydrate and lipid degradation to protein metabolism; and (iii) free-living molecular features adapting to parasitic traits.

The adaptation of *H. contortus* to its environment during its developmental transition in vitro appears to be tightly regulated by miRNA networks. Specifically, the enriched HSF1-dependent transactivation and xenobiotic biodegradation and metabolism pathways in the xL3 stage indicate stress responses (Dayalan

ACCEPTED MANUSCRIPT

Naidu and Dinkova-Kostova, 2017), which might be caused by exogenous stimuli such as NaClO, CO_2 and/or high temperature during artificial exsheathment; under this condition, the exclusive, high transcription of *hco-miR-34* in the xL3 stage strongly suggests a functional role in stress responses. The upregulated transcription of *miR-34* has also been reported in the dauer stage of *C. elegans* and the infective stage of *S. ratti* (see Ahmed et al., 2013; Isik et al., 2016), and might be involved in the morphological and metabolic adaptation through interactions with *daf-2* and *daf-16*, in response to environmental stresses (Isik et al., 2016). Additionally, the high transcription of *hco-miR-252* in the L4 stage might associate with pathogen-related responses, which accords with previous studies of *C. elegans* (see Kudlow et al., 2012; Sun et al., 2016). However, the functional roles of *hco-miR-34* and *hco-miR-252* need to be explored to understand why the Rap1 signalling pathway is suppressed in both xL3 and L4 stages, and what the downstream regulatory networks are. Moreover, *hco-miR-5888, -5901, -5932* and *-5976* are highly transcribed in parasitic stages, but nothing is yet known about the functional roles of these small RNAs in *H. contortus* (see Winter et al., 2012). These "novel" miRNAs might play crucial roles in the development of this parasite, and thus warrant detailed investigation.

In conclusion, the integrated transcriptomic, proteomic and bioinformatic investigation reported here has given detailed insights into molecular alterations during the developmental transition from the free-living to parasitic larval stages of *H. contortus* in vitro. Dominant post-transcriptional regulations were inferred, and miRNAs likely play critical roles during the in vitro developmental transition. This transition involves a range of biological changes associated with metabolism, signal transduction and genetic information processing. Specifically, *hco-miR-34* and *hco-miR-252* appear to play key roles in the molecular alterations, particularly in stress responses and environmental adaptation. Taken together, these findings significantly enhance our understanding of the developmental biology of this important parasite at the molecular level under well-defined conditions outside of the host. Future work should compare the molecular developmental transitions in vitro with those that occur in vivo within the host animal.

ACCEPTED MANUSCRIPT

Acknowledgements

Funding from the National Health and Medical Research Council (NHMRC) of Australia, the Australian Research Council, Melbourne Water Corporation and The University of Melbourne (BIP) is gratefully acknowledged (R.B.G.). Support from the Melbourne Bioinformatics Platform is gratefully acknowledged. P.K.K. holds an NHMRC Early Career Researcher Fellowship (ECRF), and N.D.Y. holds an NHMRC Career Development Fellowship (CDF1).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/XXXXX/j.ijpara.2018.XX. XXX.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

References

Ahmed, R., Chang, Z., Younis, A.E., Langnick, C., Li, N., Chen, W., Brattig, N., Dieterich, C., 2013. Conserved miRNAs are candidate post-transcriptional regulators of developmental arrest in free-living and parasitic nematodes. Genome Biol. Evol. 5, 1246–1260.

Andrés-León, E., Núñez-Torres, R., Rojas, A.M., 2016. miARma-Seq: a comprehensive tool for miRNA, mRNA and circRNA analysis. Sci. Rep. 6, 25749.

Bargmann, C.I., 2006. Chemosensation in *C. elegans*. WormBook 25, 1–29.

Bartel, D.P., 2009. MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233.

Beall, M.J., Pearce, E.J., 2002. Transforming growth factor-beta and insulin-like signalling pathways in parasitic helminths. Int. J. Parasitol. 32, 399–404.

Bento, G., Ogawa, A., Sommer, R.J., 2010. Co-option of the hormone-signalling module dafachronic acid-DAF-12 in nematode evolution. Nature 466, 494–497.

Blaxter, M.L., De Ley, P., Garey, J.R., Liu, L.X., Scheldeman, P., Vierstraete, A., Vanfleteren, J.R., Mackey, L.Y., Dorris, M., Frisse, L.M., Vida, J.T., Thomas, W.K., 1998. A molecular evolutionary framework for the phylum Nematoda. Nature 392, 71–75.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Bürglin, T.R., Lobos, E., Blaxter, M.L., 1998. *Caenorhabditis elegans* as a model for parasitic nematodes. Int. J. Parasitol. 28, 395–411.

Butcher, R.A., 2017. Small-molecule pheromones and hormones controlling nematode development. Nat. Chem. Biol. 13, 577–586.

Campbell, B.E., Nagaraj, S.H., Hu, M., Zhong, W., Sternberg, P.W., Ong, E.K., Loukas, A., Ranganathan, S., Beveridge, I., McInnes, R.L., Hutchinson, G.W., Gasser, R.B., 2008. Gender-enriched transcripts in *Haemonchus contortus*--predicted functions and genetic interactions based on comparative analyses with *Caenorhabditis elegans*. Int. J. Parasitol. 38, 65–83.

