Available online at www.sciencedirect.com
ScienceDirect
International Journal for Parasitology 38 (2008) 1329–1341
www.elsevier.com/locate/ijpara

# Microarray analysis of gender- and parasite-specific gene transcription in *Strongyloides ratti* ☆

Helen Evans [a], Luciane V. Mello [a], Yongxiang Fang [b], Ernst Wit [b], Fiona J. Thompson [c], Mark E. Viney [c], Steve Paterson [a,*]

[a] *School of Biological Sciences, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK*
[b] *Centre for Applied Statistics, University of Lancaster, Fylde College, Lancaster LA1 4YF, UK*
[c] *School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK*

Received 2 January 2008; received in revised form 17 February 2008; accepted 17 February 2008

**Abstract**

The molecular mechanisms by which parasitic nematodes reproduce and have adapted to life within a host are unclear. In the present study, microarray analysis was used to explore differential transcription among the different stages and sexes of *Strongyloides ratti*, a parasitic nematode of brown rats. Specifically, gender-biased transcription between free-living females and free-living males, and parasitic-biased transcription between parasitic females and free-living females was determined. Of the estimated 3,688 distinct transcripts represented on the microarray, 743 (20%) exhibited male-biased transcription of >1.4-fold ($2^{0.5}$), 689 (19%) female-biased transcription, 418 (11%) parasitic-biased transcription and 305 (8%) free-living-biased transcription. Among those transcripts that exhibited the highest levels of differential transcription, an orthologue of major sperm protein was identified in males, distinct aspartic protease orthologues in either parasitic or in free-living females, and orthologues of *hsp-17* chaperone in parasitic females. These 3,688 transcripts were separated into 12 clusters, such that the pattern of transcription between life-stages and biological replicates was similar among transcripts within a cluster and dissimilar between clusters. Using annotation inferred from *Caenorhabditis elegans*, gene ontology terms over-represented in one or more clusters were identified and showed that female-biased transcription was associated with genes involved in reproductive processes and larval development, male-biased transcription was linked to genes involved in metabolism, and free-living-biased transcription related to genes involved in the regulation of body fluids and response to external stimulus. The association of gene ontology with parasite-biased transcription was less clear. The present findings for *S. ratti* provide a basis for a detailed exploration of differentially regulated molecules and might assist in the search for novel drug or vaccine targets in parasitic nematodes.


*Keywords:* Reproduction; Spermatogenesis; Oogenesis; Transcriptomics; Fecundity; Aspartyl protease; Cathepsin

## 1. Introduction

For parasitic nematodes with a direct life-cycle, successful transmission is generally dependent on their ability to reproduce and release large numbers of infective stages into the environment to ensure successful infections of new hosts (Roberts, 1995; Poulin, 2007). One of the first steps in understanding the molecular basis of reproduction in parasitic nematodes is to identify genes with specific roles in male and female reproduction (Justine, 2002). One route to identifying such gender-specific genes is to contrast the transcriptomes of male and female worms in order to define genes exhibiting differential patterns of transcription between the sexes. The resources to perform transcriptomic analyses in parasitic nematodes are now becoming available and should provide powerful means to identify those genes linked to reproductive processes and so, potentially,

☆ *Note:* Nucleotide sequence data reported in this paper are available in the GenBank™, EMBL and DDBJ databases under the Accession Nos. FC809424–FC822028. Microarray data are available from ArrayExpress, accession number E-MAXD-37.

* Corresponding author. Tel.: +44 151 795 4521; fax: +44 151 795 4408.
*E-mail address:* s.paterson@liv.ac.uk (S. Paterson).

0020-7519/$34.00 © 2008 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijpara.2008.02.004

identify novel drug or vaccine targets to block transmission (Nisbet et al., 2004; Cottee et al., 2006; Campbell et al., 2008).

A point of reference for most genetic studies of parasitic nematodes is the free-living nematode *Caenorhabditis elegans* (see Bürglin et al., 1998; Blaxter, 2003). However, its utility for examining gender-specific patterns of transcription is limited in that adults exist only as hermaphrodites and males, i.e. there are no naturally occurring females, although it is recognised that the hermaphrodite essentially acts as a female (LaMunyon and Ward, 1998; Hope, 1999; LaMunyon and Ward, 1999). Thus, comparisons of males against hermaphrodites using a microarray representing ~18,000 genes (Reinke et al., 2004) indicated that 7% of *C. elegans* genes were upregulated in males and 11% upregulated in hermaphrodites. Much less is known about gender-specific differences in gene transcription in parasitic nematodes. Complementary DNA (cDNA) libraries constructed for *Trichostrongylus vitrinus* and *Oesophagostomum dentatum* derived by suppressive subtractive hybridisation (SSH) (Nisbet and Gasser, 2004; Cottee et al., 2006), have been reported to contain a number of either male- or female-specific transcripts, with gender-specific transcription confirmed using small-scale microarrays. Recently, Campbell et al. (2008) used an oligonucleotide array to identify differential transcription between males and females in 301 of 1,885 putative genes from *Haemonchus contortus*. Li et al. (2005) also used a microarray representing 3,569 expressed sequence tags (ESTs) from *Brugia malayi* (a filarial nematode) to identify 1,170 genes with gender-specific transcription. While these studies of *T. vitrinus*, *O. dentatum*, *H. contortus* and *B. malayi* describe genes with orthologues in *C. elegans* that are also involved in reproduction, a sizeable proportion of the genes found in parasitic nematodes appear not to have a *C. elegans* orthologue (cf. Parkinson et al., 2004). This latter finding would argue for studies in a wider range of parasitic nematode taxa: first, because nematodes are a diverse phylum and so there may be significant differences in the sets of genes involved in reproduction between different taxa; and, second, because many of the genes of most relevance to understanding the biology of parasitic nematodes are those which are unique to parasitic nematodes.

Another key difference between *C. elegans* and most directly transmitted nematodes is that adaptation to a parasitic mode of existence involves the ability to survive at different points in the life-cycle, both within a host and in the external environment. These contrasting environments differ substantially, particularly in temperature, in the availability of oxygen and nutrients and in the types of abiotic and biotic stressors which they contain (Viney, 1999). Understanding the molecular basis of how parasitic nematodes adapt to environments within and outside the host is challenging, but is particularly relevant for the development of new drugs or vaccines targeted against parasitic stages.

The genus *Strongyloides* (order Rhabditida, family Strongyloididea) contains a number of species of medical or veterinary importance. *Strongyloides stercoralis* infects 30–100 million humans worldwide (Bethony et al., 2006). These infections may often be asymptomatic, but can lead to disseminated infection and acute symptoms, particularly in immunocompromised individuals (Lewthwaite et al., 2005). *Strongyloides westeri* and *Strongyloides ransomi* infect horses and pigs, respectively, and *Strongyloides papillosus* has a host range covering sheep, goats and cattle (Anderson, 2000). Species of *Strongyloides* are unusual in that the parasitic stages are female only, and produce eggs by mitotic parthenogenesis (see Viney, 1994, 1999). These eggs can either develop directly into infective L3s (iL3s) or into free-living males and females, which mate and produce eggs that subsequently develop into iL3s (Fig. 1). This unusual life-cycle affords the opportunity to study molecular differences between (i) free-living males and free-living females to infer biological differences due to gender and (ii) free-living females and parasitic females to predict biological differences due to adaptations to life within and outside of a host. In the present study, a microarray analysis was performed for parasitic females, free-living females and free-living males of *Strongyloides ratti*, which is a natural parasite of brown rats (*Rattus norvegicus*) and a commonly used laboratory model for *Strongyloides* infection (Dawkins, 1989).

## 2. Materials and methods

### 2.1. Parasites

A laboratory line (LIV4) of *S. ratti* was used for all analyses. This line was generated by crossing lines ED248 and ED321 (Paterson and Viney, 2003; Paterson and Barber, 2007) and maintained by passage of 500 heterogonic iL3s (i.e. those produced by free-living adults) in female Wistar rats (Charles River, UK) every 5 weeks. To generate parasite material for the microarray analyses, eight rats (split into two groups of four) were infected with 500 iL3s on day 0 and a further 500 iL3s on day 3 p.i. On day 8 p.i., the faeces were collected from each rat and cultured at

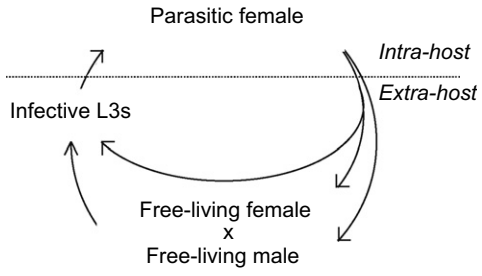


Fig. 1. Life-cycle of *Strongyloides ratti*. Eggs produced by mitotic parthenogenesis by parasitic females can develop either directly into infective L3s by homogonic development or into free-living males and females, which reproduce sexually to produce eggs that develop into infective L3s.

