DATASET BRIEF

Animal Proteomics

Proteomics
www.proteomics-journal.com

Proteomic Characterization of *Caenorhabditis elegans* Larval Development

Tian Xia, Edward R. Horton, Anna Elisabetta Salcini, Roger Pocock, Thomas R. Cox, and Janine T. Erler**

The nematode *Caenorhabditis elegans* is widely used as a model organism to study cell and developmental biology. Quantitative proteomics of *C. elegans* is still in its infancy and, so far, most studies have been performed on adult worm samples. Here, we used quantitative mass spectrometry to characterize protein level changes across the four larval developmental stages (L1–L4) of *C. elegans*. In total, we identified 4130 proteins, and quantified 1541 proteins that were present across all four stages in three biological replicates from independent experiments. Using hierarchical clustering and functional ontological analyses, we identified 21 clusters containing proteins with similar protein profiles across the four stages, and highlighted the most overrepresented biological functions in each of these protein clusters. In addition, we used the dataset to identify putative larval stage-specific proteins in each individual developmental stage, as well as in the early and late developmental stages. In summary, this dataset provides system-wide analysis of protein level changes across the four *C. elegans* larval developmental stages, which serves as a useful resource for the *C. elegans* research community. MS data were deposited in ProteomeXchange (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the primary accession identifier PXD006676.

The nematode *Caenorhabditis elegans* is used as a eukaryotic model organism to study cell and developmental biology. The short life span, invariant cell lineages, self-fertilizing hermaphroditic lifestyle, and genetic amenability make this organism a powerful system for genetic and functional analyses.[1,2] Under laboratory culturing conditions, postembryonic development follows four larval stages (L1–L4) before animals become sexually mature adults. Under adverse conditions, such as starvation, *C. elegans* can enter a dauer larval stage, which is a developmentally arrested stage. Biochemical studies using quantitative mass spectrometry (MS) approaches have only recently begun to be used in *C. elegans* research, and have so far been applied mainly to study the molecular profile of *C. elegans* in adult worms.[3–11] Quantitative MS has been used to study the proteomes of specific worm strains,[3,4] as well as protein changes in response to heat shock,[5,6] in age-related processes,[7,8] in dauer larva,[9] and during starvation.[10] To our knowledge, only two studies[11,12] have used a proteomics-based methodology to decipher protein changes during larval development before adulthood. In particular, Grün et al. used a combination of quantitative MS and mRNA sequencing to compare protein and transcript changes during the life cycle of the two closely related nematode species *C. elegans* and *Caenorhabditis briggsae*, which revealed conserved mRNA and protein changes during evolution.[11]

Here, we used quantitative MS to characterize the protein level changes across the four larval developmental stages of *C. elegans* in biological triplicate, which substantially advances knowledge of proteome changes during *C. elegans* postembryonic development. We used the stable-isotope labeling with amino acids in cell culture (SILAC)-based methodology[4,5] to label wild-type N2 Bristol animals. *C. elegans* were grown at 20 °C on nematode growth medium plates seeded with the lysine auxotroph *Escherichia coli* strain ET505 (F^-, λ^- lysA0::Tn10. IN(rrnD-rrnE)1) as food. In two of the three biological repeats, *E. coli* was grown in M63 minimum medium supplemented with heavy lysine (lysine-$^{13}C_6{}^{15}N_2$), and in one biological repeat light lysine (lysine-$^{12}C_6{}^{14}N_2$) was used. Nematodes were grown on plates with heavy or light lysine labeled bacteria for two generations[4] before protein samples were taken (**Figure 1**A). We performed an incorporation test to determine the efficiency of SILAC labeling, and found that proteins collected from all four stages had incorporation rates of at least 96% (L1, 98.3%; L2, 97.5%; L3, 97.3%; L4, 96.8%).

T. Xia, Dr. E. R. Horton, Dr. A. E. Salcini, Dr. T. R. Cox, Prof. J. T. Erler
Biotech Research and Innovation Centre
University of Copenhagen
Copenhagen, Denmark
E-mail: t.cox@garvan.org.au; janine.erler@bric.ku.dk
Dr. R. Pocock
Department of Anatomy and Developmental Biology
Development and Stem Cells Program
Monash Biomedicine Discovery Institute
Monash University
Melbourne, Australia
Dr. T. R. Cox
Cancer Division
Garvan Institute of Medical Research and The Kinghorn Cancer Centre
St Vincent's Clinical School
Faculty of Medicine
UNSW Sydney
Sydney, Australia