Cantacessi, C., Campbell, B.E., Young, N.D., Jex, A.R., Hall, R.S., Presidente, P.J., Zawadzki, J.L., Zhong, W., Aleman-Meza, B., Loukas, A., Sternberg, P.W., Gasser, R.B., 2010. Differences in transcription between free-living and CO_2-activated third-stage larvae of *Haemonchus contortus*. BMC Genomics 11, 266.

Clarke, N.D., Berg, J.M., 1998. Zinc fingers in *Caenorhabditis elegans*: finding families and probing pathways. Science 282, 2018–2022.

Corsi, A.K., Wightman, B., Chalfie, M., 2015. A transparent window into biology: a primer on *Caenorhabditis elegans*. WormBook 18, 1–31.

Crook, M., 2014. The dauer hypothesis and the evolution of parasitism: 20 years on and still going strong. Int. J. Parasitol. 44, 1–8.

Csárdi, G., Franks, A., Choi, D.S., Airoldi, E.M., Drummond, D.A., 2015. Accounting for experimental noise reveals that mRNA levels, amplified by post-transcriptional processes, largely determine steady-state protein levels in yeast. PLoS Genet. 11, e1005206.

Dayalan Naidu S, Dinkova-Kostova A.T., 2017. Regulation of the mammalian heat shock factor 1. FEBS J. 284, 1606–1627.

Datu, B.J., Gasser, R.B., Nagaraj, S.H., Ong, E.K., O'Donoghue, P., McInnes, R., Ranganathan, S., Loukas, A., 2008. Transcriptional changes in the hookworm, *Ancylostoma caninum*, during the transition from a free-living to a parasitic larva. PLoS Negl. Trop. Dis. 2, e130.

Doyle, S.R., Laing, R., Bartley, D.J., Britton, C., Umer, Chaudhry., Gilleard, J.S., Holroyd, N., Mable, B.K., Maitland, K., Morrison, A.A. Tait, A., Tracey, A., Berriman, M., Devaney, E., Cotton, J.A., Sargison, N.D., 2017. A genome resequencing-based genetic map reveals the recombination landscape of an outbred parasitic nematode in the presence of polyploidy and polyandry. Genome Biol. Evol. doi: 10.1093/gbe/evx269.

Edfors, F., Danielsson, F., Hallström, B.M., Käll, L., Lundberg, E., Pontén, F., Forsström, B., Uhlén, M., 2016. Gene-specific correlation of RNA and protein levels in human cells and tissues. Mol. Syst. Biol. 12, 883.

Filipowicz, W., Bhattacharyya, S.N., Sonenberg, N., 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat. Rev. Genet. 9, 102–114.

Fujita, A., Sato, J.R., Demasi, M.A., Sogayar, M.C., Ferreira, C.E., Miyano, S., 2009. Comparing Pearson, Spearman and Hoeffding's D measure for gene expression association analysis. J. Bioinform. Comput. Biol. 7, 663–684.

ACCEPTED MANUSCRIPT

Gasser, R.B., von Samson-Himmelstjerna, G., 2016. *Haemonchus contortus* and Haemonchosis – Past, Present and Future Trends. Academic Press, London.

Geary, T.G., Thompson, D.P., 2001. *Caenorhabditis elegans*: how good a model for veterinary parasites? Vet. Parasitol. 101, 371–386.

Gillan, V., Maitland, K., Laing, R., Gu, H., Marks, N.D., Winter, A.D., Bartley, D., Morrison, A., Skuce, P.J., Rezansoff, A.M., Gilleard, J.S., Martinelli, A., Britton, C., Devaney, E., 2017. Increased expression of a microRNA correlates with anthelmintic resistance in parasitic nematodes. Front. Cell. Infect. Microbiol. 7, 452.

Gilleard, J.S., 2004. The use of *Caenorhabditis elegans* in parasitic nematode research. Parasitology 128 Suppl 1, S49–S70.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

Grosshans, H., Filipowicz, W., 2008. Molecular biology: the expanding world of small RNAs. Nature 451, 414–416.

Gu, H.Y., Marks, N.D., Winter, A.D., Weir, W., Tzelos, T., McNeilly, T.N., Britton, C., Devaney, E., 2017. 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. 11, e0006056.

Howe, K.L., Bolt, B.J., Cain, S., Chan, J., Chen, W.J., Davis, P., Done, J., Down, T., Gao, S., Grove, C., Harris, T.W., Kishore, R., Lee, R., Lomax, J., Li, Y., Muller, H.M., Nakamura, C., Nuin, P., Paulini, M., Raciti, D., Schindelman, G., Stanley, E., Tuli, M.A., Van Auken, K., Wang, D., Wang, X., Williams, G., Wright, A., Yook, K., Berriman, M., Kersey, P., Schedl, T., Stein, L., Sternberg, P.W., 2016a. WormBase 2016: expanding to enable helminth genomic research. Nucleic Acids Res. 44, D774–780.

Howe, K.L., Bolt, B.J., Shafie, M., Kersey, P., Berriman, M., 2016b. WormBase ParaSite - a comprehensive resource for helminth genomics. Mol. Biochem. Parasitol. 215, 2–10.

Hu, P.J., 2007. Dauer. WormBook 8, 1–19.