19 °C (Paterson and Viney, 2003). Rats were then starved overnight and culled on day 9 p.i. All procedures were conducted under UK Home Office licence and with approval from the University of Liverpool Ethics Committee. Animal care was provided in accordance with UK Home Office regulations. The small intestines were dissected from the rats, slit longitudinally, suspended in 100 ml measuring cylinders filled with RPMI (Gibco, UK) for 1 h at 37 °C, and then transferred to fresh RPMI and incubated for another hour. Parasitic females were then picked by pipette, washed in fresh RPMI and stored at −20 °C in RNAlater (Qiagen, Crawley, UK). On day 12 p.i., 200 free-living males and 200 free-living females from each rat were isolated from the cultures. These adults were 4 days old, which corresponded to the beginning of egg production in free-living females, when incubated at 19 °C. These adults were washed in distilled water and stored at −20 °C in RNAlater as before.

### 2.2. Microarray approach

A microarray was constructed from a set of 14,000 cDNA clones for which single-pass, unidirectional ESTs had previously been obtained (Thompson et al., 2005). These clones comprised a mixture of parasitic and free-living stages. In order to reduce the redundancy of the transcript sequences represented on the array, clones were clustered by using Phred and CAP3 (described in Section 2.4) to yield 5,156 'contigs'. A contig was defined as a transcript sequence distinct from other such sequences and may be formed from either the sequence of a single EST or from the consensus sequence from a group of ESTs. Therefore, each contig potentially represented different genes, splice variants or, in a few cases, non-overlapping fragments of the same transcript. For contigs consisting of more than 100 cDNA clones, 10 were randomly selected to be represented on the array. For all other contigs, up to five clones were randomly selected. In total, 7,773 unique cDNA clones, comprised of 1,909 (24.6%) free-living L1s; 1,840 (23.7%) free-living L2s; 691 (8.9%) mixed free-living adults and iL3s; 2,975 (38.3%) parasitic females at day 6 p.i. and 358 (4.6%) parasitic females at day 15 p.i. (Thompson et al., 2005) were subjected to 45 cycles of PCR amplification using the following cycling conditions: 95 °C for 5 s, 45 °C for 30 s, 72 °C for 1 min. For each clone, 75μl reactions were used, each containing 1× Tricine buffer (500 mM KCl, 300 mM Tricine pH 8.0, 20 mM $MgCl_2$), 0.2 mM dNTP, 0.2 μM forward and reverse primers (SP6 and T7 or TOPO) (Sigma–Aldrich, Dorset, UK), and 0.65 U *Taq* polymerase (Bioline Ltd., London). Aliquots of individual amplicons were checked for integrity on a 1% agarose gel. To make the microarray print plates, 100–200 ng of each product was transferred to 4 mm, 384-well plates (Genetix, New Milton Hampshire), dried completely and resuspended in 12 μl of spotting buffer (50% glycerol, 0.05% Triton X). Finally, amplicons were spotted onto GAPS II coated glass slides (Corning Inc., NY, USA) using an AJ100 Inkjet Microarray Spotter (Arrayjet Limited, Dalkeith, Scotland). Following construction of the microarray, all 7,773 clones were re-sequenced bi-directionally, both to verify the identity of each clone and to gain extra sequence data (Accession Nos. FC809424–FC822028). Sequences were clustered as before and each cDNA clone assigned to a contig. In most cases, forward and reverse sequences formed part of the same contig. For cases where forward and reverse sequences formed part of different contigs, the cDNA clone was assigned to the contig with the longest sequence. On this basis, 7,773 cDNA clones were assigned to 3,688 contigs.

Total RNA was extracted from *S. ratti* using TRIzol reagent (Invitrogen Life Technologies, Paisley, UK) and then purified using RNeasy MinElute spin columns (Qiagen, Crawley, UK) following the manufacturer's protocols. RNA concentration was estimated using the NanoDrop® ND-1000 Spectrophotometer (NanoDrop Technologies Inc., DE, USA), and the 260/280 and 260/230 ratios were calculated to assess RNA purity. Two hundred nanograms of each sample was dried to a final volume of 2 μl, and then subjected to one round of RNA amplification using the TargetAmp™ 1-Round Aminoallyl-aRNA Amplification Kit 101 (Epicentre Biotechnologies, WI, USA). RNA was quantified and the quality assessed using the 2100 Bioanalyser (Agilent Technologies, Wokingham, UK).

For microarray hybridisation, 5 μg of aminoallyl-aRNA was dried completely and resuspended in 5 μl of 0.1 M sodium bicarbonate. Samples were coupled with either Cy3 or Cy5 mono-reactive dyes (GE Healthcare, Buckinghamshire, UK) in the dark for 1 h and then purified using RNeasy MinElute spin columns (Qiagen, Crawley, UK). One micrograms of Cy3-labelled aRNA was combined with 1 μg of Cy5-labelled aRNA according to a loop design, such that samples originally derived from the same rat were hybridised within the same loop (Kerr, 2003). This approach was taken to minimise any effect of positive correlations between samples derived from the same host animal or culture. Combined samples were fragmented at 70 °C for 15 min using fragmentation buffer (Ambion, TX, USA) and then added to 90 μl Corning GAPS II hybridisation buffer (3× SSC [150 mM sodium chloride, 15 mM sodium citrate, pH 7], 0.1% SDS, 0.1 mg/ml BSA). Arrays were hybridised in chambers (Genetix, Hampshire, UK) at 50 °C for 18 h.

Following hybridisation, the arrays were washed once with 2× SSC, 0.1% SDS at 42 °C for 5 min, twice with 0.1× SSC, 0.1% SDS at room temperature for 5 min, four times with 0.1× SSC at room temperature for 1 min, and once with 0.01× SSC at room temperature for 1 min. Hybridised microarrays were scanned with an Axon 4000B scanner using the GenePix 3.1 analysis software (Axon Instruments, Foster City, CA, USA) at 532 and 635 nm for Cy3 and Cy5, respectively. Data from each probe was then extracted and subjected to background correction using the BlueFuse computer program (BlueGnome

Ltd., Cambridge, UK). A total of 38 hybridisations (15 parasitic females, 12 free-living females and 11 free-living males) from 19 slides were analysed (ArrayExpress accession E-MAXD-37).

### 2.3. Microarray validation by quantitative PCR (qPCR)

To validate the microarray results, transcripts represented by nine test contigs were selected for qPCR analysis. Specifically, five contigs (1092-1, 1407-1, 1501-1, 1654-1 and 2927-1) found to show gender- or parasite-specific transcription by microarray analysis, and four contigs (351-1, 1410-1, 1847-1, 2048-1) found to show no change in transcription were analysed. Two reference contigs were selected for normalisation: 2079-1, which encodes a small ribosomal subunit S20 protein, and 568-1, which encodes an actin-4 protein. Neither reference contig was found to vary in transcription upon microarray analysis. Four rats were infected and free-living males, free-living females and parasitic females collected, and total RNA extracted as before. For each RNA sample, 500 ng was reverse transcribed using SuperScript III Reverse Transcriptase (Invitrogen Technologies, Paisley, UK) and 100 ng random primers following the manufacturer's protocol.