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/pmic.201700238

DOI: 10.1002/pmic.201700238

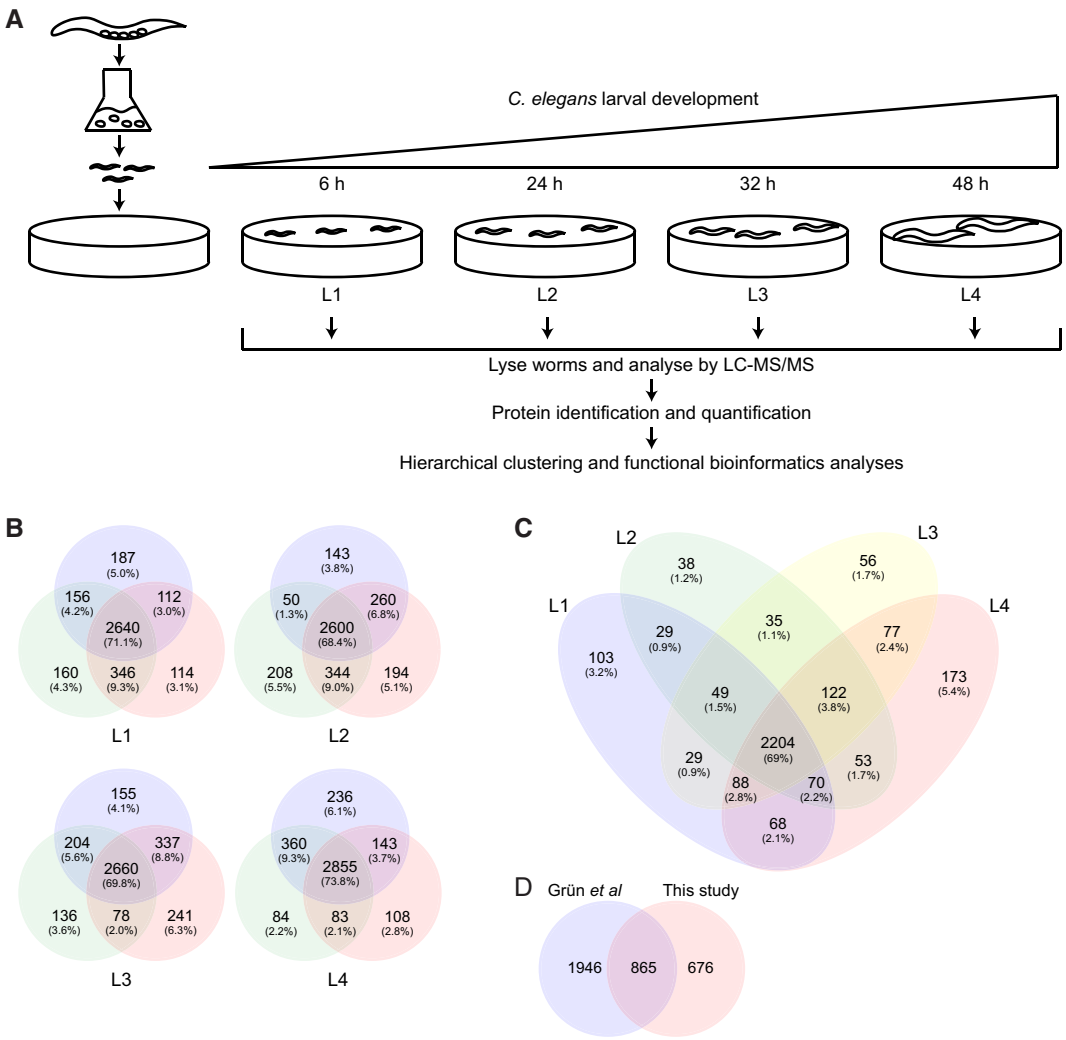

Figure 1. Overview of the experimental approach. A) To synchronize *C. elegans* animals, embryos were isolated by bleaching of mixed stage worms. Isolated embryos were grown in M9 buffer overnight without food. The hatched larvae were arrested at the L1 stage until provided with food. Arrested L1 larvae were transferred to petri dishes containing *E. coli* as food. L1, L2, L3, and L4 larva stages were sampled at 6 h, 24 h, 32 h, and 48 h post-feeding, respectively. Worm samples were then lysed and analyzed by LC-MS/MS, followed by protein identification and quantification. Data represented are from three independent biological replicates. B) Overlap of proteins identified between the three biological repeats in each larval stage is displayed as Venn diagrams. C) Overlap of proteins identified in all three biological repeats between the four larval stages is displayed as a Venn diagram. D) Overlap of proteins quantified across all four larval stages by Grün et al.[11] and the dataset in this study is displayed as a Venn diagram.