Huntley, R.P., Sitnikov, D., Orlic-Milacic, M., Balakrishnan, R., D'Eustachio, P., Gillespie, M.E., Howe, D., Kalea, A.Z., Maegdefessel, L., Osumi-Sutherland, D., Petri, V., Smith, J.R., Van Auken, K., Wood, V., Zampetaki, A., Mayr, M., Lovering, R.C., 2016. Guidelines for the functional annotation of microRNAs using the Gene Ontology. RNA 22, 667–676.

Isik, M., Blackwell, T.K., Berezikov, E., 2016. MicroRNA mir-34 provides robustness to environmental stress response via the DAF-16 network in *C. elegans*. Sci. Rep. 6, 36766.

Jan, C.H., Friedman, R.C., Ruby, J.G., Bartel, D.P., 2011. Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. Nature 469, 97–101.

Jex, A.R., Nejsum, P., Schwarz, E.M., Hu, L., Young, N.D., Hall, R.S., Korhonen, P.K., Liao, S., Thamsborg, S., Xia, J., Xu, P., Wang, S., Scheerlinck, J.-P.Y., Hofmann, A., Sternberg, P.W., Wang, J., Gasser, R.B., 2014. Genome and transcriptome of the porcine whipworm *Trichuris suis*. Nat. Genet. 46, 701–706.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2016. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 44, D457–462.

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., Segal, E., 2007. The role of site accessibility in microRNA target recognition. Nat. Genet. 39, 1278–1284.

Kim, V.N., 2005. Small RNAs: classification, biogenesis, and function. Mol. Cell 19, 1–15.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36.

Kim, K., Sato, K., Shibuya, M., Zeiger, D.M., Butcher, R.A., Ragains, J.R., Clardy, J., Touhara, K., Sengupta, P., 2009. Two chemoreceptors mediate developmental effects of dauer pheromone in *C. elegans*. Science 326, 994–998.

Kozomara, A., Griffiths-Jones, S., 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 42, D68–73.

Kudlow, B.A., Zhang, L., Han, M., 2012. Systematic analysis of tissue-restricted miRISCs reveals a broad role for microRNAs in suppressing basal activity of the *C. elegans* pathogen response. Mol. Cell 46, 530–541.

ACCEPTED MANUSCRIPT

Lahtvee, P.J., Sánchez, B.J., Smialowska, A., Kasvandik, S., Elsemman, I.E., Gatto, F., Nielsen, J., Absolute quantification of protein and mRNA abundances demonstrate variability in gene-specific translation efficiency in yeast. Cell. Syst. 4, 495–504.

Laing, R., Kikuchi, T., Martinelli, A., Tsai, I.J., Beech, R.N., Redman, E., Holroyd, N., Bartley, D.J., Beasley, H., Britton, C., Curran, D., Devaney, E., Gilabert, A., Hunt, M., Jackson, F., Johnston, S.L., Kryukov, I., Li, K., Morrison, A.A., Reid, A.J., Sargison, N., Saunders, G.I., Wasmuth, J.D., Wolstenholme, A., Berriman, M., Gilleard, J.S., Cotton, J.A., 2013. The genome and transcriptome of *Haemonchus contortus*, a key model parasite for drug and vaccine discovery. Genome Biol. 14, R88.

Law, C.W., Alhamdoosh, M., Su, S., Smyth, G.K., Ritchie, M.E., 2016. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. F1000Res. 5, 1408–1427.

Le, T.D., Zhang, J., Liu, L., Liu, H., Li, J., 2015. miRLAB: an R based dry lab for exploring miRNA-mRNA regulatory relationships. PLoS One 10, e0145386.

Lee, S.S., Schroeder, F.C., 2012. Steroids as central regulators of organismal development and lifespan. PLoS Biol. 10, e1001307.

Lewis, B.P., Burge, C.B., Bartel, D.P., 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120, 15–20.

Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Genomics 12, 323.

Lok, J.B., 2016. Signaling in parasitic nematodes: physicochemical communication between host and parasite and endogenous molecular transduction pathways governing worm development and survival. Curr. Clin. Microbiol. Rep. 3, 186–197.

Mao, X., Cai, T., Olyarchuk, J.G., Wei, L., 2005. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics 21, 3787–3793.

Morris, K.V., Mattick, J.S., 2014. The rise of regulatory RNA. Nat. Rev. Genet. 15, 423–437.

Nikolaou, S., Gasser, R., 2006. Prospects for exploring molecular developmental processes in *Haemonchus contortus*. Int. J. Parasitol. 36, 859–868.

Ogawa, A., Streit, A., Antebi, A., Sommer, R.J., 2009. A conserved endocrine mechanism controls the formation of dauer and infective larvae in nematodes. Curr. Biol. 19, 67–71.

Piano, F., Gunsalus, K.C., Hill, D.E., Vidal, M., 2006. *C. elegans* Network Biology: A Beginning. WormBook 21, 1–20.

Politz, S.M., Philipp, M., 1992. *Caenorhabditis elegans* as a model for parasitic nematodes: a focus on the cuticle. Parasitol. Today 8, 6–12.

Preston, S., Jabbar, A., Nowell, C., Joachim, A., Ruttkowski, B., Baell, J., Cardno, T., Korhonen, P.K., Piedrafita, D., Ansell, B.R., Jex, A.R., Hofmann, A., Gasser, R.B., 2015. Low cost whole-organism screening of compounds for anthelmintic activity. Int. J. Parasitol. 45, 333–343.