Oligonucleotide primers were designed for qPCR using the Primer Express™ program (Applied Biosystems, Warrington, UK) to amplify a 60–90 bp insert with an annealing temperature of 60 °C (Supplementary Table S1). To assess the efficiency of each primer pair, standard curves were produced from a 5-fold dilution series of *S. ratti* cDNA. Briefly, 10 μl reaction volumes contained 1× Power SYBR® Green PCR Master Mix (Applied Biosystems, Warrington, UK), 0.1 μM of each primer and 3 μl of cDNA. All reactions were performed in triplicate, and all reactions representing a given contig were placed on the same plate. qPCR amplification was performed using a 7500 Fast cycler (Applied Biosystems) under the following conditions: an initial denaturation at 95 °C for 10 min, then 44 cycles of denaturation at 95 °C for 5 s and annealing/extension at 60 °C for 30 s, during which the fluorescence produced by SYBR® Green was recorded. The efficiency of each primer pair was calculated from the slope of the standard curves for cycle threshold ($C_t$) versus relative concentration using the equation, $E = 10^{-1/\text{slope}}$. For the qPCR assay, 1/100 dilutions of each of the free-living female, free-living male and parasitic female cDNA were amplified in triplicate and relative transcription ratios calculated according to the equation $\log_2 \text{ratio} = \Delta C_{t_{\text{test}}} \log_2 E_{\text{test}} - \Delta C_{t_{\text{reference}}} \log_2 E_{\text{reference}}$ (Pfaffl, 2001). Contig 2079-1 was used as the reference contig since initial results indicated highly similar levels of transcription using either 2079-1 or 568-1 as a reference.

### 2.4. Data analysis

Sequencing "base-calling" was performed using Phred 1.7, and vector sequences were masked by Cross_match (http://www.phrap.org/). Default criteria for vector masking were used, except that the minimum length of matching sequence was seven and the minimum alignment score was 20. A sequence was judged not to have a poly-A tail if its longest continuous sequence of A's was <10 bases in length. Reads containing <40 high-quality bases after trimming vector segments were considered as low-quality reads and discarded. If the 5′ and 3′ ends of a sequence were given a quality score <15 they were classed as low-quality regions. ESTs were defined as high repeat ESTs if the content of any base was >80% of the sequence. A two-step clustering procedure was carried out using the sequence alignment program CAP3 (Huang and Madan, 1999) to assemble ESTs first into groups with weak sequence similarity (>40 bps overlap and 85% identity), and then in a subsequent round with more stringent clustering (>40 bps overlap and 90% identity) into contigs.

Similarity searches (BLASTx) (Altschul et al., 1997) were conducted with the consensus sequence of each contig, using an *e*-value cut-off of e-10, against the following series of databases: *C. elegans* peptides, wormpep181 (ftp.sanger.ac.uk/pub/wormbase); UniProtKB/Swiss-Prot, unip_sw_53.2 (www.uniprot.org); nr (www.ncbi.nlm.nih.gov); conserved domain database (CDD) (www.ncbi.nlm.nih.gov); and 5′ and 3′ untranslated region (UTR) (www.ba.itb.cnr.it/UTR). The databases used were those available during October 2007. If necessary, clones belonging to the 10 contigs found to exhibit the highest levels of differential transcription in each of female, male, parasitic and free-living stages were re-sequenced to ensure that at a minimum of 600 bp of sequence were available. The results for each contig from each database can be found stored under EnvBase entry egcat000053, (nebc.nox.ac.uk/cgi-bin/public_catalogue.cgi). Putative gene ontology (GO) annotation for each contig was assigned where possible using the most significant alignment in the *C. elegans* peptide database. The Entrez Gene identifiers (www.ncbi.nlm.gov/sites/entrez) corresponding to these peptides inferred from *C. elegans* were used to map contigs to GO terms using the GOPkgBuilder function within the AnnBuilder package in R (www.bioconductor.org).

Spot intensities across the microarrays were normalised using the vsn package in R (www.r-project.org), followed by dye-bias correction using loess and spatial normalisation using 2-D loess in Matlab (www.mathworks.com). Differential transcription was estimated from a mixed effect model, $Y = X\delta + \varepsilon$, given $Y$, the vector of observed log ratios; $X$, the design matrix; $\delta$, a vector having elements for the dye effect and the two independent contrasts (free-living female versus parasitic female and free-living female versus free-living male); and $\varepsilon$, the error term such that $\varepsilon \sim N(0, \Sigma)$ given the covariance matrix $\Sigma$ describing both the random noise and the correlation between the transcription of a gene between free-living males, free-living females and parasitic females taken from the same host.

Contigs exhibiting similar patterns of transcription among life-stages and replicates were identified by per-

forming *k*-means clustering (Wit and McClure, 2004), assuming 12 clusters. The mean fold change in transcription for each contig was calculated from the individual values for each clone. Directed acyclic graphs representing the hierarchical structure of the GO terms present on the arrays were constructed using the GOstats package in R (www.bioconductor.org). Cumulative hypergeometric tests were then performed to identify GO terms that were significantly over-represented in each cluster. Only GO terms that mapped to 10 or more contigs on the array were used. The set of unique Entrez Gene identifiers within each cluster were used to prevent double counting of contigs that may represent different sequences derived from the same transcript. An unadjusted *P*-value of <0.01 was taken as being indicative of significant over-representation (correction for multiple comparisons was not possible for this analysis, since GO terms are grouped in hierarchies and hence non-independent).

## 3. Results

### 3.1. Microarray analysis

The aim of this study was to identify genes exhibiting differential transcription between free-living males and free-living females or between parasitic females and free-living females, since these two comparisons represented male/female and parasitic/free-living contrasts. Accordingly, an estimated 3,688 transcripts (i.e. distinct contig sequences) were studied by microarray analysis using mixed models with fixed effects for dye bias, for the male/female and parasitic/free-living contrasts, and random effects for different hosts. For differential transcription between free-living females and free-living males (Table 1), 252 contigs exhibited >2-fold transcription in males and 230 contigs exhibited >2-fold transcription in females (false discovery rate <0.035). For differential transcription between parasitic females and free-living females, 71 contigs exhibited >2-fold transcription in parasitic females and 111 contigs exhibited >2-fold transcription in free-living females (false discovery rate <0.050). Approximately half of the contigs (1,869 of 3,688) exhibited less than a 1.41-fold ($2^{0.5}$) difference in both the male/female and the parasitic/free-living contrasts. The 10 contigs that exhibited the greatest fold changes in transcription between the male/female and parasitic/free-living contrasts are shown in Table 2. Contig 121-1, which was upregulated in free-living males, exhibited similarity to a motif associated with major sperm protein (MSP) in *C. elegans*. In parasitic females, contigs 352-1, 429-1, 429-2 and 3564-1 exhibited homology to *C. elegans hsp-17*, which encodes a small heat-shock protein. An alignment of the sequences for these contigs suggested that 429-1 and 429-2 represented the same transcript, but that 352-1 and 3564-1 each represented different transcripts. This indicated that there were three distinct paralogues and/or splice variants for the *hsp-17* orthologue in *S. ratti* parasitic females. Contigs upregulated in parasitic females (933-1) or upregulated in free-living stages (1694-1) aligned to different aspartic protease proteins from *C. elegans*.

### 3.2. qPCR analysis

To validate the transcription levels determined by microarray analysis, nine contigs were selected for qPCR analysis (Fig. 2). Those contigs that displayed differential transcription in the microarray analysis also displayed differential transcription using qPCR. This demonstrated that the microarray analysis reliably detected individual genes showing gender-biased or parasite-biased transcriptional profiles. However, and in common with other studies (Hsiao et al., 2002), the qPCR results indicated that the sensitivity range of the microarray was limited relative to qPCR due to signal saturation. Thus, the microarray underestimated the magnitude of differential transcription between the three life-stages. This was particularly evident for contig 933-1 in the parasitic/free-living contrast. The microarray analysis identified 933-1 as one of the most highly upregulated contigs in parasitic females – and this was also confirmed by qPCR – but whereas the mean difference for clones within this contig was ~10-fold ($2^{3.26}$) by microarray analysis, qPCR indicated a 5,000-fold difference ($2^{12.29}$). Similarly, but far less pronounced, the microarray underestimated the relative transcription of contigs 1092-1, 1501-1, 2048-1 and 351-1. Additionally, qPCR provided evidence for male-biased transcription for contigs 1407-1, 1847-1 and 2048-1, which was not detected by microarray analysis.