To synchronize the population of *C. elegans* used, embryos were isolated by bleaching and incubated in M9 buffer overnight without food, and hatched larva were arrested at L1 stage until provided with food. L1, L2, L3, and L4 stage nematodes were sampled 6, 24, 32, and 48 h post-feeding, respectively (Figure 1A). The stages of *C. elegans* were visually confirmed by checking morphology under a dissection microscope. At each time point, *C. elegans* were washed off the plates with M9 buffer, and washed three times in M9 buffer to minimize bacterial contamination. *C. elegans* were lysed in 1 mL lysis buffer (6 M urea/2 M thiourea + 10 mM HEPES pH 8.0) and sonicated at 4 °C using a Bioruptor at high energy setting (30 cycles each of 30 s on, 10 s off). Proteins were dissolved at room temperature for 1 h followed by centrifugation (13 000 × *g*, 4 °C, 5 min) to pellet undissolved proteins. Protein concentration was measured using the Bradford method and 100 μg of protein was collected. Proteins were reduced in final concentration 1 mM dithiothreitol (DTT) for 30 min at room temperature in the dark, followed by addition of final concentration 5 mM 2-chloroacetamide (CAA) for 30 min at room temperature in the dark. Samples were diluted four times with 50 mM ammonium bicarbonate, digested with Lys-C (1:100, 1 μg Lys-C per 100 μg protein) overnight at 37 °C, acidified with final concentration 2% trifluoroacetic acid (TFA), and centrifuged (2000 × *g*, 5 min) to clarify. Digested peptides were desalted using C18 columns as previously published.[13,14] Columns were first activated with methanol, washed with 80% CH_3CN and 0.1% formic acid (FA), followed by two washes with 3% CH_3CN and 1% TFA. Ten micrograms of peptides were loaded on the column, followed by two washes with 0.1% FA. The columns were then stored at 4 °C in a humidified environment (0.1% TFA).

Each sample was run unfractionated as a single-shot fraction on a Thermo Fisher Q-Exactive Mass Spectrometer. For each run, peptide was loaded onto a 50-cm C18 reverse-phase analytical column (Thermo EasySpray ES803) using 100% Buffer A (0.1% FA in water) at a pressure of 720 bar using the Thermo EasyLC 1000 uHPLC system in a single-column setup, and the column oven operating at 45 °C. Peptides were eluted over a 4-h gradient ranging from 6 to 60% of a solution containing 80% acetonitrile, 0.1% FA. The Q-Exactive (Thermo Fisher Scientific) was run using the DD-MS2 top10 method. Full MS spectra were collected at a resolution of 70 000, with an automatic gain control (AGC) target of 3×10^6 or a maximum injection time of 20 ms and a scan range of 300–1750 $m\,z^{-1}$. The MS2 spectra were obtained at a resolution of 17 500, with an AGC target value of 1×10^6 or a maximum injection time of 60 ms, a normalized collision energy of 25%, and an underfill ratio of 0.1%. Dynamic exclusion was set to 45 s, and ions with a charge state of <2 or unknown were excluded. MS performance was verified for consistency by running complex cell lysate quality-control standards, and chromatography was monitored to check for reproducibility.

Raw MS/MS spectra were processed using MaxQuant version 1.5.1.2 software and searched against the CAEEL reference proteome UP000001940 6239 available at http://www.ebi.ac.uk/reference_proteomes. Peptides identified with a false-discovery rate (FDR) of <1.0% were assembled into identified proteins. MS/MS spectra searches allowed for K8 heavy labeled Lysine, carbamidomethylation of cysteines and possible acetylation of N-termini as fixed/mix modifications. Allowed variable modifications were oxidized methionine, deamidation of glutamine and asparagine, pyroglutamic acid modification at N-terminal glutamine and hydroxyproline. To ensure high-confidence identifications and quantification, a MaxQuant score of >50 was required.

In total, we identified 4130 proteins across all developmental stages in wild-type worms (Supporting Information, Table 1). We identified 3715 proteins from the L1 stage, of which 2640 were identified in all three repeats; 3799 proteins from the L2 stage, of which 2600 were identified in all three repeats; 3811 from the L3 stage, of which 2660 were identified in all three repeats; and 3869 proteins from the L4 stage, of which 2855 were identified in all three repeats (Figure 1B). These data show high reproducibility in the proteins identified across biological repeats in each larval stage. In total, 2204 proteins were identified in all three biological repeats across all four larval stages (Figure 1C).

To quantitatively describe reliable and biologically relevant protein expression changes between the larval development stages, we filtered the dataset to select proteins identified in all three biological repeats with at least two unique peptides across all four stages, which resulted in a high-confidence quantifiable list of 1541 proteins (Supporting Information, Table 2). Data were quantified using the protein intensity value provided by MaxQuant for each protein at each stage and in each biological replicate. For each condition, the individual intensity value for each protein was normalized to the summed intensity in that condition, as described previously.[15–17] Normalized values for each protein were averaged across the three biological repeats. The normalized mean intensity value for each protein at each larval stage was then normalized to the maximum value for each individual protein across the four time points, as we have achieved previously.[18] This approach results in a protein's expression as a proportion of its maximum across the four larval stages (see Supporting Information, Table 2).