Procko, C., Lu, Y., Shaham, S., 2012. Sensory organ remodeling in *Caenorhabditis elegans* requires the zinc-finger protein ZTF-16. Genetics 190, 1405–1415.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K., 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.

Schwarz, E.M., Korhonen, P.K., Campbell, B.E., Young, N.D., Jex, A.R., Jabbar, A., Hall, R.S., Mondal, A., Howe, A.C., Pell, J., Hofmann, A., Boag, P.R., Zhu, X.Q., Gregory, T., Loukas, A., Williams, B.A., Antoshechkin, I., Brown, C., Sternberg, P.W., Gasser, R.B., 2013. The genome and developmental transcriptome of the strongylid nematode *Haemonchus contortus*. Genome Biol. 14, R89.

Su, S., Law, C.W., Ah-Cann, C., Asselin-Labat, M.L., Blewitt, M.E., Ritchie, M.E., 2017. Glimma: interactive graphics for gene expression analysis. Bioinformatics 33, 2050–2052.

Sun, L., Zhi, L., Shakoor, S., Liao, K., Wang, D., 2016. microRNAs involved in the control of innate immunity in candida infected *Caenorhabditis elegans*. Sci. Rep. 6, 36036.

Tang, Y.T., Gao, X., Rosa, B.A., Abubucker, S., Hallsworth-Pepin, K., Martin, J., Tyagi, R., Heizer, E., Zhang, X., Bhonagiri-Palsikar, V., Minx, P., Warren, W.C., Wang, Q., Zhan, B., Hotez, P.J., Sternberg, P.W., Dougall, A., Gaze, S.T., Mulvenna, J., Sotillo, J., Ranganathan, S., Rabelo, E.M., Wilson, R.K., Felgner, P.L., Bethony, J., Hawdon, J.M., Gasser, R.B., Loukas, A., Mitreva, M., 2014. Genome of the human hookworm *Necator americanus*. Nat. Genet. 46, 261–269.

Tyagi, R., Joachim, A., Ruttkowski, B., Rosa, B.A., Martin, J.C., Hallsworth-Pepin, K., Zhang, X., Ozersky, P., Wilson, R.K., Ranganathan, S., Sternberg, P.W., Gasser, R.B., Mitreva, M., 2015. Cracking the

ACCEPTED MANUSCRIPT

nodule worm code advances knowledge of parasite biology and biotechnology to tackle major diseases of livestock. Biotechnol. Adv. 33, 980–991.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., Pachter, L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28, 511–515.

Uchiyama, T., Irie, M., Mori, H., Kurokawa, K., Yamada, T., 2015. FuncTree: functional analysis and visualization for large-scale omics data. PLoS One 10, e0126967.

van Assche, R., Broeckx, V., Boonen, K., Maes, E., De Haes, W., Schoofs, L., Temmerman, L., 2015. Integrating -omics: systems biology as explored through *C. elegans* research. J. Mol. Biol. 427, 3441–3451.

Vastrik, I., D'Eustachio, P., Schmidt, E., Joshi-Tope, G., Gopinath, G., Croft, D., de Bono, B., Gillespie, M., Jassal, B., Lewis, S., Matthews, L., Wu, G., Birney, E., Stein, L., 2007. Reactome: a knowledge base of biologic pathways and processes. Genome Biol. 8, R39.

Veglia, F., 1915. The anatomy and life-history of the *Haemonchus contortus* (Rud.). Rep. Dir. Vet. Res. 3–4, 347–500.

Viney, M., 2017. How can we understand the genomic basis of nematode parasitism? Trends Parasitol. 33, 444–452.

Wang, F., Xu, L., Song, X., Li, X., Yan, R., 2016. Identification of differentially expressed proteins between free-living and activated third-stage larvae of *Haemonchus contortus*. Vet. Parasitol. 215, 72–77.

Watts, J.L., Ristow, M., 2017. Lipid and carbohydrate metabolism in *Caenorhabditis elegans*. Genetics 207, 413–446.

Winter, A.D., Weir, W., Hunt, M., Berriman, M., Gilleard, J.S., Devaney, E., Britton, C., 2012. Diversity in parasitic nematode genomes: the microRNAs of *Brugia pahangi* and *Haemonchus contortus* are largely novel. BMC Genomics 13, 4.

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., Kong, L., Gao, G., Li, C.Y., Wei, L., 2011. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 39, W316–322.

Yu, G., He, Q.Y., 2016. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol. Biosyst. 12, 477–479.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

FIGURE LEGENDS

Fig. 1. Stage-specific transcriptome, microRNAome and proteome of *Haemonchus contortus*. Heat maps display transcription or expression profiles (log_2-transformed counts per million (CPM) or abundance) of (A) mRNAs (top 500), (B) miRNAs (top 80) and (C) proteins (top 500) representing the L3, exsheathed L3 (xL3) and L4 stages of *H. contortus*. Profiles for L3, xL3 and L4 clustered into three groups (blue, green and orange bars). Venn diagrams show the numbers of (D) mRNAs, (E) miRNAs and (F) proteins unique to the L3 (blue), xL3 (green) or L4 (orange) of *H. contortus*, and shared by two or three of these developmental stages (intersections).