Table 1
Numbers of contigs that exhibited differential transcription in free-living versus parasitic females (columns) and in free-living males versus free-living females (rows)

| | Free-living-biased (>2-fold) | Free-living-biased (1.4- to 2-fold) | No differential transcription (<1.4-fold) | Parasitic-biased (1.4- to 2-fold) | Parasitic-biased (>2-fold) | Sub-total |
|---|---|---|---|---|---|---|
| Male-biased (>2-fold) | 45 | 32 | 152 | 15 | 8 | 252 |
| Male-biased (1.4- to 2-fold) | 15 | 32 | 409 | 29 | 6 | 491 |
| No differential transcription (<1.4-fold) | 46 | 110 | 1,869 | 193 | 38 | 2,256 |
| Female-biased (1.4- to 2-fold) | 4 | 15 | 371 | 65 | 4 | 459 |
| Female-biased (>2-fold) | 1 | 5 | 164 | 45 | 15 | 230 |
| Sub-total | 111 | 194 | 2,965 | 347 | 71 | 3,688 |

Table 2
Contigs that exhibited greatest differential expression in male/female contrasts (a and b) and in parasitic/free-living contrasts (c and d)

| *S. ratti* contig | Mean $\log_2$ fold change | Cluster | Accession number | Description of protein |
|---|---|---|---|---|
| *(a) Male-biased expression* | | | | |
| 2449-1 | 5.66 | *k*1 | WP:CE08181[a], Q9ERI6[b] | Mitochondrial short-chain dehydrogenase (DHS-22)[a], retinol dehydrogenase 14[b] |
| 2383-1 | 5.46 | *k*1 | EDP37601[c] | Metallopeptidase family (M24)[c] |
| 2485-1 | 5.24 | *k*1 | No hit | – |
| 2394-1 | 5.18 | *k*1 | No hit | – |
| 2546-1 | 5.08 | *k*1 | WP:CE24948[a], Q80TP3[b] | Predicted ubiquitin-protein ligase[a], ubiquitin-protein ligase (EDD1)[b] |
| 2425-1 | 4.97 | *k*1 | No hit | – |
| 2397-1 | 4.71 | *k*1 | No hit | – |
| 36-1 | 4.68 | *k*1 | No hit | – |
| 2581-1 | 4.61 | *k*1 | WP:CE37516[a], Q8CAK1[b] | Transcription factor in CCR4 complex[a], putative transferase homologue (C1orf69)[b] |
| 121-1 | 4.60 | *k*1 | WP:CE10896[a], AAN08882[c] | Major sperm protein (MSP) domain[a], MSP-domain protein 4[c] |
| *(b) Female-biased expression* | | | | |
| 302-1 | 4.02 | *k*3 | No hit | – |
| 2361-1 | 3.87 | *k*3 | No hit | – |
| 83-1 | 3.69 | *k*3 | No hit | – |
| 2520-1 | 3.63 | *k*3 | WP:CE41741[a], Q6RI86[b] | Molting defective (MLT-4)[a], transient receptor potential cation channel subfamily A member 1[b] |
| 2621-1 | 3.56 | *k*3 | No hit | – |
| 2626-1 | 3.40 | *k*3 | WP:CE30321[a], XP_001667627[c] | Unknown[a], Hypothetical protein in *C. briggsae*[c] |
| 178-1 | 3.36 | *k*3 | No hit | – |
| 2779-1 | 3.28 | *k*3 | No hit | – |
| 600-1 | 3.22 | *k*3 | WP:CE39506[a], XP_001678358[c] | Unknown[a], Hypothetical protein in *C. briggsae*[c] |
| 3046-1 | 3.18 | *k*3 | WP:CE06247[a], NP_492047[c] | Chondroitin proteoglycan family member (CPG-3)[a,c] |
| *(c) Parasitic-biased expression* | | | | |
| 2634-1 | 3.96 | *k*6 | No hit | – |
| 773-1 | 3.91 | *k*3 | No hit | – |
| 933-1 | 3.26 | *k*3 | WP:CE33320[a] | Aspartic protease (ASP-2)[a] |
| 352-1 | 3.18 | *k*6 | WP:CE04635[a] | Heat-shock protein (HSP-17)[a] |
| 429-2 | 2.97 | *k*3 | NP_001023958[c] | Heat-shock protein (HSP-17)[c] |
| 3564-1 | 2.84 | *k*6 | WP:CE04635[a] | Heat-shock protein (HSP-17)[a] |
| 219-1 | 2.59 | *k*3 | No hit | – |
| 3606-1 | 2.32 | *k*6 | No hit | – |
| 429-1 | 2.26 | *k*3 | WP:CE04635[a] | Heat-shock protein (HSP-17) |
| 889-1 | 2.17 | *k*6 | WP:CE17723[a], Q5FW05[b] | Unknown[a], ATP-binding domain protein 3[b] |
| *(d) Free-living-biased expression* | | | | |
| 207-2 | 4.46 | *k*2 | No hit | – |
| 2550-1 | 4.33 | *k*2 | No hit | – |
| 2394-1 | 4.02 | *k*1 | No hit | – |
| 1534-1 | 3.69 | *k*2 | WP:CE31783[a] | Collagen (COL-17)[a] |
| 1694-1 | 3.52 | *k*2 | WP:CE09542[a] | Aspartic protease (ASP-6)[a] |
| 4107-1 | 3.50 | *k*2 | No hit | – |
| 4178-1 | 3.47 | *k*2 | WP:CE05408[a], XP_001676898[c] | Unknown[a], hypothetical protein in *C. briggsae*[c] |
| 4060-1 | 3.28 | *k*2 | No hit | – |
| 2581-1 | 3.25 | *k*1 | WP:CE37516[a] | Transcription factor in CCR4 complex[a] |
| 2449-1 | 3.25 | *k*1 | WP:CE08181[a], Q9ERI6[b] | Mitocondrial short-chain dehydrogenase (DHS-22)[a], retinol dehydrogenase 14[b] |

Annotation derived from [a] Wormpep; [b] Swissprot; or [c] nr databases.

### 3.3. GO analysis

As this study identified a large number of differentially regulated transcripts, the general patterns inherent within these data were interpreted by testing whether genes that shared a similar pattern of transcription also shared similar biological characteristics. Accordingly, the 3,688 contigs were grouped into 12 *k*-means clusters (*k*1–*k*12) based on their pattern of co-transcription among the 38 hybridisations (Table 3). Thus, each cluster contained contigs that exhibited a similar pattern of transcription among life-stages and replicates. The size of these 12 clusters ranged from 15 to 793. The *k*-means clustering algorithm was repeated 1,000 times to give a measure of robustness for each cluster, defined as the average proportion of contigs from each cluster in the original run of the algorithm assigned to that same cluster in subsequent runs. On this basis, the robustness of the 12 clusters ranged from 81% to 99%. Fig. 3 shows the gender- and

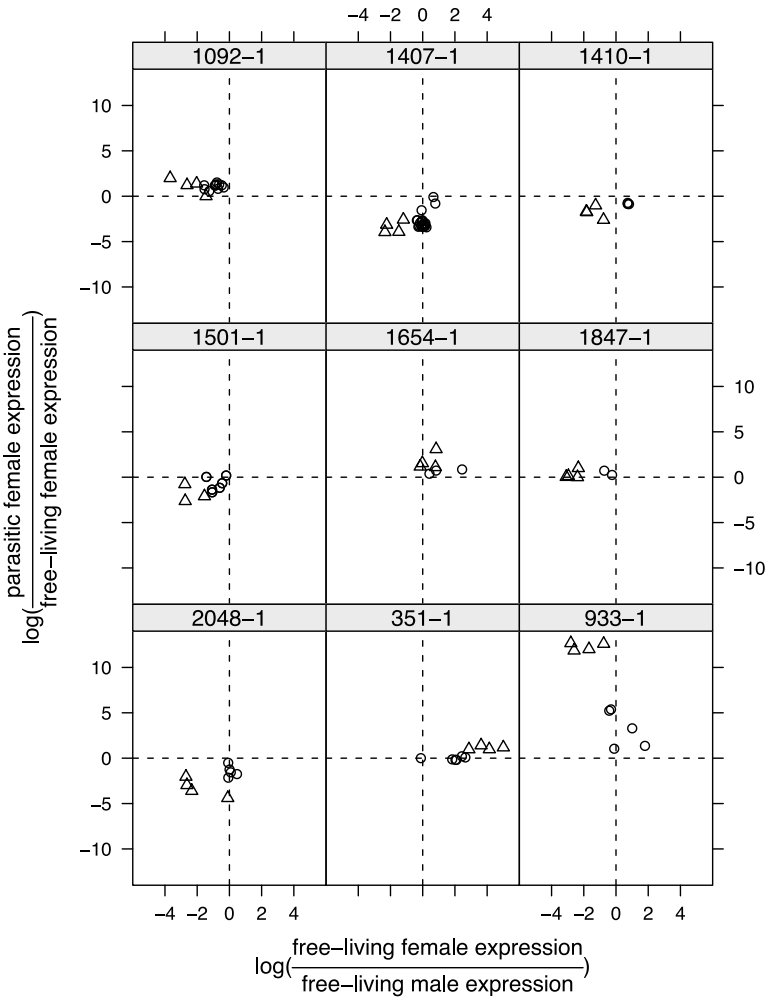


Fig. 2. Comparison between $\log_2$ fold changes against the male/female and parasitic/free-living contrasts measured by microarray analysis and by quantitative PCR. Data are shown for fold changes measured by quantitative PCR from individual samples (triangles) and for fold changes estimated by microarray for individual clones within a contig (circles).