We calculated the coefficient of variation (CV) for each protein at each larval stage, and found average CVs of 25, 26, 26, and 24% for L1, L2, L3, and L4 stages, respectively, which demonstrates high reproducibility between biological replicates for most proteins. For proteins with high CV values and therefore high variability between repeats, this may have been due to true biological variation, discrepancies in sampling times between different runs, or difficulties in accurately quantifying low abundance proteins by MS. We also calculated which proteins showed statistically significant changes in normalized intensity values between pairwise larval stages by using the Benjamini, Krieger, and Yekutieli test, with a significance threshold of FDR < 5%. The most number of significant changes occurred between L1 and L4 stages (Supporting Information, Table 2), whereas no significant differences were observed between L2 and L3 stages. These data suggest little change in the proteome between L2 and L3 stages, while larger changes occur during early and late events of development.

Overall, we quantified fewer proteins than the 2811 proteins quantified across all four stages previously by Grün et al.[11] (Figure 1D), which is likely because we applied a more stringent inclusion criteria that proteins should be quantified in all three independent biological repeats with at least two unique peptides to ensure that we identified biologically relevant, high-confidence and reproducible changes between larval stages. We also did not fractionate samples to ensure that each protein within each condition was processed simultaneously to reduce sources of variation. In addition, 865 proteins were quantified across all four larval stages by both studies (Figure 1D). We compared quantitative values by calculating the correlation (Pearson's R^2 and Spearman's rank ρ correlation coefficients) between both datasets for each protein. We identified 332 ($R^2 > 0.4$) and 415 ($\rho > 0.4$) proteins that showed a positive correlation between both datasets, which validates the identified protein expression profiles of these proteins. Proteins that did not show similar trends in protein expression changes between the two studies may be due to differences in sampling times at each larval stage, differences in sample preparation and MS data acquisition approaches, difficulty in quantifying low abundance proteins accurately, or due to true biological variability in their expression during development. We quantified 676 proteins that were not quantified across all four larval stages previously[11] (Figure 1D), of which 580 proteins were not identified at all, and therefore describes for the first time how these proteins change during larval development (see Supporting Information, Table 2).

To assess changes over time, data were hierarchically clustered using an uncentered Pearson correlation and a complete-linkage matrix, which was performed using Cluster 3.0[19] and Java TreeView.[20] In total, 21 protein groups were identified that contain proteins with similar protein profiles across the four stages (**Figure 2**; Supporting Information, Figure 1). To examine the functional role of co-clustered proteins during development, we performed functional enrichment analysis on individual clusters using the Database for Annotation, Visualization and Integrated Discovery (DAVID),[21] which identified significantly overrepresented functional categories from the Gene Ontology (biological