Fig. 2. Relationships between mRNA transcription and protein expression levels in the L3, exsheathed L3 (xL3) and L4 stages of *Haemonchus contortus*. Levels of mRNAs (log_2-transformed counts per million, CPM) are related to protein levels (log_2 abundance). Correlations of the mRNA and protein levels are shown for (A) L3, (B) xL3 and (C) L4 stages; correlations of the changes (log_2-transformed fold change, FC) in mRNA and protein levels between (D) L3 and xL3, (E) xL3 and L4, and (F) L3 and L4 stages are shown; correlations of the changes in mRNA and protein levels of miRNA targets between (G) L3 and xL3, (H) xL3 and L4 and (I) L3 and L4 stages are displayed. The number of mRNA-protein pairs (*n*), correlation coefficients (*r*) and significance (*P*) are indicated. Dots and values for differentially transcribed/expressed molecules (*n*, *r* and *P*) are highlighted in blue.

Fig. 3. Enriched pathways (Reactome) of differential transcription and expression in *Haemonchus contortus*. Enriched pathways for differentially transcribed/expressed mRNAs and proteins in the (A) L3, (B) exsheathed L3 (xL3) and (C) L4 stages of *Haemonchus contortus*. Dot sizes indicate counts of significantly highly (fold change (FC) of > 2; false discovery rate (FDR) of < 0.05) transcribed/expressed mRNAs and proteins. The numbers of differentially transcribed/expressed mRNAs and/or proteins mapped to Reactome pathways are indicated. Statistical significance (false discovery rate, FDR) of enrichment is indicated in colour.

Fig. 4. Enriched biological functions (via Kyoto Encyclopedia of Genes and Genomes, KEGG) of differential transcription and expression in *Haemonchus contortus*. Enriched biological processes and associated pathways of differential mRNAs, microRNAs and proteins in the (A) L3, (B) exsheathed L3 (xL3) and (C) L4 stages of *Haemonchus contortus*. Annotations are illustrated by hierarchical biological categories, biological processes and KEGG pathways. Enriched annotations of highly transcribed mRNAs and miRNAs (fold change (FC) of > 2; false discovery rate (FDR) of < 0.05), or highly expressed proteins (fold change (FC) of > 2; false discovery rate (FDR) of < 0.05) are indicated by green, orange and purple dots, respectively, and listed. Dot sizes indicate counts of significantly highly transcribed/expressed mRNAs/proteins (see also Supplementary Tables S10, S11 and S12).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