parasitic-biased transcription of contigs within these 12 *k*-means clusters. Contigs within a cluster tended to co-locate and form distinct groups, e.g., *k*2 represented a group of contigs upregulated in free-living stages and *k*4 a group upregulated in males. Therefore, clustering on the basis of co-transcription was able to distinguish subsets of contigs on the array without using arbitrary fold changes. Cumulative hypergeometric tests were then used to identify GO terms within the biological process (BP) ontology that were over-represented within each cluster. A total of 76 GO terms were identified within the BP ontology (Fig. 4) as significantly associated ($P < 0.01$) with one or more *k*-means cluster (Fig. 5). Equivalent analyses for the molecular function and cellular compartment ontologies are presented in Supplementary Figs. S1–S4.

## 4. Discussion

Approximately half of the *S. ratti* transcripts represented on the microarray exhibited gender-biased transcription and/or parasite-biased transcription. Two complementary approaches were applied to understand the biological significance of these data: first, the putative function(s) of transcripts that showed the greatest level of differential transcription was investigated by searching for orthologues in *C. elegans* and other model organisms (Table 2); and second, characteristics shared between transcripts that exhibited a similar pattern of transcription were studied (Figs. 3–5) (Allison et al., 2006). From these findings, the present study provides some new insights into the differences between males and females, and between parasitic and free-living stages.

How do males and females differ in *S. ratti*? Morphologically, their most obvious difference is the presence of a testis in males and a uterus in females, which comprise a large fraction of the body of an adult nematode (Justine, 2002). Thus, a significant proportion of genes associated with reproductive processes exhibited gender-biased transcription. Clusters *k*4 and *k*6 displayed male- and female-biased transcription, respectively, and were associated with GO terms for reproductive processes or behaviour (GO:0019098, GO:0032504, GO:0033057 and GO:0048609), genitalia development (GO:0040035 and GO:0048806) and sex differentiation (GO:0007548)

Table 3
Cluster analysis of contigs indicating the robustness of clusters (%) and the number of differentially transcribed contigs within each cluster between the male/female contrast and the free-living/parasitic contrast

| Cluster | *k*1 | *k*2 | *k*3 | *k*4 | *k*5 | *k*6 | *k*7 | *k*8 | *k*9 | *k*10 | *k*11 | *k*12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No. of contigs | 15 | 43 | 61 | 101 | 180 | 188 | 207 | 305 | 502 | 596 | 697 | 793 |
| Robustness | 99% | 98% | 97% | 84% | 81% | 95% | 90% | 83% | 90% | 94% | 84% | 84% |
| Free-living (>2-fold) | 15 | 43 | 0 | 17 | 36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Free-living (1.4- to 2-fold) | 0 | 0 | 2 | 36 | 78 | 1 | 3 | 3 | 0 | 36 | 0 | 35 |
| No differential transcription (<1.4-fold) | 0 | 0 | 37 | 46 | 66 | 126 | 194 | 285 | 359 | 553 | 548 | 751 |
| Parasitic (1.4- to 2-fold) | 0 | 0 | 10 | 1 | 0 | 40 | 10 | 15 | 122 | 7 | 135 | 7 |
| Parasitic (>2-fold) | 0 | 0 | 12 | 1 | 0 | 21 | 0 | 2 | 21 | 0 | 14 | 0 |
| Male (>2-fold) | 15 | 16 | 0 | 82 | 4 | 2 | 31 | 81 | 2 | 0 | 15 | 4 |
| Male (1.4- to 2-fold) | 0 | 8 | 0 | 19 | 31 | 2 | 97 | 137 | 5 | 2 | 112 | 78 |
| No differential transcription (<1.4-fold) | 0 | 16 | 2 | 0 | 141 | 22 | 79 | 87 | 305 | 392 | 539 | 673 |
| Female (1.4- to 2-fold) | 0 | 3 | 1 | 0 | 3 | 57 | 0 | 0 | 151 | 175 | 31 | 38 |
| Female (>2-fold) | 0 | 0 | 58 | 0 | 1 | 105 | 0 | 0 | 39 | 27 | 0 | 0 |

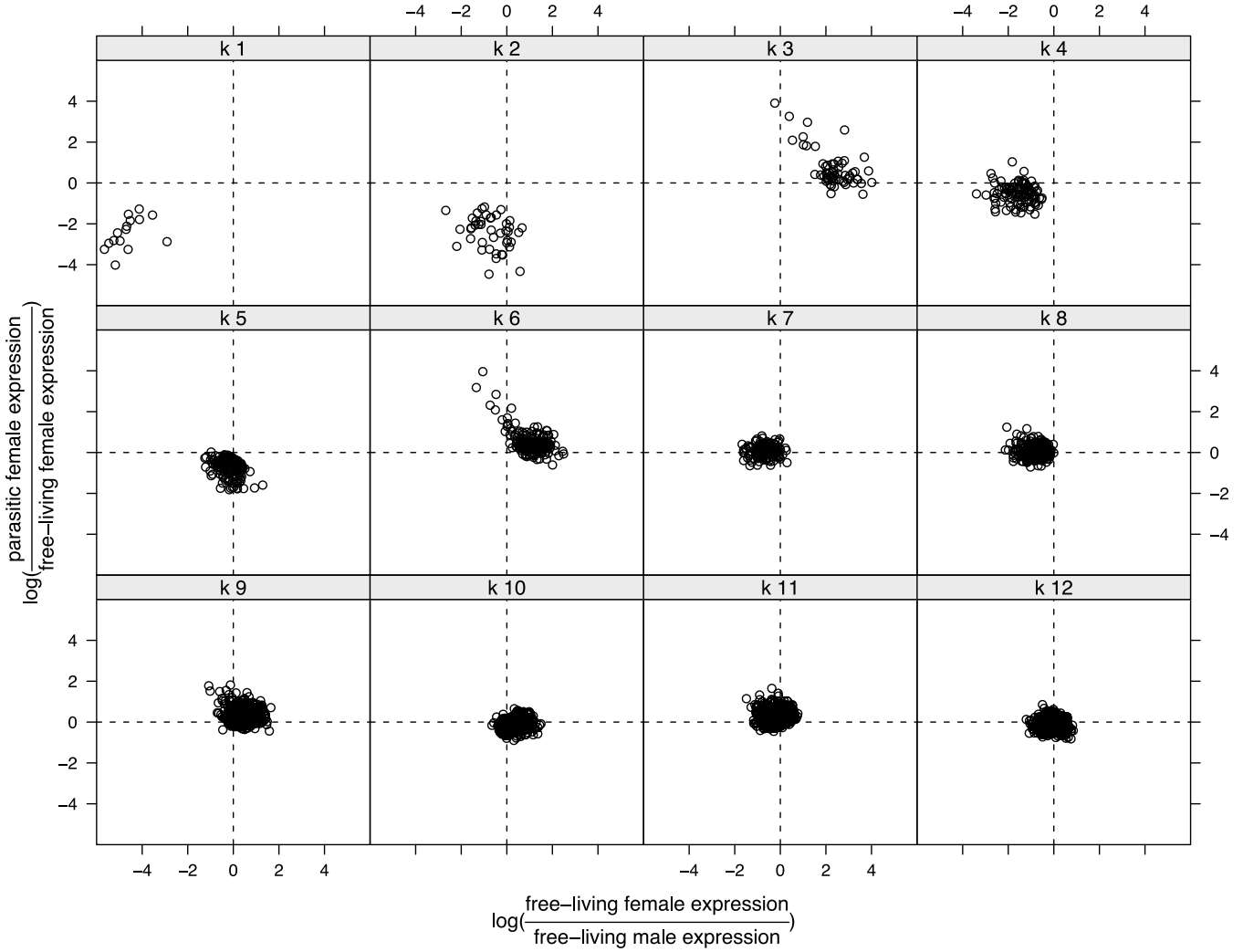


Fig. 3. Contigs ($n = 3{,}688$) assigned to 12*k*-means clusters plotted against the male/female and parasitic/free-living contrasts. Clusters are arranged by increasing numbers of contigs assigned to each cluster.