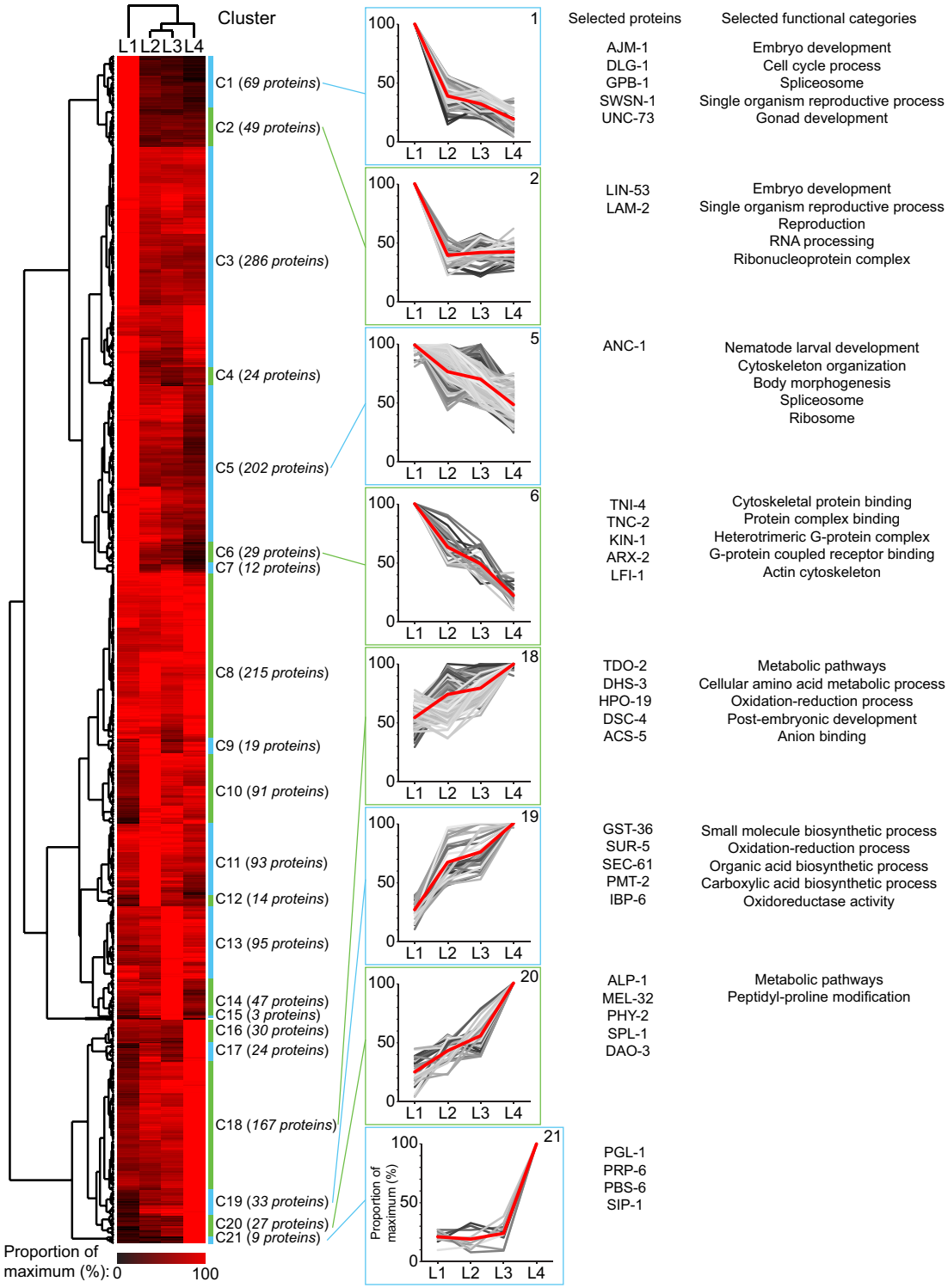

Figure 2. Temporal profiling of *C. elegans* proteins during larval development. Isolated *C. elegans* larvae at each developmental stage (L1–L4) were lysed and analyzed by mass spectrometry (data are from three independent experiments; see Supporting Information, Table 2). In total, 1541 proteins were quantified and analyzed by unsupervised hierarchical clustering, revealing distinct temporal profiles of protein expression during development. Twenty-one clusters, labelled C1–C21, were chosen on the basis of a Pearson correlation threshold greater than 0.9 and are indicated by blue and green bars. Clusters are shown alongside corresponding profile plots for selected groups, with the mean temporal profile for each cluster indicated by a red line. Profile plots for all 21 clusters are shown in Supporting Information, Figure 1. Quantitative heat map shows normalized mean intensity values as a proportion of the maximum value for each given protein. Selected proteins and significantly overrepresented functional annotations for particular clusters are listed. Full details of enriched functional terms are provided in Supporting Information, Table 3.