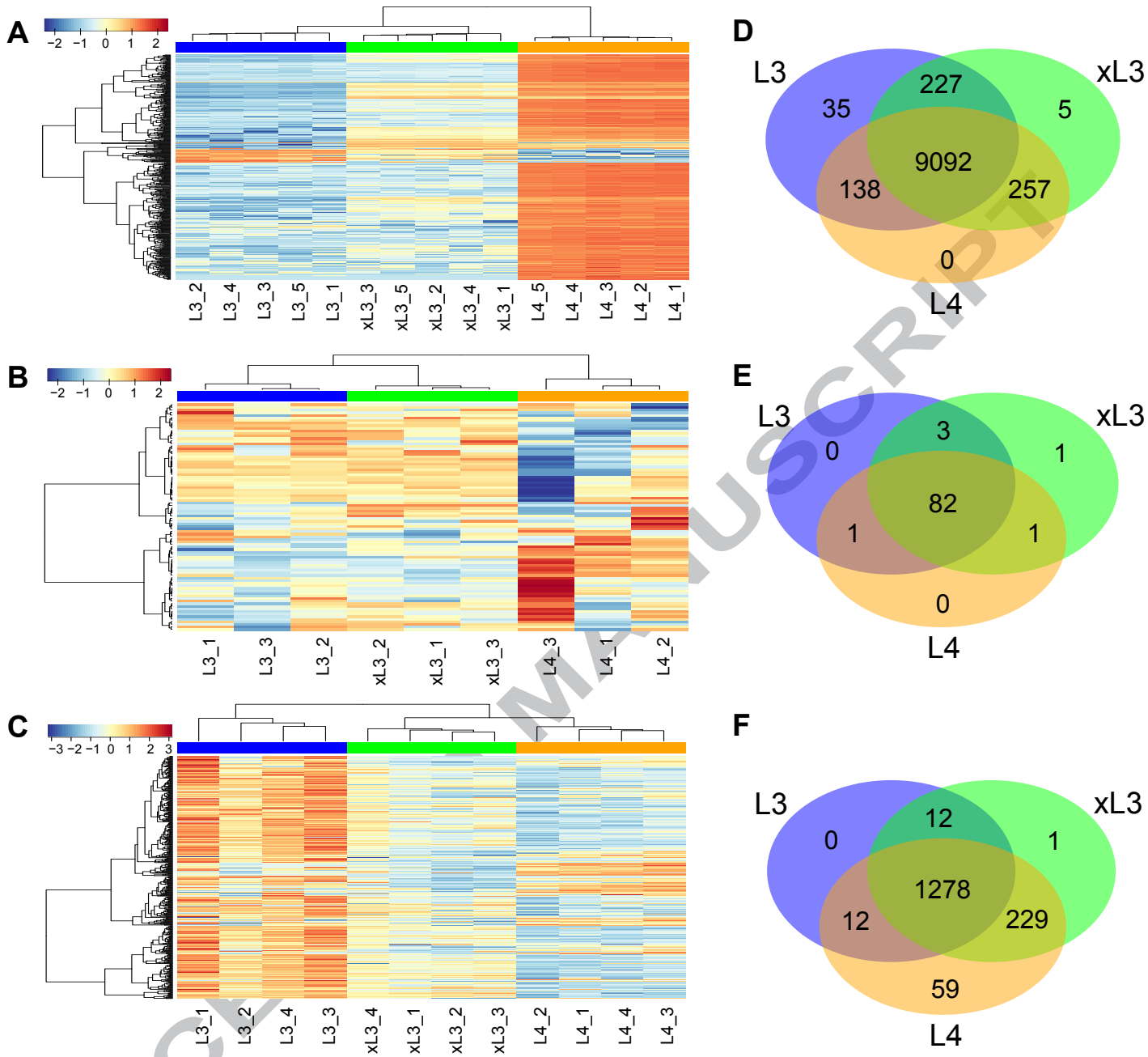
Highlights

- Integrated transcriptomic, proteomic and bioinformatic analysis
- Larval development of *Haemonchus contortus* in vitro
- Molecular alterations are under tight post-transcriptional control

ACCEPTED MANUSCRIPT

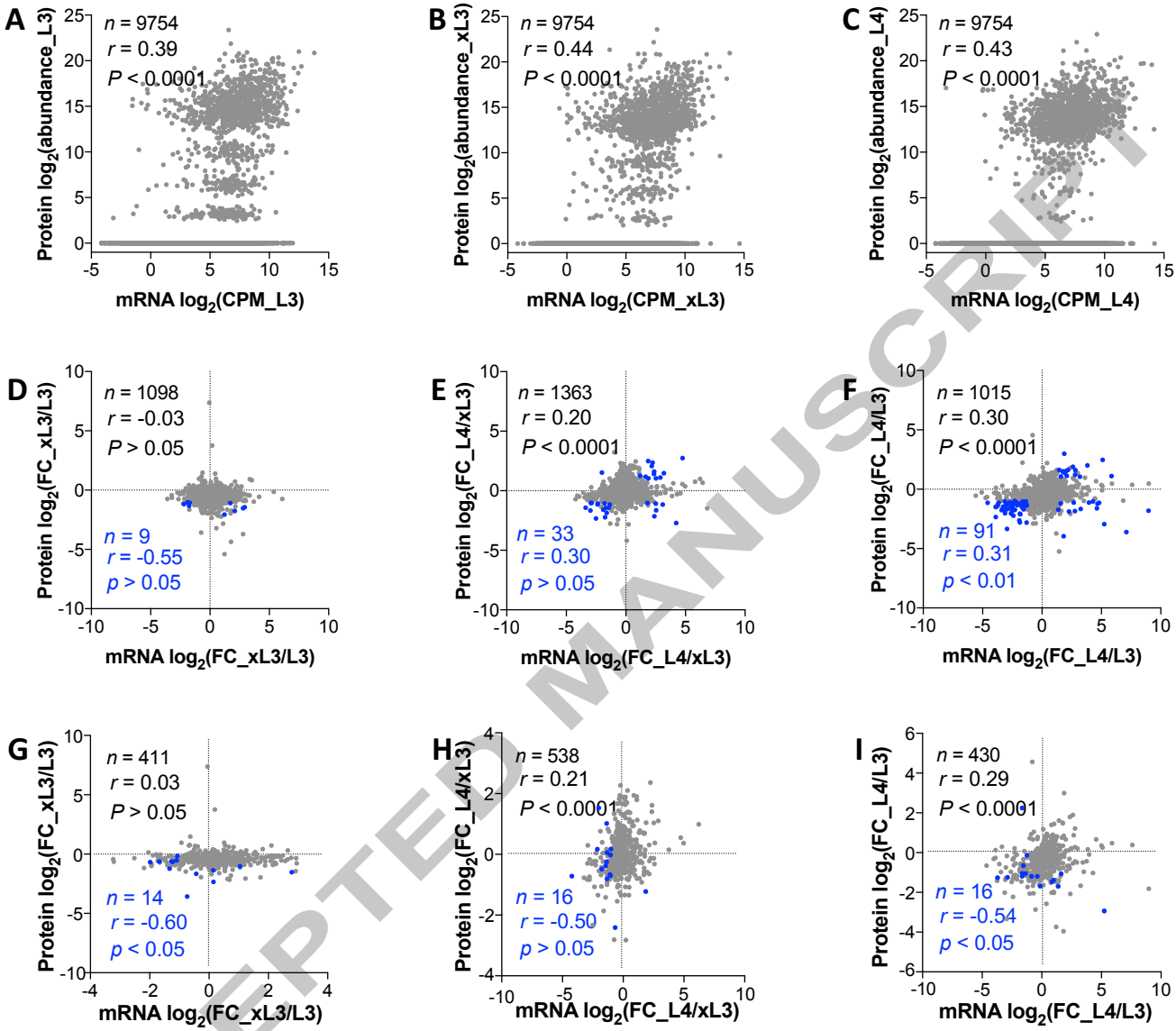
ACCEPTED MANUSCRIPT

Fig. 1 - Ma et al.