(Fig. 5). Cluster *k*10 exhibited female-biased transcription and was associated with GO:0000003 'Reproduction', and also GO terms associated with embryonic and larval development (GO:0009790, GO:0009791 and GO:0009792). At the individual transcript level, an orthologue of *cpg-3* (chrondroitin protoegylcan-3; R06C7.4) exhibited strong female-biased transcription (Table 2). In *C. elegans*, chrondroitin protoegylcans are critical to embryonic development and are implicated in interactions with chitin present in nematode eggshells (Olson et al.,

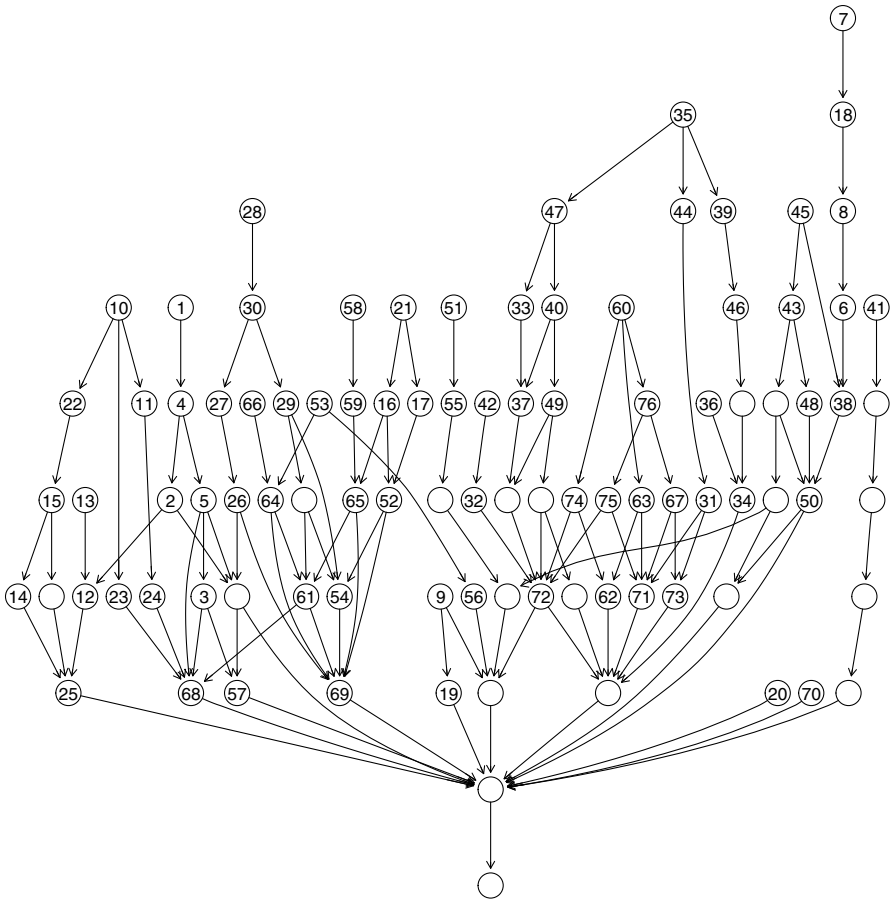


Fig. 4. The hierarchical structure for gene ontology (GO) terms in the biological process hierarchy for which significant over-representation ($P < 0.01$) of contigs within one or more of the 12$k$-means clusters was observed. Numbered nodes are GO terms with significant over-representation, blank nodes are their non-significant, parent nodes within the GO hierarchy. The GO terms and descriptions associated with each numbered node can be found in Fig. 5.

2006). In addition to morphological differences, there may also be different constraints on the life-histories of males and females (Poulin, 2007). Whereas free-living females need to survive long enough to produce a batch of fertilised eggs, males only need to survive long enough to fertilize these eggs. Consistent with this, free-living male *S. ratti* die more rapidly than free-living females. Moreover, populations of males within a faecal pellet are often synchronous, such that competition for access to females may occur, which may result in a higher metabolic activity in males. Thus, cluster $k8$ exhibited male-biased expression and contained several GO terms associated with metabolism and the tricarboxylic acid cycle, including GO:0006091 'generation of precursor metabolites and energy'.

How do parasitic and free-living stages differ from each other in *S. ratti*? Nematodes in the external environment are subjected to abiotic fluctuations in temperature, salinity and pH, which are not experienced by parasitic stages (Waller and Faedo, 1996). Thus, cluster $k5$ was upregulated in free-living stages and associated with GO terms GO:0050878 'regulation of body fluids' and GO:0006820 'anion transport', which may help free-living stages adapt to changes in pH and salinity, and GO:0009605 'response to external stimulus' and GO:0007626 'locomotory behaviour', which may help free-living stages move into suitable environments. How parasitic stages differ from free-living stages is less clear. Cluster $k9$ was upregulated in parasitic females versus free-living females, and also upregulated in free-living females versus free-living males. This cluster was associated with GO:0051276 'chromosome organisation and biogenesis' and GO:0051293 'establishment of spindle localisation', which may be related to mitotic parthenogenesis employed by parasitic females of *S. ratti*. There are three GO terms associated with both $k9$ and $k5$; GO:009886 'post-embryonic morphogenesis', GO:0010171 'body morphogenesis' and GO:0040032 'post-embryonic body morphogenesis'. These might represent genes involved in the different body morphologies of free-living males, free-living females and parasitic females. Various orthologues of a *C. elegans* heat-shock gene, *hsp-17*, were upregulated in parasitic females (Table 2), which was also seen in an analysis of *S. ratti* ESTs by Thompson et al. (2005). Whether the several, distinct contigs associated with *hsp-17* represent different duplicated genes within *S. ratti* (paralogues) or different splice variants of the same gene is unclear at present but should be resolved once a complete genome sequence of *S. ratti* becomes available.

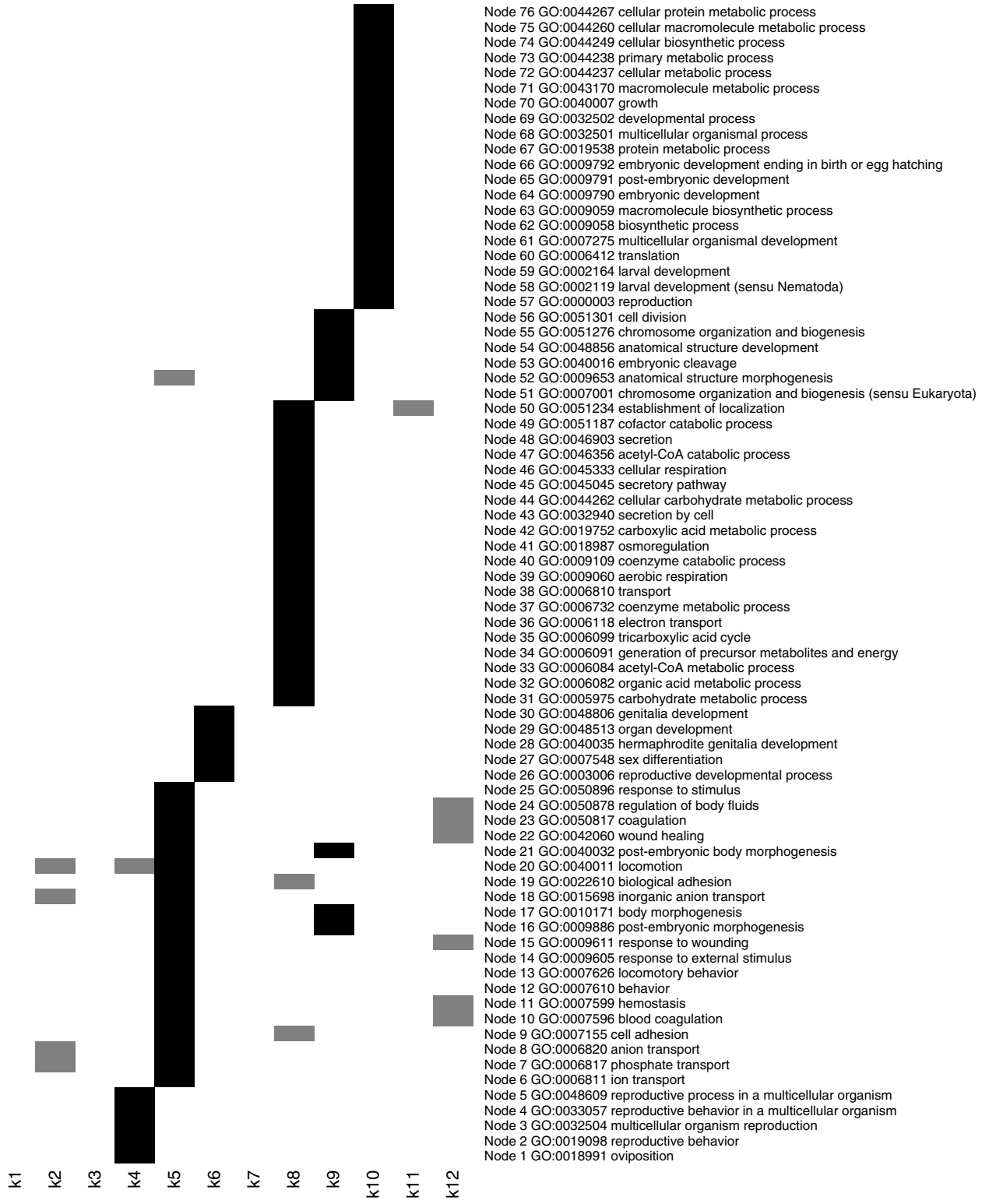


Fig. 5. A heatmap showing the gene ontology terms in the biological process hierarchy for which significant over-representation ($P < 0.01$) of contigs within one or more of the 12$k$-means clusters ($k1$–$k12$) was observed. Black areas indicate significance at $P < 0.01$. Grey areas indicate an additional association with another cluster at $P < 0.05$.

Small heat-shock proteins, despite the name, are not necessarily involved in stress responses in parasitic nematodes since various orthologues were shown to be highly upregulated in female *T. vitrinus* and *O. dentatum* (see Nisbet and Gasser, 2004; Cottee et al., 2006) compared with males from the same host and HSP-20 protein in *H. contortus* was not upregulated in response to heat-shock (Hartman et al., 2003). It is notable that an aspartic protease was

upregulated in parasitic *S. ratti* stages (933-1). Aspartic proteases have been reported to aid degradation of host haemoglobin in hookworms and are currently being developed as hookworm vaccines (Williamson et al., 2003; Loukas et al., 2006). However, *S. ratti* does not feed on host blood and thus it is possible that, in *S. ratti* parasitic females, aspartic proteases help digest mucus. The proposal that parasitism is believed to have arisen separately in the lineages leading to hookworms and *Strongyloides* (see Blaxter et al., 1998; Dorris et al., 2002) suggests a widespread role for aspartic proteases in nematode parasitism.

Two recent studies (Li et al., 2005; Campbell et al., 2008) reported microarray analyses of gender-biased transcription in *B. malayi* and *H. contortus*, respectively; some notable similarities exist between those and the present study. Campbell et al. (2008) identified an interaction network involving acetyl-coA synthetase and the upregulation of genes involved in the tricarboxylic acid (TCA) cycle in males. Similarly, in *S. ratti*, GO:0006084 'acetyl-coA metabolic process' and GO:0006084 'tricarboxylic acid cycle' were associated with cluster *k*8, which exhibited male-biased transcription. In addition, cluster *k*8 contained *sdhb-1* (succinate dehydrogenase; F42A8.2) a component of the TCA cycle which Campbell et al. (2008) also found upregulated in males. This suggests an increased requirement for carbohydrate-based energy production through the TCA cycle in males. Li et al. (2005) reported an upregulation of myosin genes in *B. malayi* males, several of which were shown to be upregulated in *S. ratti* males, specifically the myosin light chain genes *mlc-2* (C36E6.5) and *mlc-3* (F09F7.2), and the myosin heavy chain gene *unc-54* (F11C3.3). Additionally, both Campbell et al. (2008) and the present study showed that tropomyosin, *lev-11* (Y105E8B.1), was upregulated in males. Myosin and tropomyosin are critical to muscle function (Nikolaou et al., 2006) and locomotion, and are energetically costly. Thus, increased energy production and muscle function are consistent with the requirement for males to find females and to compete with other males.

Li et al. (2005) identified a number of female-specific transcripts that were also identified in the present study and associated with GO:004032 ('post-embryonic body morphogenesis'); *noah-1* (C34G6.6), *noah-2* (F52B11.3), *cct-4* (K01C8.10), *pab-1* (Y106G6H.2), *sma-1* (R31.1) and *unc-70* (K11C4.3). The genes *noah-1* and *noah-2* are not well characterised in *C. elegans* but are homologues of *Drosophila NOmpA*, which is a mechanoreceptor gene. Similarly, *cct-4* is a T complex chaperone, reported as being essential for mechanoreception (Zhang et al., 2002). The gene *pab-1* encodes a polyadenylate binding protein which acts through translational regulation and is crucial to female germline differentiation (Ciosk et al., 2004; Maine et al., 2004). The genes *sma-1* and *unc-70* both encode beta-spectrins, which are essential for the formation of the spectrin cytoskeleton and to epidermal morphogenesis (McKeown et al., 1998; Moorthy et al., 2000). Whether further molecular relationships exist between these genes and/or their products is presently unclear, but their female-biased transcription in both *S. ratti* and *B. malayi* suggests a significant role in the body morphology of parasitic nematodes.

In future, rapid transcriptomic analysis across many species of parasitic nematodes will be possible, enabled by new ultra-high throughput sequencing technologies (Margulies et al., 2005), in situ-synthesis of oligonucleotide arrays (Campbell et al., 2008) and digital transcriptomics (Nielsen et al., 2006). The bioinformatic approaches used herein can be applied to such technologies, and future extensions could include methods to help assign function to those genes that share little or no homology to those of *C. elegans*. Clustering on the basis of co-transcription may be particularly appropriate; thus, two parasitic nematode genes, one with an orthologue in *C. elegans* and the other without, might share a similar function if they share a similar pattern of transcription (Kim et al., 2001). Similarly, many parasitic nematode genes may exhibit homology to *C. elegans* genes but may have adopted new biological functions associated with the evolution of parasitism, and the identification of such genes should be a priority. One solution may be to incorporate molecular evolutionary analyses; genes that have undergone rapid, adaptive evolution between free-living and parasitic taxa are more likely to have acquired new biological functions. Ultimately, however, one would wish to determine the functions of, and interactions among parasite genes directly through molecular and biochemical approaches. Microarray studies thus provide a basis for further, more detailed and fundamental studies into the biology of parasitic nematodes. As such, there are good prospects for identifying novel drug or vaccine targets against parasite reproduction and/or survivorship within a host that are applicable across a range of parasitic nematode species.

## Acknowledgements

This work was funded by the Wellcome Trust and the Natural Environment Research Council, UK. The authors are grateful for the expertise and advice provided by staff of the NERC Molecular Genetics Facility in the University of Liverpool and the NERC Environmental Bioinformatics Centre, Oxford and for technical assistance by Rebecca Barber and Jo Sanders.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijpara.2008.02.004.

## References

Allison, D.B., Cui, X., Page, G.P., Sabripour, M., 2006. Microarray data analysis: from disarray to consolidation and consensus. Nat. Rev. Genet. 7, 55–65.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Anderson, R.C., 2000. Nematode Parasites of Vertebrates: Their Development and Transmission. CABI, Cambridge, MA.

Bethony, J., Brooker, S., Albonico, M., Geiger, S.M., Loukas, A., Diemert, D., Hotez, P.J., 2006. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. Lancet 367, 1521–1532.

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.

Blaxter, M.L., 2003. Nematoda: genes, genomes and the evolution of parasitism. Adv. Parasitol. 54, 101–195.

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

Campbell, B.E., Nagaraj, S.H., Hu, M., Zhong, W., Sternberg, P.W., Ong, E.K., Loukas, A., Ranganathan, S., Beveridge, I., McInnes, R., 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.

Ciosk, R., DePalma, M., Priess, J.R., 2004. ATX-2, the *C. elegans* ortholog of ataxin 2, functions in translational regulation in the germline. Development 131, 4831–4841.

Cottee, P.A., Nisbet, A.J., El-Osta, Y.G.A., Webster, T.L., Gasser, R.B., 2006. Construction of gender-enriched cDNA archives for adult *Oesophagostomum dentatum* by suppressive-subtractive hybridization and a microarray analysis of expressed sequence tags. Parasitology 132, 691–708.

Dawkins, H.J.S., 1989. *Strongyloides ratti* infections in rodents: value and limitations as a model of human strongyloidiasis. In: Grove, D.I. (Ed.), Strongyloidiasis: A Major Roundworm Infection of Man. Taylor & Francis, London, pp. 287–332.

Dorris, M., Viney, M.E., Blaxter, M.L., 2002. Molecular phylogenetic analysis of the genus *Strongyloides* and related nematodes. Int. J. Parasitol. 32, 1507–1517.

Hartman, D., Cottee, P.A., Savin, K.W., Bhave, M., Presidente, P.J.A., Fluton, L., Walkiewicz, M., Newton, S.E., 2003. *Haemonchus contortus*: molecular characterisation of a small heat shock protein. Exp. Parasitol. 104, 96–103.

Hope, I.A., 1999. Background on *Caenorhabditis elegans*. In: Hope, I.A. (Ed.), *C. elegans*; a Practical Handbook. Oxford University Press, Oxford.

Hsiao, L., Jensen, R., Yoshida, T., Clark, K., Blumenstock, J., Gullans, S., 2002. Correcting for signal saturation errors in the analysis of microarray data. Biotechniques 32, 330–336.

Huang, X., Madan, A., 1999. CAP3: a DNA assembly program. Genomic Res. 9, 868–877.

Justine, J., 2002. Male and female gametes and fertilisation. In: Lee, D.L. (Ed.), The Biology of Nematodes. Taylor and Francis, London, pp. 73–119.

Kerr, M.K., 2003. Design considerations for efficient and effective microarray studies. Biometrics 59, 822–828.

Kim, S.K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J.M., Eizinger, A., Wylie, B.N., Davidson, G.S., 2001. A gene expression map for *Caenorhabditis elegans*. Science 293, 2087–2092.

LaMunyon, C.W., Ward, S., 1998. Larger sperm outcompete smaller sperm in the nematode *Caenorhabditis elegans*. Proc. R. Soc. B 265, 1997–2002.

LaMunyon, C.W., Ward, S., 1999. Evolution of sperm size in nematodes: sperm competition favours larger sperm. Proc. R. Soc. B 266, 263–267.

Lewthwaite, P., Gill, G.V., Hart, C.A., Beeching, N.J., 2005. Gatrointestinal parasites in the immunocompromised. Curr. Opin. Infect. Dis. 18, 427–435.

Li, B.W., Rush, A.C., Crosby, S.D., Warren, W.C., Williams, S.A., Mitreva, M., Weil, G.J., 2005. Profiling of gender-regulated gene transcripts in the filarial nematode *Brugia malayi* by cDNA oligonucleotide array analysis. Mol. Biochem. Parasitol. 143, 49–57.

Loukas, A., Bethony, J., Hotez, P.J., 2006. Hookworm vaccines: past, present and future. Lancet Infect. Dis. 6, 733–741.

Maine, E.M., Hansen, D., Springer, D., Vought, V.E., 2004. *Caenorhabditis elegans atx-2* promotes germline proliferation and the oocyte fate. Genetics 168, 817–830.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380.

McKeown, C., Praitis, V., Austin, J., 1998. *sma-1* encodes a betaH-spectrin homolog required for *Caenorhabditis elegans* morphogenesis. Development 125, 2087–2098.

Moorthy, S., Chen, L., Bennett, V., 2000. *Caenorhabditis elegans* beta-G spectrin is dispensable for establishment of epithelial polarity, but essential for muscular and neuronal function. J. Cell Biol. 149, 915–930.

Nielsen, K.L., Laursen Hogh, A., Emmersen, J., 2006. DeepSAGE – digital transcriptomics with high sensitivity, simple experimental protocol and multiplexing of samples. Nucleic Acids Res. 34, e133.

Nikolaou, S., Hu, M., Chilton, N.B., Hartman, D., Nisbet, A.J., Presidente, P.J.A., Gasser, R.B., 2006. Class II myosins in nematodes – genetic relationships, fundamental and applied implications. Exp. Parasitol. 24, 338–350.

Nisbet, A.J., Cottee, P., Gasser, R.B., 2004. Molecular biology of reproduction and development in parasitic nematodes: progress and opportunities. Int. J. Parasitol. 34, 125–138.

Nisbet, A.J., Gasser, R.B., 2004. Profiling of gender-specific gene expression for *Trichostrongylus vitrinus* (Nematoda: Strongylida) by microarray analysis of expressed sequence tag libraries constructed by suppressive-subtractive hybridisation. Int. J. Parasitol. 34, 633–643.

Olson, S.K., Bishop, J.R., Yates, J.R., Oegema, K., Esko, J.D., 2006. Identification of novel chrondroitin proteoglycans in *Caenorhabditis elegans*: embryonic cell division depends on CPG-1 and CPG-2. J. Cell Biol. 6, 985–994.

Parkinson, J., Mitreva, M., Whitton, C., Thomson, M., Daub, J., Martin, J., Schmid, R., Hall, N., Barrell, B., Waterston, R.H., McCarter, J.P., Blaxter, M.L., 2004. A transcriptomic analysis of the phylum Nematoda. Nat. Genet. 36, 1259–1267.

Paterson, S., Viney, M.E., 2003. Functional consequences of genetic diversity in *Strongyloides ratti* infections. Proc. R. Soc. B 270, 1023–1032.

Paterson, S., Barber, R., 2007. Experimental evolution of parasite life-history traits in *Strongyloides ratti* (Nematoda). Proc. R. Soc. B 274, 1467–1474.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, 2002–2007.

Poulin, R., 2007. Evolutionary Ecology of Parasites. Princeton University Press, Princeton, NJ.

Reinke, V., Gil, I.S., Ward, S., Kazmer, K., 2004. Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. Development 131, 311–323.

Roberts, M.G., 1995. A pocket guide to host-parasite models. Parasitol. Today 11, 172–177.

Thompson, F.J., Mitreva, M., Barker, G.L., Martin, J., Waterson, R.H., McCarter, J.P., Viney, M.E., 2005. An expressed sequence tag analysis of the life-cycle of the parasitic nematode *Strongyloides ratti*. Mol. Biochem. Parasitol. 142, 32–46.

Viney, M.E., 1994. A genetic analysis of reproduction in *Strongyloides ratti*. Parasitology 109, 511–515.

Viney, M.E., 1999. Exploiting the life cycle of *Strongyloides ratti*. Parasitol. Today 15, 231–235.

Waller, P.J., Faedo, M., 1996. The prospects for biological control of the free-living stages of nematode parasites of livestock. Int. J. Parasitol. 26, 915–925.

Williamson, A.L., Brindley, P.J., Loukas, A., 2003. Hookworm cathepsin D aspartic proteases: contributing roles in the host-specific degradation of serum proteins and skin macromolecules. Parasitology 126, 179–185.

Wit, E., McClure, J., 2004. Statistics for Microarrays; Design, Analysis and Inference. John Wiley, Chichester, UK.

Zhang, Y., Ma, C., Delohery, T., Nasipak, B., Foat, B.C., Bounoutas, A., Bussemaker, H.J., Kim, S.K., Chalfie, M., 2002. Identification of genes expressed in *C. elegans* touch receptor neurons. Nature 418, 331–335.