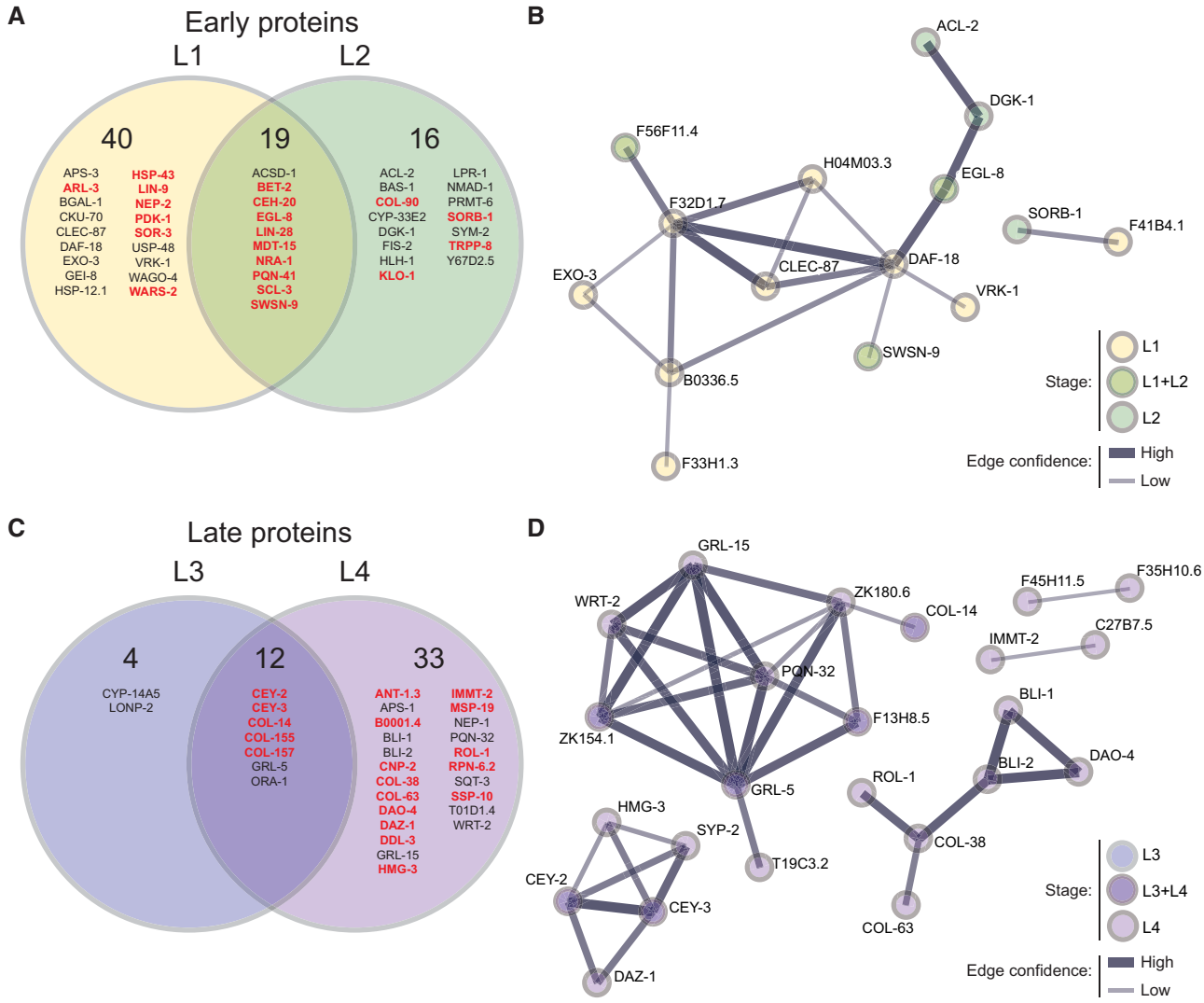

Figure 3. Identification of putative larval stage–specific proteins expressed during worm development. A) The number of proteins identified only in L1 stage, in both L1 and L2 stage, or only in L2 stage is shown. B) STRING analysis[27] of proteins identified only in L1 stage, in both L1 and L2 stage, or only in L2 stage. C) The number of proteins identified only in L3 stage, in both L3 and L4 stage, or only in L4 stage is shown. D) STRING analysis[27] of proteins identified only in L3 stage, in both L3 and L4 stage, or only in L4 stage. A protein was classed as being expressed in a larval stage provided the intensity value was greater than zero in at least two biological replicates for that stage, and the intensity value was equal to zero across all three replicates of other stages. Proteins from each category are labeled by protein name (see Supporting Information, Table 4). Nodes represent proteins and are colored according to stage-specific protein expression. Edges between nodes represent protein associations based on experiments, databases, and co-expression. Thicker lines indicate higher confidence associations. Unconnected proteins are not shown. Proteins that display maximum transcript expression levels (using data from a stage-specific RNA-seq gene expression dataset,[28] see Supporting Information, Table 4) in the stage we determined the protein to be specifically expressed are highlighted in bold, red text.

process, cellular component, and molecular function)[22] and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Categories were considered significantly overrepresented provided they satisfied the following criteria: fold enrichment ≥ 1.5, Bonferroni-corrected p-value <0.05, EASE score (modified Fisher's exact test) <0.05 and at least two proteins per keyword. A full list of significantly overrepresented terms is provided in Supporting Information, Table 3. These analyses demonstrate that many proteins previously known to be involved in developmental processes were identified here, as terms such as nematode larval development were overrepresented in multiple clusters. We also identified proteins associated with postembryonic development and metabolic pathways in cluster C18, which contains proteins whose abundance increased during development. For example, the tryptophan-converting enzyme tryptophan 2,3-dioxygenase (TDO-2) was identified in cluster C18, showing an increasing abundance during development. In support of our data showing an increased abundance of TDO-2 during larval development, TDO-2 has been shown previously to regulate adult life span and motility during ageing, and it has an age-accumulative effect in regulating motility.[23,24] In addition, we found that protein groups that tended to decrease in abundance during

development were mainly associated with the actin cytoskeleton and cell cycle regulation. Such clusters included terms such as cytoskeleton organization in cluster C5, cytoskeletal protein binding in cluster C6, and cell cycle process in cluster C1 (Figure 2). This suggests that cells are proliferating and migrating as nematodes are rapidly growing during this initial stage of development. However, clusters that contained proteins that increased during development were mainly enriched for functional terms associated with metabolism, such as metabolic pathways in clusters C19 and C20, and biosynthetic process in cluster C19. These data suggest a switch from rapid growth and cytoskeletal reorganization to an increased dependency on metabolism during the later stages of development. These analyses characterize the changing functional landscape of the *C. elegans* proteome during larval development.