ACCEPTED MANUSCRIPT

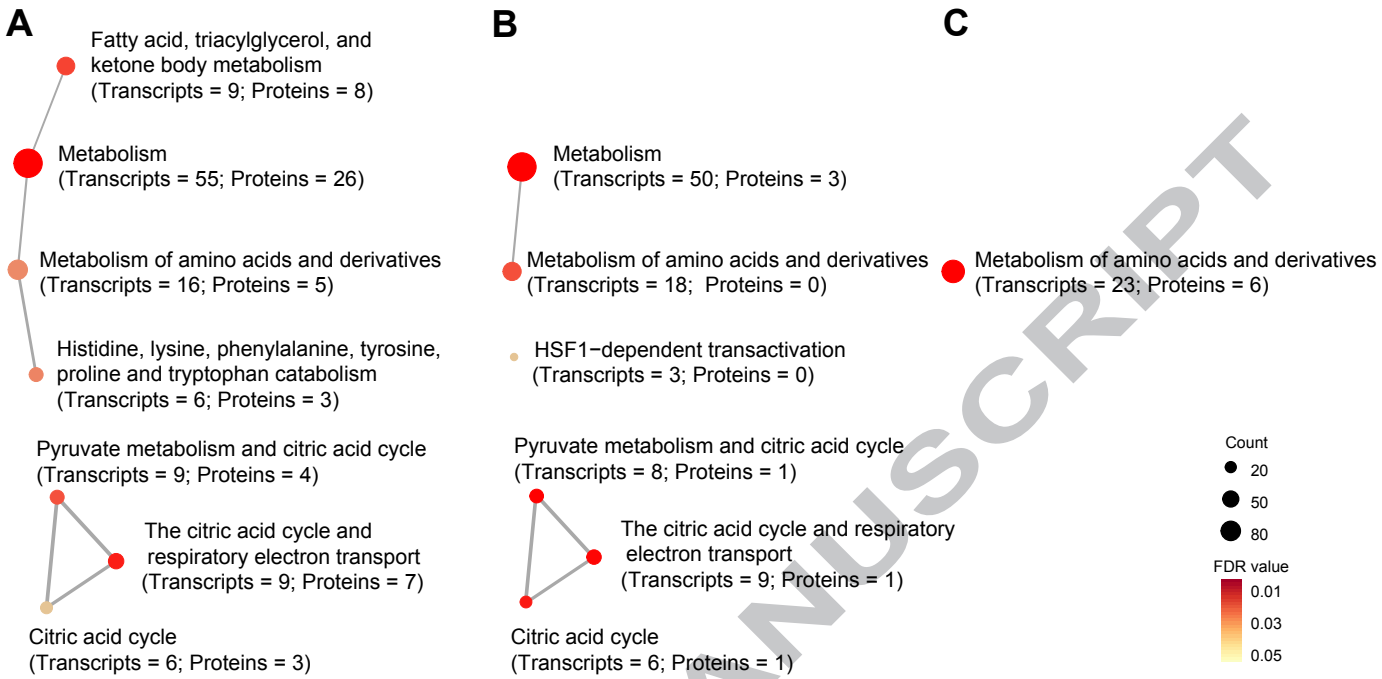
Fig. 2 - Ma et al.



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Fig. 3 - Ma et al.

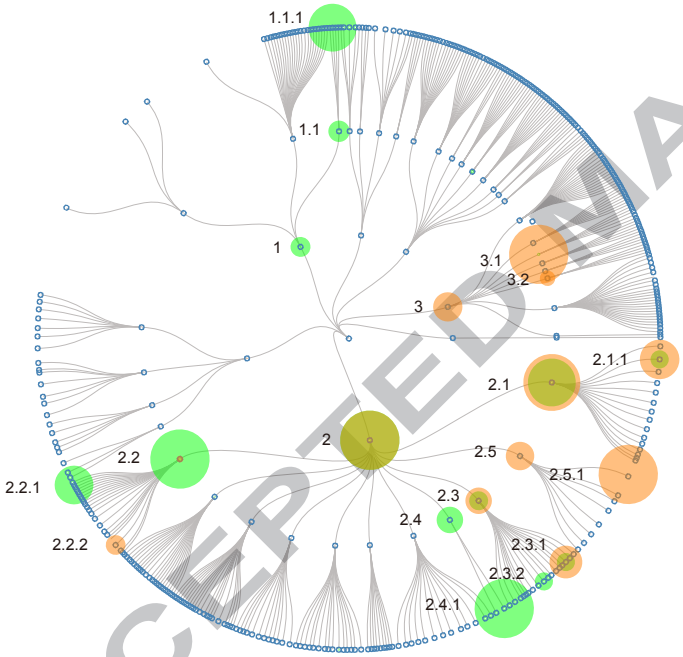


ACCEPTED MANUSCRIPT

Fig. 4 - Ma et al.

ACCEPTED MANUSCRIPT

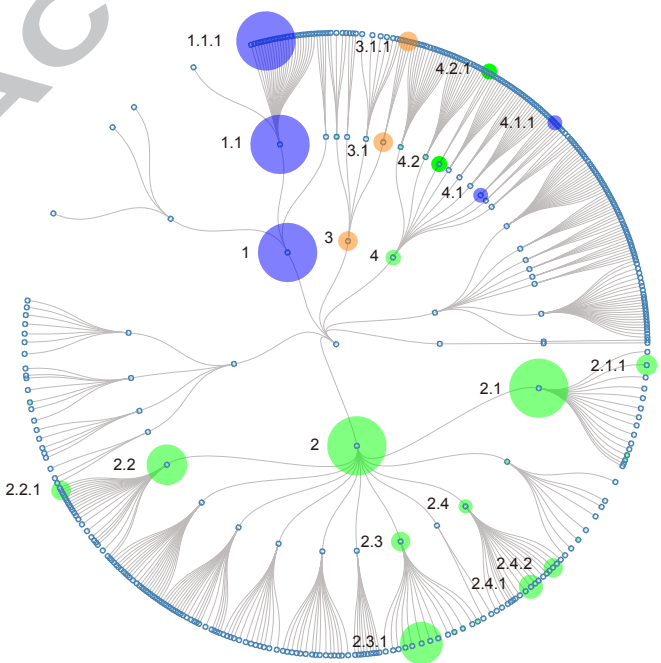
A



Highly transcribed mRNAs (FC > 2 and $P < 0.05$)
Highly transcribed miRNAs (FC > 2 and $P < 0.05$)
Highly expressed proteins (FC > 2 and $P < 0.05$)