Next, to demonstrate how our dataset can be used as a valuable resource by the scientific community, we sought to identify proteins that may be expressed at specific larval stages. We chose to classify proteins as expressed at a larval stage provided intensity values greater than zero were detected in at least two biological replicates, and we classified proteins as not expressed at a larval stage if the intensity value was equal to zero across all three replicates of a specific stage. Using this approach, we identified putative larval stage–specific proteins from our data (**Figure 3**; Supporting Information, Table 4). We characterized these as either "early proteins" when detected in only L1, in both L1 and L2, or only L2, and thus may be important for the early stages of worm development (Figure 3A,B); or "late proteins" when detected in only L3, in both L3 and L4, or only L4, and thus may be important for the later stages of worm development (Figure 3C,D). Several germline-related proteins are included in the L4-specific group. Spermatogenesis in *C. elegans* initiates during the L4 larval stage[25] and, indeed, sperm-specific proteins such as major sperm protein 19 (MSP-19) and sperm-specific class P protein 10 (SSP-10) appear in the L4-specific group (Figure 3C). Some of the proteins here have already been shown to exhibit stage-specific expression patterns. For example, the *bli-1* gene encoding for cuticle collagen (BLI-1) has been shown to be expressed only in the mid-L4 stage,[26] which agrees with our data highlighting BLI-1 as an L4-specific protein. STRING analysis[27] revealed protein associations (direct physical interactions or co-expression) between several larval stage–specific proteins, which further supports that these proteins are present at these stages. For example, a group of L4-specific proteins associated with BLI-1, all of which are collagen/cuticle-related proteins (Figure 3D).

To provide further evidence that these proteins are expressed at certain stages, and to provide a benchmark for our data in the context of other previously published studies, we compared our data to stage-specific RNA sequencing (RNA-seq) gene expression data from the modENCODE project.[28] We found 79% of putative L1–L2-specific proteins and 67% of putative L4-specific proteins have maximum transcript expression levels in the stage we determined the protein to be specifically expressed (Figure 3A,C; Supporting Information, Table 4). Moreover, 68 and 63% of the early and late proteins, respectively, were also classed as early/late in the RNA-seq data[28] (see Figure 3). These data show that many of the proteins identified as uniquely expressed at a specific larval stage by proteomics have maximal transcript abundance levels at these stages too, which provides further evidence that these proteins are likely to be functionally relevant at these larval stages. In several cases, however, mRNA transcripts were detected at larval stages while we did not identify the associated protein at that stage. This may be due to regulation at the translation level that results in inhibited mRNA translation, which could explain why mRNA levels were detected but protein levels were not, or because the protein was expressed at a level too low to be identified by MS at these stages.

In summary, the data presented in this study provide systems-wide proteome analysis of *C. elegans* larval development using quantitative MS-based proteomics. Analysis of this dataset reveals multiple groups of proteins that show different temporal expression profiles during development, and allowed us to identify putative larval stage–specific protein expression. These data provide a useful resource to the developmental and worm scientific communities, and add to the expanding number of studies using proteomics in nematodes.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

T.X. and E.R.H. have contributed equally to this work. This work was supported by Lundbeckfonden (R191-2015-1609) [T.X.], an EMBO long-term fellowship (ALTF 922–2016) [E.R.H.], the National Health and Medical Research Council (NHMRC) Australia (GNT1105374) and veski innovation fellowship (VIF 23) [R.P] the Innovation Foundation Denmark (1311-00010B) [T.R.C.], the National Health and Medical Research Council (NHMRC) Australia (GNT1129766) [T.R.C.], and a Hallas Møller Stipend from the Novo Nordisk Foundation [J.T.E.]. We thank Rune Linding (BRIC, University of Copenhagen) for providing access to mass spectrometry facilities. MS data were deposited in ProteomeXchange (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[29] with the primary accession identifier PXD006676.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

animal proteomics, bioinformatics, *Caenorhabditis elegans*, development, mass spectrometry

Received: June 20, 2017
Revised: November 3, 2017

[1] E. M. Jorgensen, S. E. Mango, *Nat. Rev. Genet.* **2002**, *3*, 356.
[2] B. Goldstein, N. King, *Trends Cell Biol.* **2016**, *26*, 818.
[3] M. Q. Dong, J. D. Venable, N. Au, T. Xu, S. K. Park, D. Cociorva, J. R. Johnson, A. Dillin, J. R. Yates III, *Science* **2007**, *317*, 660.
[4] J. Fredens, K. Engholm-Keller, A. Giessing, D. Pultz, M. R. Larsen, P. Højrup, J. Møller-Jensen, N. J. Færgeman, *Nat. Methods* **2011**, *8*, 845.

[5] M. Larance, A. P. Bailly, E. Pourkarimi, R. T. Hay, G. Buchanan, S. Coulthurst, D. P. Xirodimas, A. Gartner, A. I. Lamond, *Nat. Methods* **2011**, *8*, 849.
[6] V. Liang, M. Ullrich, H. Lam, Y. L. Chew, S. Banister, X. Song, T. Zaw, M. Kassiou, J. Götz, H. R. Nicholas, *Cell Mol. Life Sci.* **2014**, *71*, 3339.
[7] D. M. Walther, P. Kasturi, M. Zheng, S. Pinkert, G. Vecchi, P. Ciryam, R. I. Morimoto, C. M. Dobson, M. Vendruscolo, M. Mann, F. U. Hartl, *Cell* **2015**, *161*, 919.
[8] V. Narayan, T. Ly, E. Pourkarimi, A. B. Murillo, A. Gartner, A. I. Lamond, C. Kenyon, *Cell Syst.* **2016**, *3*, 144.
[9] C. Erkut, A. Vasilj, S. Boland, B. Habermann, A. Shevchenko, T. V. Kurzchalia, *PLoS One* **2013**, *8*, e82473.
[10] M. Larance, E. Pourkarimi, B. Wang, A. Brenes Murillo, R. Kent, A. I. Lamond, A. Gartner, *Mol. Cell. Proteomics* **2015**, *14*, 1989.
[11] D. Grün, M. Kirchner, N. Thierfelder, M. Stoeckius, M. Selbach, N. Rajewsky, *Cell Rep.* **2014**, *6*, 565.
[12] Y. Tabuse, T. Nabetani, A. Tsugita, *Proteomics* **2005**, *5*, 2876.
[13] T. R. Cox, R. M. H. Rumney, E. M. Schoof, L. Perryman, A. M. Høye, A. Agrawal, D. Bird, N. A. Latif, H. Forrest, H. R. Evans, I. D. Huggins, G. Lang, R. Linding, A. Gartland, J. T. Erler, *Nature* **2015**, *522*, 106.
[14] A. E. Majorca-Guliani, C. D. Madsen, T. R. Cox, E. R. Horton, F. A. Venning, J. T. Erler, *Nat. Med.* **2017**, *23*, 890.
[15] E. R. Horton, J. D. Humphries, B. Stutchbury, G. Jacquemet, C. Ballestrem, S. T. Barry, M. J. Humphries, *J. Cell Biol.* **2016**, *212*, 349.
[16] D. N. Itzhak, S. Tyanova, J. Cox, G. H. Borner, *eLIFE* **2016**, *5*, e16950.
[17] J. F. Krey, P. A. Wilmarth, J. B. Shin, J. Klimek, N. E. Sherman, E. D. Jeffery, D. Choi, L. L. David, P. G. Barr-Gillespie, *J. Proteome Res.* **2014**, *13*, 1034.
[18] E. R. Horton, A. Byron, J. A. Askari, D. H. J. Ng, A. Millon-Frémillon, J. Robertson, E. J. Koper, N. R. Paul, S. Warwood, D. Knight, J. D. Humphries, M. J. Humphries, *Nat. Cell Biol.* **2015**, *17*, 1577.
[19] M. J. L. de Hoon, S. Imoto, J. Nolan, S. Miyano, *Bioinformatics* **2004**, *20*, 1453.
[20] A. J. Saldanha, *Bioinformatics* **2004**, *20*, 3246.
[21] D. W. Huang, B. T. Sherman, R. A. Lempicki, *Nat. Protoc.* **2009**, *4*, 44.
[22] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, *Nat. Genet.* **2000**, *25*, 25.
[23] H. Michels, R. I. Seinstra, J. C. Uitdehaag, M. Koopman, M. van Faassen, C. N. Martineau, I. P. Kema, R. Buijsman, E. A. Nollen, *Sci. Rep.* **2016**, *6*, 39199.
[24] A. T. van der Goot, W. Zhu, R. P. Vázquez-Manrique, R. I. Seinstra, K. Dettmer, H. Michels, F. Farina, J. Krijnen, R. Melki, R. C. Buijsman, M. Ruiz Silva, K. L. Thijssen, I. P. Kema, C. Neri, P. J. Oefner, E. A. Nollen, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14912.
[25] S. W. L'Hernault, *Mol. Cell. Endocrinol.* **2009**, *306*, 59.
[26] B. M. Jackson, P. Abete-Luzi, M. W. Krause, D. M. Eisenmann, *G3 (Bethesda)* **2014**, *4*, 733.
[27] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen, C. von Mering, *Nucleic Acids Res.* **2015**, *43*, D447.
[28] H. Hutter, J. Suh, *Worm* **2016**, *5*, e1234659.
[29] J. A. Vizcaíno, R. G. Côté, A. Csordas, J. A. Dianes, A. Fabregat, J. M. Foster, J. Griss, E. Alpi, M. Birim, J. Contell, G. O'Kelly, A. Schoenegger, D. Ovelleiro, Y. Pérez-Riverol, F. Reisinger, D. Ríos, R. Wang, H. Hermjakob, *Nucleic Acids Res.* **2013**, *41*, D1063.