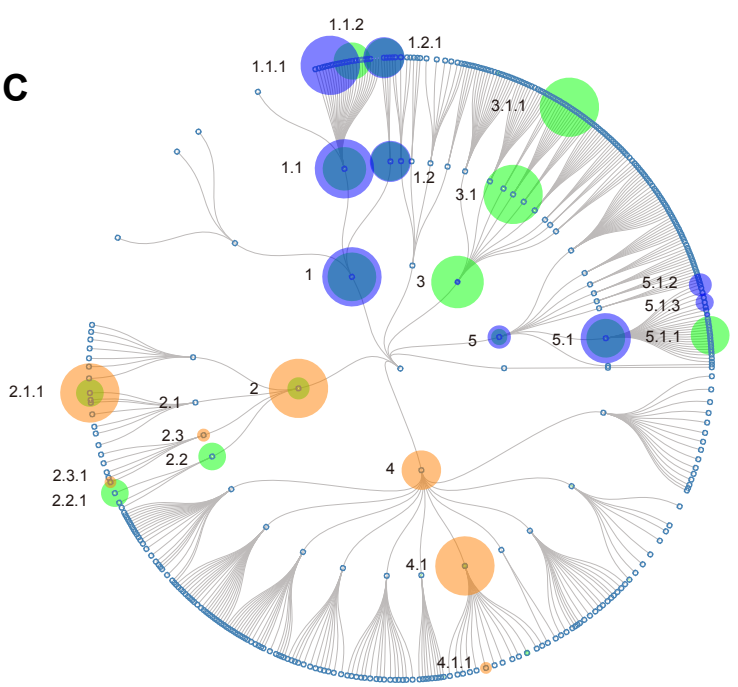
1. Environmental information processing
1.1. Signalling molecules and interaction
1.1.1. Neuroactive ligand-receptor interaction
2. Metabolism
2.1. Carbohydrate metabolism
2.1.1. Citrate (TCA) cycle
2.2. Xenobiotics biodegradation and metabolism
2.2.1. Drug metabolism
2.2.2. Benzoate degradation
2.3. Lipid metabolism
2.3.1. Fatty acid degradation
2.3.2. Steroid hormone biosynthesis
2.4. Nucleotide metabolism
2.4.1 Purine metabolism
2.5. Energy metabolism
2.5.1 Oxidation phosphorylation
3. Diseases
3.1. Neurodegenerative diseases
3.2. Endocrine and metabolic diseases

B



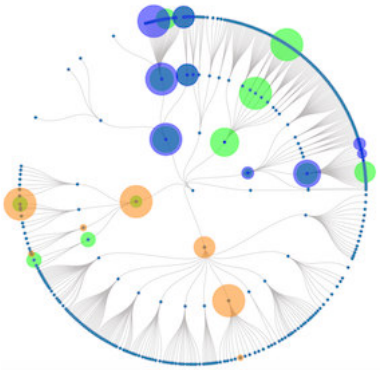
1. Environmental information processing
1.1. Signal transduction
1.1.1. Rap1 signalling pathway
2. Metabolism
2.1. Carbohydrate metabolism
2.1.1. Citrate (TCA) cycle
2.2. Xenobiotics biodegradation and metabolism
2.2.1. Drug metabolism
2.3. Aminoacid metabolism
2.3.1 Arginine and proline metabolism
2.4. Lipid metabolism
2.4.1. Steroid hormone biosynthesis
2.4.2. Fatty acid degradation
3. Cellular processes
3.1. Cell communication
3.1.1 Tight junction
4. Organismal systems
4.1. Sensory system
4.1.1. Phototransduction
4.2. Endocrine system
4.2.1. PPAR signalling pathway

C



1. Environmental information processing
1.1. Signal transduction
1.1.1. Rap1 signalling pathway
1.1.2. Hormone signal transduction
1.2. Signalling molecules and interactions
1.2.1. Neuroactive ligand-receptor interaction
2. Genetic information processing
2.1. Folding, sorting and degradation
2.1.1. Protein processing in endoplasmic reticulum
2.2. Transcription
2.2.1. Spliceosome
2.3. Translation
2.3.1. Aminoacyl-tRNA biosynthesis
3. Organismal systems
3.1. Digestive system
3.1.1. Protein digestion and absorption
4. Metabolism
4.1. Amino acid metabolism
4.1.1. Cysteine and methionine metabolism
5. Diseases
5.1. Infectious diseases
5.1.1. Amoebiasis
5.1.2. Pathogenic *Escherichia coli* infection
5.1.3. Bacterial invasion of epithelial cells

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT