Supplemental Information for

Conservation of mRNA and protein expression during development of *C. elegans*

Dominic Grün, Marieluise Kirchner, Nadine Thierfelder, Marlon Stoeckius, Matthias Selbach, Nikolaus Rajewsky*

*corresponding author: rajewsky@mdc-berlin.de

Supplemental Experimental Procedures

Worm cultivation

C. elegans and *C. briggsae* were cultured as described previously (Brenner, 1974). Gravid adult worms were treated with bleach to isolate embryos (Lewis and Fleming, 1995). Part of the embryos were used for RNA and protein extraction and further analysis, while the remaining embryos were incubated in M9 buffer without food (Brenner, 1974) at room temperature for 12h to hatch and arrest at the L1 stage. The starved, synchronized L1 larvae were set out in S-medium (S-Basal supplemented with 3 mM $MgCl_2$, 3 mM CaCl, 10 mM potassium citrate, modified from (Sulston and Brenner, 1974), supplemented with a mix of antibiotics and antimycotics (Antibiotic-Antimycotic, Invitrogen, USA). Worms were fed with the *E. coli* strain AT713 and cultivated at 20°C. Staging was determined by examining at least 100 animals under an inverted microscope (Axio Observer, Carl Zeiss, Germany) at intervals for size determination, germline and vulval development. Animals were collected at the L1, L2, L3, early L4, late L4 and young adult (pre-gravid) stages and washed several times in M9 buffer. Exact times for growth after addition of food to starved L1s for wild type *C. elegans* N2 were as follows: L1 larvae were grown for 7h, L2 for 20h, L3 for 29h, early L4 for 41h, late L4 for 52h and young Adults for 57 h. For wild type *C. briggsae* AF16 the harvesting time points were: 6h for L1, 20h for L2, 28h for L3, 41h for early L4, 53h for late L4 and young adult after 63h. This timeline was repeated once to obtain biological replicate samples, the initial samples are referred to as biological replicate samples A, the samples from the second time course as biological replicate samples B. For temporal resolution experiments at embryo-to-L1 transition samples were collected in 3 hour intervals (0h: embryos; 3, 6, 9, 12, 15h). All samples were cleaned from bacteria and debris by either several M9 buffer washing steps or sucrose gradient centrifugation,

pelleted by centrifugation at 3000 rpm, 3 min, 4°C and stored in TRIzol Reagent (Invitrogen, USA), 1 ml per 100 µL worm pellet, at -80°C for RNA and protein isolation.

Caenorhabditis strains

The *C. elegans* strains that have been used in this study were obtained from the CGC, including wild type N2, BC10571 (*dlc-1* promoter fusion GFP), BC15026 (*elo-1* promoter fusion GFP), BC15277 (*grl-21* promoter fusion GFP), BS1080 (*gld-1* protein fusion GFP), OH3112 (*die-1* protein fusion GFP), OH4838 (*zig-1* protein fusion GFP), RT130 (*vit-2* protein fusion GFP), TX189 (*oma-1* protein fusion GFP), and SS104 (*glp-4* protein fusion GFP). The wild type *C. briggsae* strain that has been used in this study is AF16. The GFP fusion strains BS1080 and TX189 were donated by T.Schedl and R.Lin.

Stable isotope labeling by amino acids in nematodes

To label worm cultures, the arginine-lysine-auxotroph *E. coli* strain AT713 was used as food source for *C. elegans* and *C. briggsae*. Light labeled bacteria were obtained by cultivating AT713 in M9 minimal medium complemented with 0.2% glucose and 20 standard amino acids (Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val), obtained from Sigma-Aldrich, Inc., St. Louis, MO, USA, to a final concentration of 150 mg/L. For the culture of heavy labeled bacteria normal L- lysine was substituted by 150 mg/ml $^{15}N_2/^{13}C_6$- L-lysine (heavy lysine; Cambridge Isotope Laboratories, USA)(Hanke et al., 2008). All synchronously grown and staged worm samples were fed with light labeled bacteria. In detail, bacteria were added to the S-medium (final OD_{600} = 5.0) containing 1500 worms/ml. For worm cultures to be harvested at L1 stage OD_{600} of bacteria was 2.5 for 3000 worms/ml. Worm culture media was supplemented with Antibiotic-Antimycotic (Invitrogen, USA). The heavy labeled unsynchronous worm sample (heavy reference; containing worms of all stages) for SILAC protein quantification was produced by feeding worms with heavy labeled *E. coli*. As lysine is an essential amino acid for worms, heavy lysine is incorporated into the nematode proteome, resulting in complete heavy labeling after 2 generations. Worms

were cultivated on self-made peptone-free NGM plates (per 1L: 3g NaCl, 20g Agar, 1 ml 5 mg/ml cholesterol solution in 95% ethanol, 25 ml K_2PO_4, 1 ml $MgSo_4$, 1 ml $CaCl_2$, 10 ml Antibiotic-Antimycotic), seeded with an excess of heavy labeled bacteria: 35 ml of a culture with OD_{600} = 2.4, concentrated to 2.5 ml. After 3 generations of heavy labeling, adults were set out on fresh plates (50 per plate) and grown until the culture was unsynchronous, after 6 days of cultivation at 20°C. Worms were re-fed on day 4 to ensure that bacteria were present in excess at all times. Worms were harvested by washing them off the plates with M9 buffer. Animals were washed several times until all bacteria were removed, pelleted by letting them settle down in a falcon tube on ice or by centrifuging at 3000 rpm, 3 min, 4°C. Worm samples were stored in TRIzol Reagent (Invitrogen, USA) at -80°C (1 ml per 100 µL worm pellet) to be used for RNA and protein isolation. In total, we measured protein ratios versus heavy reference at the embryo-to-L1 transition in both nematode species for two independent biological samples (A, B). Each sample was measured in at least two replicates (A1, A2, A3 and B1, B2). For direct SILAC quantification of fold changes between embryos and L1 larvae, heavy and light labeled embryos and L1 staged larvae were obtained in a similar way, i. e. heavy and light worms. Worms were both cultivated on peptone-free NGM plates as described above, using heavy or light labeled *E. coli* AT713 as food source, respectively. Heavy embryos mixed 1:1 with light L1 larvae and vice versa are referred to as "direct" SILAC samples. They represent another set of biological replicates for the embryo-to-L1 transition and are later referred to as L/H (C) and H/L (D), respectively.

Protein Isolation

Proteins were isolated in parallel to RNA with TRIzol Reagent (Invitrogen, USA) from the exact same worm samples according to the instructions of the manufacturer with slight modifications. After complete removal of the aqueous phase, the organic and interphase were vortexed and proteins were precipitated by adding 4 volumes of ice-cold acetone, inverting 6 times and incubating at least 30 min at -20°C. Proteins were pelleted by centrifugation in an Eppendorf table top centrifuge (Eppendorf, Germany) at full speed (16.100 x *g*) at 4°C for 10 min. The remaining steps were done according to the protocol

of the manufacturer: The supernatant was removed and the protein pellet washed 3 times in a solution containing 0.3 M guanidine hydrochloride in 95% ethanol, followed by one final wash in 100% ethanol and vacuum drying the pellet for 10 min at 45°C in a speed vac (Concentrator plus, Eppendorf, Germany).

Mass Spectrometry

Sample preparation

Protein extracts obtained from TRIzol extraction were solubilized in 4% LDS loading buffer, sonicated and centrifuged for 10 min at 14000 rpm. Protein concentration of the supernatant was measured by amido black (Dieckmann-Schuppert and Schnittler, 1997). “Heavy” reference sample was mixed with each of the “light” staged samples at a 1:2 ratio. For the direct comparison of embryo and L1 larvae, protein extracts were mixed at a ratio of 1:1. For every sample approximately 150 µg total protein was separated under reducing conditions by SDS-Page on a 4–12% NuPAGE gradient gel (Invitrogen, USA) according to the manufacturer’s instructions. Proteins were fixed in 50% methanol, 10% acetic acid and stained by Colloidal Coomassie Blue (Invitrogen, USA). Gel lanes were cut into 16 slices and samples were processed following the standard in-gel digest protocol (Shevchenko et al., 2006) using Lysyl endopeptidase (LysC) (Wako, Osaka, Japan). Stop and go extraction (STAGE) tips containing C_{18} empore disks (3M, USA) were used to purify and store peptide extracts (Rappsilber et al., 2003).

LC – MS/MS

Online LC-MS/MS analysis was performed by separating peptide mixtures by reversed phase chromatography using the Eksigent NanoLC – 1D Plus system (Eksigent, Dublin, CA, USA) on in-house manufactured 10 cm fritless silica microcolumns with an inner diameter of 75 µm. Columns were packed with ReproSil-Pur C_{18}-AQ 1.9 µm resin (Dr. Maisch GmbH, Germany). Separation was performed using a 10–60% ACN gradient (240 min or 360 min) with 0.5% acetic acid at a flow rate of 200 nl/min. Eluting peptides were directly ionized by electrospray ionization and transferred into the orifice of a LTQ-

Orbitrap hybrid mass spectrometer (classic, XL or Velos instruments, Thermo Fisher, Waltham, MA, USA). Mass spectrometry was performed in the data dependent mode with one full scan in the Orbitrap (m/z = 300–1,700; R = 60,000; target value = 1 x 10^6). The five (Classic,XL) or twenty (Velos) most intense ions with a charge state greater than one were selected (target value 5,000; monoisotopic precursor selection enabled) and fragmented in the LTQ using CID (35% normalized collision energy, wideband activation enabled). Dynamic exclusion for selected precursor ions was 60s.

Processing of MS data

The MaxQuant software package (version 1.0.13.13) was used to identify and quantify proteins (Cox and Mann, 2008; Cox et al., 2009). SILAC duplets were extracted from isotope patterns, re-calibrated and quantified by the Quant module (heavy label Lys-8; maximum of four labeled amino acids per peptide; polymer detection enabled; top 6 MS/MS peaks per 100 Da). Peak lists were searched on a MASCOT search engine (version 2.2, MatrixScience, USA) against an in-house curated database for *C. elegans*, *C. briggsae* (see "Gene models in *C. elegans* and *C. briggsae*" and "Annotation of protein sequences") and *E. coli* (MG1655) plus common contaminants (e.g. trypsin, BSA). All protein sequences were also reversed to generate a target-decoy database (Elias and Gygi, 2007). Carbamidomethylation of cysteine was selected as fixed modification, oxidation of methionine and acetylation of the protein N-terminus were used as variable modifications. LysC was selected as protease (full specificity) with a maximum of 3 missed cleavages. A mass tolerance of 0.5 Da was selected for fragment ions. A minimum of six amino acids per identified peptide and at least one peptide per protein group were required. False discovery rate was set to 5% (FDR of 5) at both the peptide and protein level. Protein ratios were calculated from the median of all normalized peptide ratios using only unique peptides or peptides assigned to the protein group with the highest number of peptides ("Occam's razor" peptides). Only protein groups with at least three SILAC counts were considered for further analysis.

Western Blotting

To validate proteomics data, protein samples of staged *C. elegans* wild type strain N2 and *C. briggsae* wild type strain AF16 were analyzed by western blotting. In addition,

several GFP-protein fusion strains and GFP-promoter fusion strains were included (obtained from the CGC, USA). Worms were grown in liquid culture as described above, and harvested at the indicated time points. Harvested and washed worms were stored in TRIzol Reagent (Invitrogen, USA) at -80°C for subsequent protein isolation and western blotting. Several proteins were analyzed using either protein specific antibodies or anti-GFP antibody for translational or promoter-GFP fusion strains (see below). 20 µg total protein for each stage was separated by SDS Page and gels were blotted onto PVDF membrane using a semi-wet system according to the manufacturer's instructions (Invitrogen, USA). Unspecific binding sites were blocked with 3% milk powder overnight at 4°C, all primary antibodies were applied at a 1:1000 dilution in TBST (140 mM NaCl, 25 mM Tris-HCl, 0.1 % Tween 20, pH 7.4) for 3h at room temperature, blots were washed 3x in TBST and incubated either with an anti-mouse, anti-rabbit or anti-goat secondary antibody (GE Healthare Europe GmbH, Germany; Santa Cruz Biotechnology, USA) conjugated to horseradish peroxidase diluted 1:5000 in TBST for 1h. After three additional washing steps in TBST the bound secondary antibodies were detected with the Western Blot Chemiluminescence Reagent Plus for ECL immunostaining (PerkinElmer, USA). Autoradiographs were scanned and quantitative analysis was performed using ImageJ software. Mean values and standard deviation were calculated for two independent experiments, and normalized against TBA-2 or HSP-60 (as indicated).

RNA Isolation

Total RNA isolation was performed with TRIzol Reagent (Invitrogen, USA) by following the manufacturer's instructions. In brief, 10 volumes of TRIzol Reagent were added per volume of pelleted worms. After two rounds of homogenizing in the Precellys 24-Dual homogenizer (PEQLAB Biotechnologie GmbH Germany) for 5s at 6000 rpm. RNA was precipitated with isopropanol and the pellet washed with 75% ethanol. RNA concentration was measured by means of absorption spectrometry at a wavelength of 260 nm in a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, USA).

RNA integrity was determined by capillary gel electrophoresis on a Bioanalyzer (Agilent, USA).

Construction of RNA-seq libraries

For each library, polyA mRNA was purified from 10 µg of total RNA using the Dynabeads mRNA Purification Kit from Total RNA preps (Invitrogen, USA). Isolated mRNA was subsequently mixed with 5X Fragmentation buffer (200 mM Tris acetate pH 8.2, 500 mM potassium acetate, 150 mM magnesium acetate) and incubated for exactly 3.5 min at 94°C in a thermocycler, the reaction was stopped by immediately chilling the samples on ice. Cleaved mRNAs were precipitated overnight at -80°C, pelleted and washed with 70% ethanol, quality of the RNA fragmentation was checked using the RNA 6000 Pico Chip Kit on the Agilent 2100 Bioanalyzer (Agilent Technologies, USA). First strand cDNA synthesis was accomplished using Superscript II Reverse Transcriptase and random primers, being followed by second cDNA synthesis using DNA Polymerase I and RNaseH (all reagents from Invitrogen, USA). Double-stranded DNA was purified with the QIAquick PCR purification kit (Qiagen GmbH, Germany) according to the manufacturer's instructions and eluted in 50 µL of EB. cDNA quality was checked on the Agilent 2100 Bioanalyzer with the Agilent DNA 1000 kit (Agilent Technologies, USA), and concentration was determined spectrophotometrically on a NanoDrop ND-1000 (PEQLAB Biotechnologie GmbH, Germany). Library preparation was then performed using the Genomic DNA Sample Prep Kit (Illumina, USA). In short, cDNA was end-repaired by incubating the samples for 30 min at 20°C with T4 DNA polymerase, Klenow DNA Polymerase and T4 polynucleotide kinase followed by purification on a QIAquick spin column according to the instructions of the QIAquick PCR purification kit (Qiagen GmbH, Germany). Subsequently, cDNA products were incubated with Klenow fragment (3' to 5' exo minus) to add a single "A" base to the 3' ends, samples were then purified via the MinElute PCR Purification Kit according to the manufacturer's protocol (Qiagen GmbH, Hilden, Germany). Adapters were ligated to the DNA fragments using the PE Adapter Oligo Mix and purified on a QIAquick MinElute column as mentioned before. cDNA templates were size fractionated on 2% agarose gels (certified low range ultra

agarose from Biorad, Germany), the 270 to 350 bp fractions were excised and cDNA was extracted using the QIAGEN Gel Extraction kit according to the manufacturer's protocol (Qiagen GmbH, Germany). Adapter-modified DNA fragments were then enriched by PCR with Phusion DNA Polymerase using Illumina's genomic DNA primer set to create the final cDNA library. As a last step PCR products were purified on QIAquick spin columns (Qiagen GmbH, Germany), DNA quality and quantity was assessed by using the Agilent DNA 1000 kit (Agilent Technologies, USA) and NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies, USA) and diluted to 10 nM. For replicate samples, the protocol was slightly modified to start with a reduced amount of material (1 µg of total RNA instead of 10 µg): After the mRNA fragmentation step, RNA was recovered by using Agencourt RNAClean beads (Beckman Coulter Genomics, Germany). For all following cDNA purification steps, Agencourt AMPure XP beads (Beckman Coulter Genomics, Germany) were used, replacing the purification via QIAquick columns (Qiagen GmbH, Germany). Cluster generation and sequencing was performed on the Illumina cluster station and Genome Analyzer IIx according to the manufacturer's instructions. Read lengths were 76 bases. RNA sequencing was performed across all stages for one sample (A) and in replicate for two independent biological samples (A and B) for embryos and L1 larvae. The samples correspond to the exact same samples used for proteomics. For sample A, two technical replicates for embryos, L1 larvae and reference samples with independently prepared libraries were sequenced (A1, A2).

Reverse Transcription

For First Strand cDNA synthesis, 1.0 to 2.5 µg of total RNA was reverse transcribed using SuperScript III Reverse Transcriptase and anchored oligo dT_{20} primers following the manufacturer's instructions (Invitrogen, USA).

Quantitative Reverse Transcription PCR

TaqMan assays were performed for selected orthologous target genes in *C. elegans* and *C. briggsae* for samples of different time courses. Primers and TaqMan probes were designed using the Primer3 software, primers having T_ms of approximately 60°C,TaqMan probes 70°C (Breslauer et al., 1986). Primers were designed to span 2 exons and TaqMan probes were labeled with FAM (5') and TAMRA (3'). Primers and probes were ordered from the companies Biotez (Berlin, Germany) or metabion (Munich, Germany). TaqMan assays were performed on a StepOnePlus Real-Time PCR System from Applied Biosystems using the TaqMan Universal PCR Master Mix (2X) (Applied Biosystems, USA). Each assay was performed as 20 µl reaction with 30 ng single stranded cDNA using the "Standard Run" program: 2 min 50°C, 10 min activation of the AmpliTaq Gold enzyme (present in the TaqMan Universal PCR Master Mix) at 95°C followed by 40 cycles of 15 sec 95°C denaturation, 1 min 60°C of annealing and elongation. Quantification of gene expression was accomplished via the Comparative C_T method ($\Delta\Delta C_T$) (Schmittgen and Livak, 2008), whereas the reference nematode samples were set as "reference" sample. As endogenous control, the genes *Y57G11C.34* and *Cbr-ero-1* were picked for *C. elegans* and *C. briggsae*, respectively. Each experiment was performed in 3 technical replicates.

For RT-qPCRs with SYBR green to test orthologous target genes of samples from the 3h time courses, primers were designed using Primer3 as described above and ordered at the mentioned companies. SYBR green RT-qPCRs were run on the StepOnePlus Real-Time PCR System from Applied Biosystems (Applied Biosytems, USA) using the Maxima SYBR Green/Rox qPCR Master Mix (2X) from Fermentas (Fermentas GmbH, Germany). Starting material per assay were 10 ng single stranded cDNA, each performed in 3 technical replicates. The standard run method for SYBR green Reagents was used: 10 min 95°C activation of the Maxima Hot Start DNA Polymerase, 40 cycles of 15 sec 95°C denaturation and 1 min of annealing and elongation at 60°C, followed by a melt curve stage of 15 sec 95°C, 1 min 60°C, 15 sec 95°C. Quantification of gene expression was accomplished via the Comparative C_T method ($\Delta\Delta C_T$) (Schmittgen and Livak, 2008) as described above, endogenous controls were *tba-2* for *C. elegans* and *Cbr-ero-1* for *C. briggsae*.

Daf-16 RNAi

Daf-16 RNAi was performed by feeding worms on dsRNA expressing bacteria using the *daf-16* clone from the Ahringer RNAi library as described previously (Fraser et al., 2000; Kamath et al., 2003). Bacterial containing the empty L4440 vector was used as the negative RNAi control. L1 stage N2 worms were plated onto seeded and induced RNAi plates and grown at 20˚C for three days until they were gravid. Embryos were then extracted by bleaching and processed as described above.

Antibodies for Western blotting

Primary antibodies: α-ASNA-1 (Rabbit, Simon Tuck, Umeå University, Sweden), α-DAF-18 (Rabbit, Jeff Boudreau, Queen's University, Kingston, Canada), α-DAF-16 (Rabbit, Santa Cruz Biotechnology, sc-33738, polyclonal, raised against C-terminal epitopes covering amino acids 211-510), α-DAO-5 (Mouse, The Developmental Studies Hybridoma Bank, Iowa City, USA), α-DYN-1 (Mouse, The Developmental Studies Hybridoma Bank, Iowa City, USA), α-ERM-1 (Mouse, The Developmental Studies Hybridoma Bank, Iowa City, USA), α-GFP (Mouse, Roche Diagnostics GmbH, Mannheim, Germany), α-HEL-1 (Rabbit, Tom Blumenthal, University of Colorado, Denver, USA), α-HSP-60 (Mouse, The Developmental Studies Hybridoma Bank, Iowa City, USA), α-HTZ-1 (Mouse, William Kelly, Emory University, Atlanta, USA), α-INX-13 (Rabbit, Todd Starich, University of Minnesota, USA), α-ISW-1 (Rabbit, Robert Horvitz, MIT, Cambridge, USA), α-MYO-2 (Mouse, David Miller, Vanderbilt University, Nashville, USA), α-MEI-1 (Goat, Santa Cruz Biotechnology, sc-32420), α-MEX-1 (Goat, Santa Cruz Biotechnology, sc-9239), α-PIE-1 (Goat, Santa Cruz Biotechnology, sc-32420), α-RME-1 (Mouse, The Developmental Studies Hybridoma Bank, Iowa City, USA), α-TBA-2 (Mouse, The Developmental Studies Hybridoma Bank, Iowa City, USA), α-TNI-4 (Rabbit, Hiroaki Kagawa, Okayama University, Okayama, Japan), α-TOP-1 (Rabbit, Hyeon-Sook Koo, Yonsei University, Seoul, Korea), α-VHA-8 (Rabbit, Joohong Ahnn, Hanyang University, Seoul, Korea). Secondary antibodies: ECL Mouse IgG, HRP-Linked (Sheep, GE Healthcare Europe GmbH, Munich, Germany), ECL Rabbit IgG, HRP-

Linked (Donkey, GE Healthcare Europe GmbH, Munich, Germany), donkey anti-goat IgG-HRP (Donkey, Santa Cruz Biotechnology, Inc., Santa Cruz, USA)

Gene models in C. elegans and C. briggsae

In *C. elegans*, we used revised and extended gene models (Gerstein et al., 2010), comprising 64,826 transcripts which correspond to 21,774 different genes. To annotate transcripts in *C. briggsae*, we ran the spliced read alignment software TopHat (Trapnell et al., 2009) on the RNA-seq data of all different stages and the reference sample and input the pooled read alignment results to the transcript prediction software Cufflinks (Trapnell et al., 2010) with disabled coverage search. We only kept predictions supported by a reference annotation, produced by mapping the nucleotide sequence of all coding exons of *C. elegans* to *C. briggsae*, using the spliced alignment software SPALN (Gotoh, 2008). For the comparison to the reference annotation we used the software Cuffcompare, which is part of the Cufflinks software package. To further increase the sensitivity of our transcript predictions, we downloaded the genomic coordinates of the tblastn mappings for all *C. elegans* proteins to the *C. briggsae* genome from the UCSC genome browser (Kent et al., 2002). The Cufflinks predictions supported by this reference annotation were added to our set of *C. briggsae* transcripts if an ortholog for the corresponding gene had not been detected in the previous comparison to the more specific SPALN mappings. In total, we annotated 35,577 transcript variants in *C. briggsae*, corresponding to 13,938 genes. We defined a gene locus as the projection of all isoforms of a given gene to the genome. Gene loci in *C. elegans* and *C. briggsae* cover on average 1,413 and 1,743 nucleotides, respectively. The fraction of mapped reads was comparable for the two species, indicating a high quality of the *C. briggsae* transcript annotation (Table S1). Transcript expression was aggregated for each gene locus and measured in reads per kilobase of gene locus per million mapped reads (RPKM) (Pepke et al., 2009). For a direct comparison of transcript and protein expression, we quantified protein expression by the intensity of the staged light samples. Protein fold changes at developmental transitions, on the other hand, were always computed based on SILAC ratios.

Annotation of protein sequences

For the annotation of protein sequence, we translated the transcript sequence in all 6 frames and extracted the longest consecutive open reading frame, uninterrupted by a stop codon.

Quantification of transcript expression

Reads of all samples were mapped to the transcriptome using the read alignment software bwa (Li and Durbin, 2010). Prior to read mapping, consecutive strings of basecalls with lowest Phred quality score were removed from the 3' end of the reads and trimmed reads shorter than 30 bases were discarded. For bwa, a minimum seed length of 30 bases was required and default parameters were used otherwise. Reads mapping to multiple loci with the same edit distance were assigned to all of these loci with reduced weight given by one over the number of loci. A gene locus was defined by the projection of all isoforms to the genome. The number of reads mapping to a gene locus was determined by aggregating the number of reads mapping to all isoforms. Expression of a gene locus was quantified in reads per kilobase of gene locus sequence per million mapped reads (RPKM) (Pepke et al., 2009). Consensus expression across replicates of a given stage was calculated by transforming the weighted average of normalized read counts into RPKM. Read counts of a given replicate were normalized by the total number of reads for this replicate and weights were determined by dividing the number of reads of a given replicate by the sum of reads across all replicates. The standard deviation was computed accordingly and used as an estimator for the uncertainty of expression. Since both technical and biological replicates were available, we first computed the average across technical replicates of a given sample and thereafter averaged the consensus of technical replicates across different biological samples. Reproducibility of transcript quantification by RNA-seq is supported by a high correlation of $\log_2$-fold changes between technical replicates (sample A1, A2) at the embryo-L1 transition (ρ=0.83). A comparable correlation between biological replicates (B and A1: ρ=0.89; B and A2: ρ=0.91, Figure S2A) indicates that the predominant source of

noise is variability induced by the process of library preparation whereas actual variation of transcript abundance between samples was low.

Quantification of protein abundance

Quantification of protein abundance was based on intensity values for protein groups calculated by MaxQuant (Cox and Mann, 2008; Cox et al., 2009), which correspond to the sum of intensities measured for all peptides belonging to a protein group. From each protein group, we only selected the gene corresponding to the first isoform of the group, which represents all peptides identified for this group. Accordingly, we selected SILAC ratios to compute fold changes between different stages either directly, if no reference sample was used, or by taking ratios of ratios for reference based quantification. Consensus fold changes across replicates of a given sample were calculated by averaging SILAC ratios and the standard deviation was used as an estimator of uncertainty of protein fold changes. Technical replicates were averaged first and consensus across different biological samples was thereafter calculated by averaging consensus of technical replicates across biological samples.

Correlation between expression profiles and clustering

Expression of transcripts was normalized by the sum of expression values across all stages. We used SILAC ratios versus reference sample to estimate relative protein expression after normalizing by the sum of SILAC ratios across all stages. To compute correlation of expression profiles of two genes we concatenated for each gene the vectors with normalized expression of transcripts and proteins and calculated Pearson's correlation coefficient (ρ_P) of the concatenated vectors corresponding to the two genes. Hierarchical clustering of temporal expression profiles for all genes with quantified protein abundance for at least three developmental stages was performed using a gene-to-gene distance given by ($1-\rho_P$). We inspected all clusters at different levels of the hierarchy and decided that a level of seven clusters provides sufficient resolution to discriminate all fundamentally distinct expression profiles.

Annotation of RNAi phenotypes

We downloaded annotated RNAi phenotypes from wormbase (WS190) and extracted all phenotypes screened in at least two experiments. We only considered a phenotype when it was observed in more than 50% of the experiments. The final list was then mapped to the gene models used in this study.

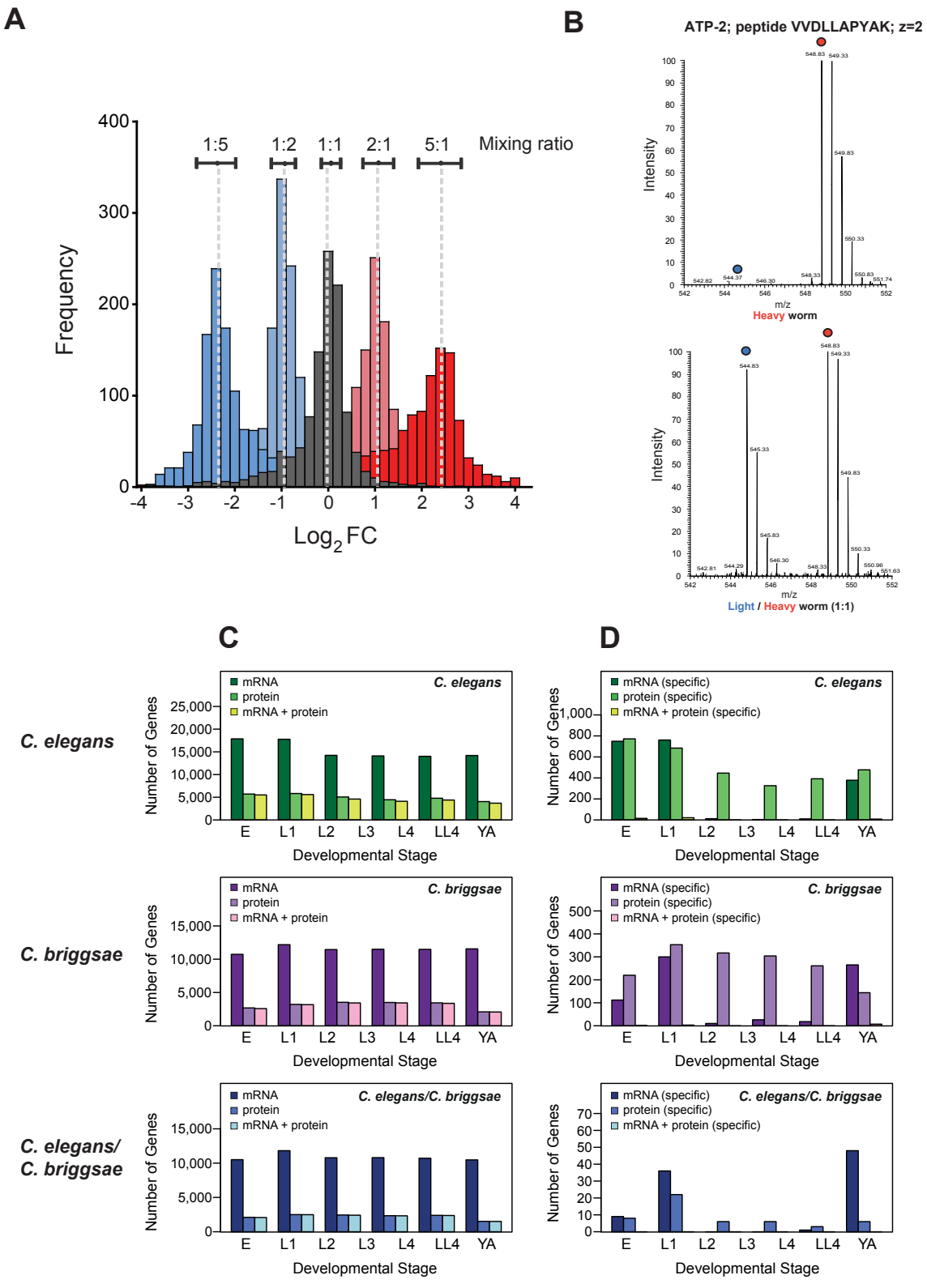
Identification of germline specific genes

To identify germline expressed genes we used RNA-seq derived transcript expression in young adults for wild type worms and a temperature sensitive mutant strain, *glp-4* (bn2ts), which develops without a germline. Sequencing and mapping was done as described elsewhere (Jungkamp et al., 2011). We then extracted all genes with more the 25 mapped reads and computed fold changes for all gene loci between wild type and mutants. The group of genes with fourfold up-regulation in wild type was considered to be specifically expressed in the germline, while genes with fourfold down-regulation in wild type where assumed to be specific to somatic cell lineages (Table S5). These two groups were used in the comparison to genes with differential expression in LL4 larvae and young adults.

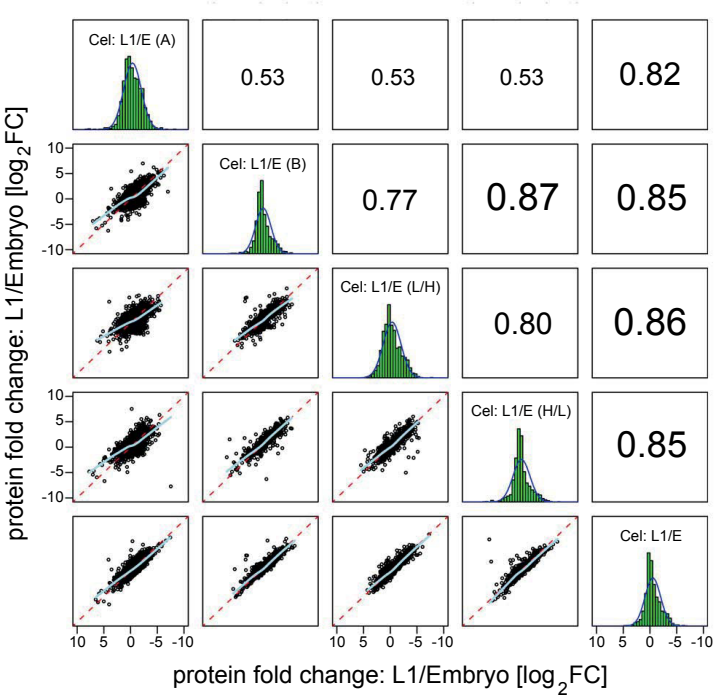
Identification of zygotically transcribed genes

In order to identify embryonically expressed genes that were not maternally deposited but newly transcribed in the zygote after zygotic genome activation, we computed transcript fold changes between the mixed embryo population and the 1-cell embryo. Quantification of transcript expression for the 1-cell embryo by RNA-seq was done in our lab (Stoeckius et al., *in preparation*). The list of all *C. elegans* genes was ordered by decreasing transcript fold changes between mixed embryos and 1-cell embryos and the top half of the list was considered as zygotically transcribed genes whereas the bottom half was assumed to comprise preferentially maternal genes.

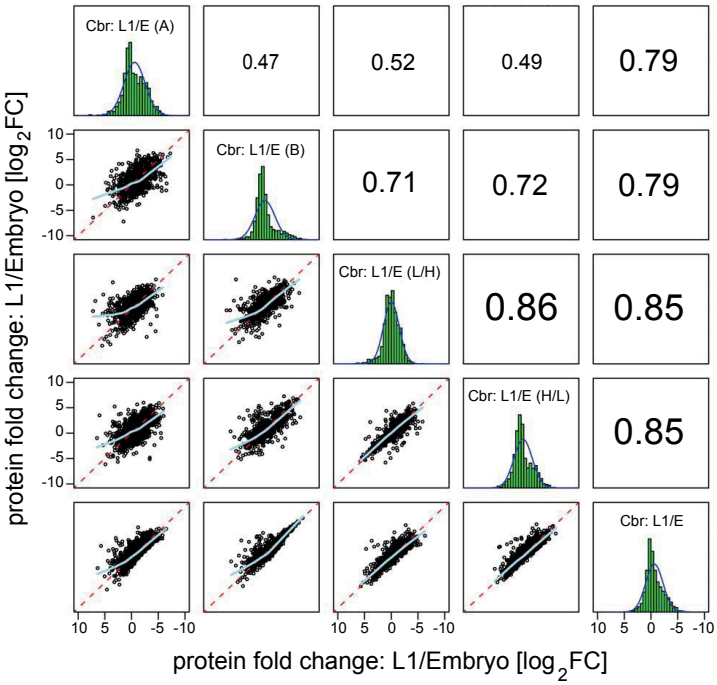
Supplemental Figures

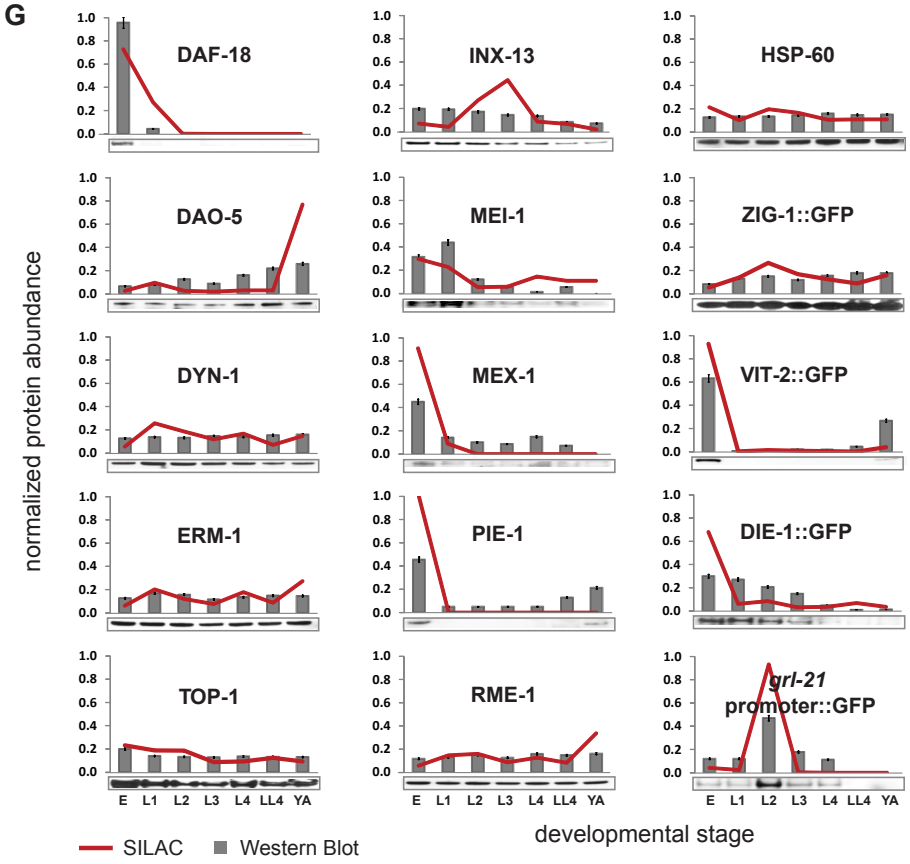

A
Frequency
Log2 FC
1:5
1:2
1:1
2:1
5:1
Mixing ratio
B
ATP-2; peptide VVDLLAPYAK; z=2
Intensity
m/z
Heavy worm
Light / Heavy worm (1:1)
C
D
C. elegans
C. briggsae
C. elegans/ C. briggsae
mRNA
protein
mRNA + protein
mRNA (specific)
protein (specific)
mRNA + protein (specific)
Number of Genes
Developmental Stage
E
L1
L2
L3
L4
LL4
YA

E


Cel: L1/E (A)
0.53
0.53
0.53
0.82
Cel: L1/E (B)
0.77
0.87
0.85
Cel: L1/E (L/H)
0.80
0.86
Cel: L1/E (H/L)
0.85
Cel: L1/E
protein fold change: L1/Embryo [log2FC]
protein fold change: L1/Embryo [log2FC]

F


Cbr: L1/E (A)
0.47
0.52
0.49
0.79
Cbr: L1/E (B)
0.71
0.72
0.79
Cbr: L1/E (L/H)
0.86
0.85
Cbr: L1/E (H/L)
0.85
Cbr: L1/E
protein fold change: L1/Embryo [log2FC]
protein fold change: L1/Embryo [log2FC]


G
normalized protein abundance
DAF-18
INX-13
HSP-60
DAO-5
MEI-1
ZIG-1::GFP
DYN-1
MEX-1
VIT-2::GFP
ERM-1
PIE-1
DIE-1::GFP
TOP-1
RME-1
grl-21 promoter::GFP
1.0
0.8
0.6
0.4
0.2
0.0
E L1 L2 L3 L4 LL4 YA
developmental stage
SILAC
Western Blot

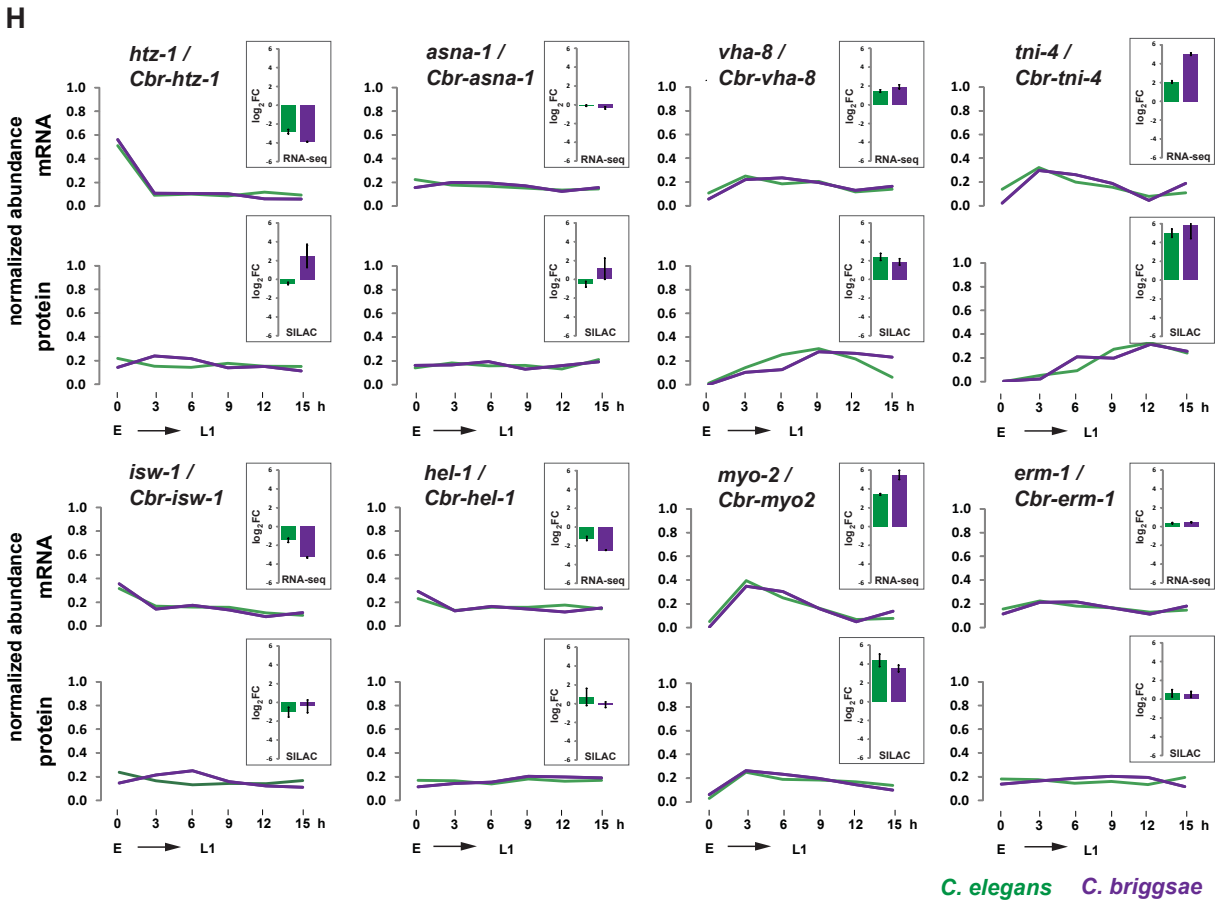
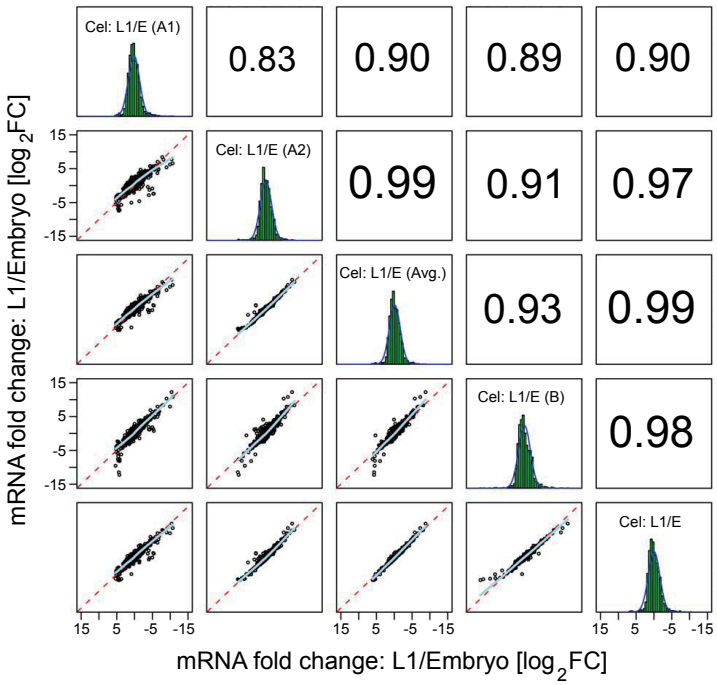


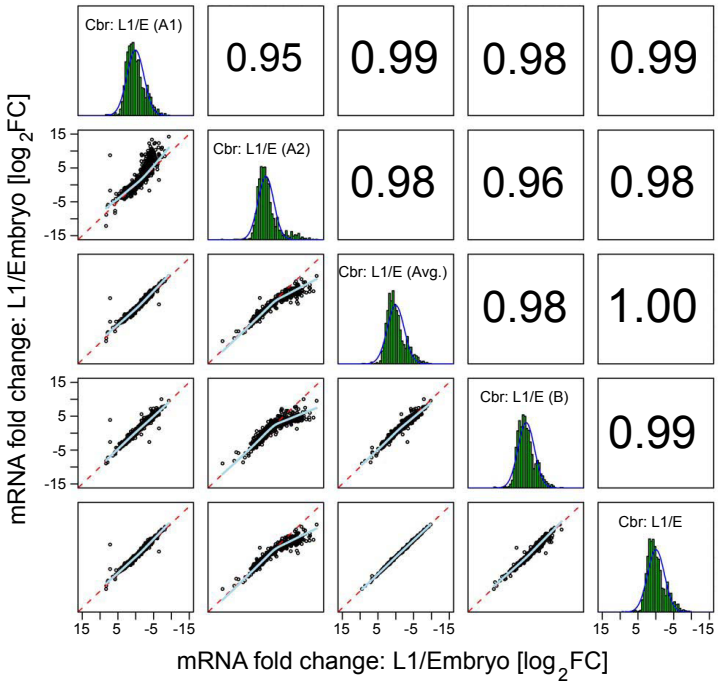
Figure S1, Related to Figure 2. Accuracy of *in vivo* SILAC quantification. (A) Histogram shows distribution of protein log_2-ratios from heavy labeled reference mixed with light L1 larvae protein lysate in defined ratios (1:5, 1:2, 1:1, 2:1 and 5:1) compared to a 1:1 mix control sample. The mean log_2-ratio value for each sample correlates well with theoretical numbers, demonstrating accuracy of SILAC. (B) Representative mass spectra of the ATP-2 peptide VVDLLAPYAK from heavy labeled *C. elegans* protein extract (top) and a 1:1 mixture of light and heavy labeled *C. elegans* protein extract (bottom). Red circles mark the heavy peptide and blue circles the light peptide, respectively. The mass difference is 8 Da (m/z = 4 Th). For the heavy labeled sample (top) the light form of the peptide is undetectable. Peptide intensities of the 1:1 (L/H) mixed sample (bottom) reflect the applied mixing ratio. Graphs show full scan of isotope cluster of the ATP-2 VVDLLAPYAK peptide. (C) Histograms indicating the number of genes with identified transcripts, proteins and both, transcripts and proteins in *C. elegans* (upper panel, in green), in *C. briggsae* (middle panel, in purple) and the overlap

between *C. elegans* and *C. briggsae* (lower panel, in blue). (D) Same as in (C), but only genes specific to each stage are shown. (E-F) Log_2-fold changes of proteins at the embryo-L1 transition computed from SILAC ratios for all biological replicates (reference based: A, B; directly measured transitions: H/L with labeled larvae and unlabeled embryos and L/H with unlabeled larvae and labeled embryos) and for the average of all samples (lower right corner) are shown. The diagonal panels display histograms of fold changes, above the diagonal the correlation between replicates is given and below the diagonal scatter plots are depicted with a moving average (blue line) and the diagonal (red broken line). (E) *C. elegans.* (F) *C. briggsae*. Ratios were averaged across technical replicates of each biological sample yielding consensus ratios for sample A and B. Log_2-fold changes computed from these ratios exhibit only a moderate correlation. However, correlation between sample B and either sample L/H and H/L is substantially higher. A likely reason for the low correlation of sample A to all other samples is an improved protein isolation protocol that has been applied for all other samples. (G) Relative expression profiles during development are shown for selected *C. elegans* proteins based on either SILAC (red lines) or western blot (grey bars) quantification. SILAC and western blot values were obtained from independent biological replicate samples; values are normalized against tubulin (TBA-2). (H) Comparison of relative expression profiles during early development covering the embryo-L1 transition in 3h resolution in *C. elegans* (green) and *C. briggsae* (purple). Profiles of transcript expression (upper panels) were obtained from RT-qPCR measurements, profiles of protein expression were obtained by western blot quantification (lower panels). RT-qPCR values were normalized against TBA-2 mRNA for *C. elegans* and Cbr-ERO-1 mRNA for *C. briggsae*, western blot values were normalized against HSP-60. The bar charts show the log_2-fold change values from RNA-seq measurements (upper panels) and SILAC quantification (lower panels) for the embryo-L1 transition. High-resolution 3h-profiles from RT-qPCRs and westerns could overall validate original gene expression data obtained by RNA-seq and SILAC. Expression profiles of the 8 chosen genes are very similar in both species.

A


Cel: L1/E (A1)
0.83
0.90
0.89
0.90
Cel: L1/E (A2)
0.99
0.91
0.97
Cel: L1/E (Avg.)
0.93
0.99
Cel: L1/E (B)
0.98
Cel: L1/E
15
5
-5
-15
mRNA fold change: L1/Embryo [log2FC]
mRNA fold change: L1/Embryo [log2FC]

B


Cbr: L1/E (A1)
0.95
0.99
0.98
0.99
Cbr: L1/E (A2)
0.98
0.96
0.98
Cbr: L1/E (Avg.)
0.98
1.00
Cbr: L1/E (B)
0.99
Cbr: L1/E
15
5
-5
-15
mRNA fold change: L1/Embryo [log2FC]
mRNA fold change: L1/Embryo [log2FC]

C

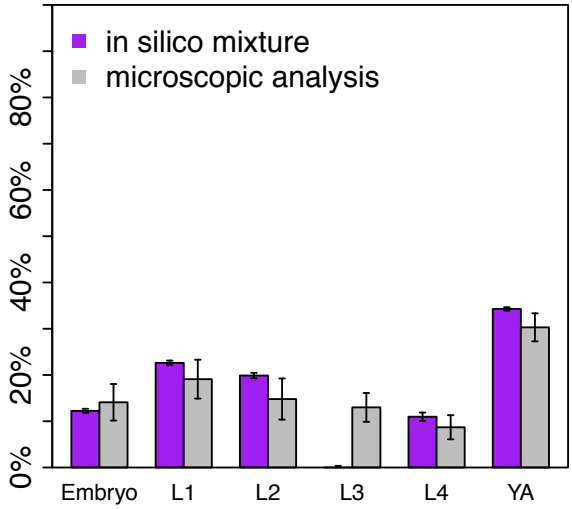
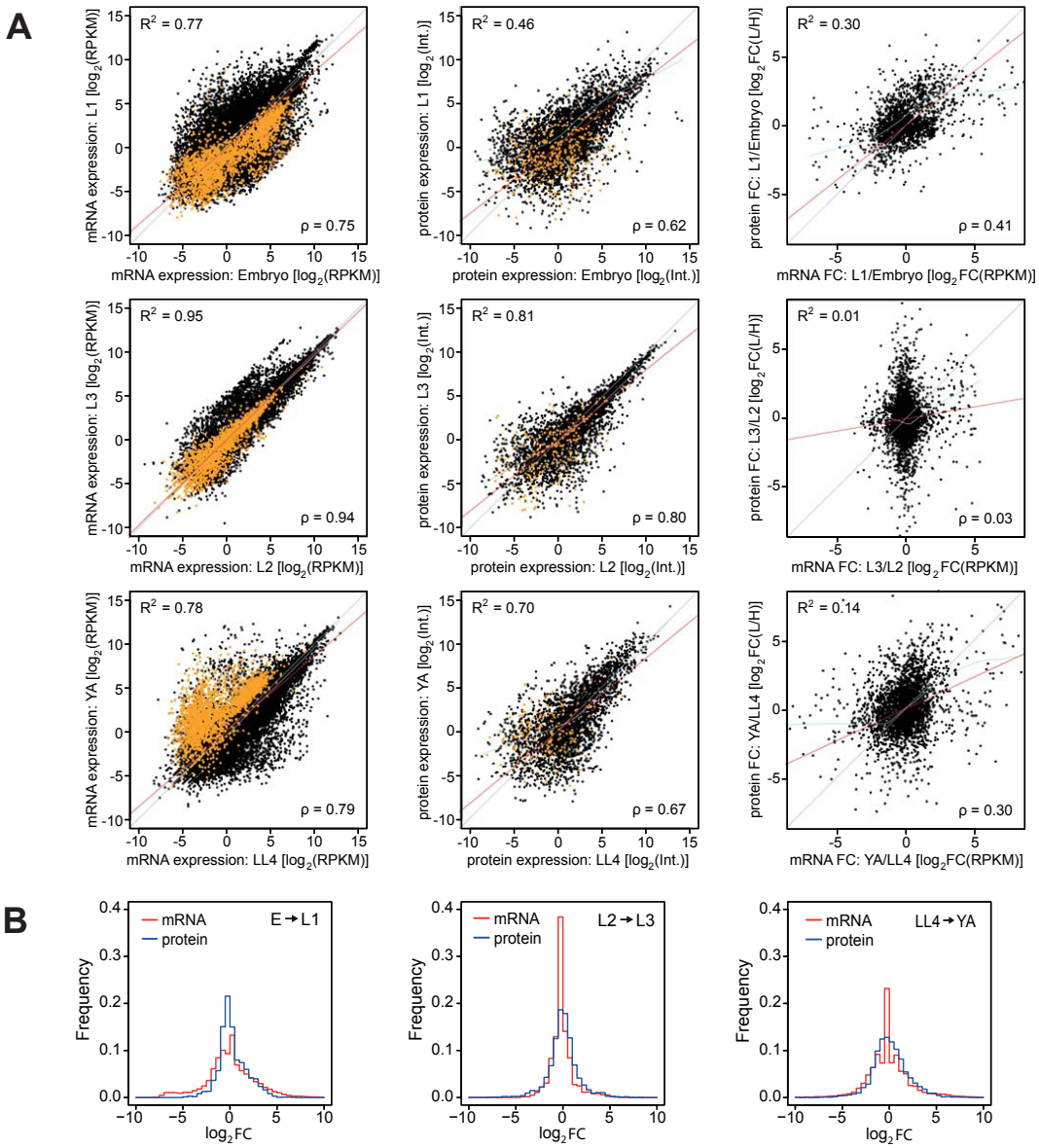
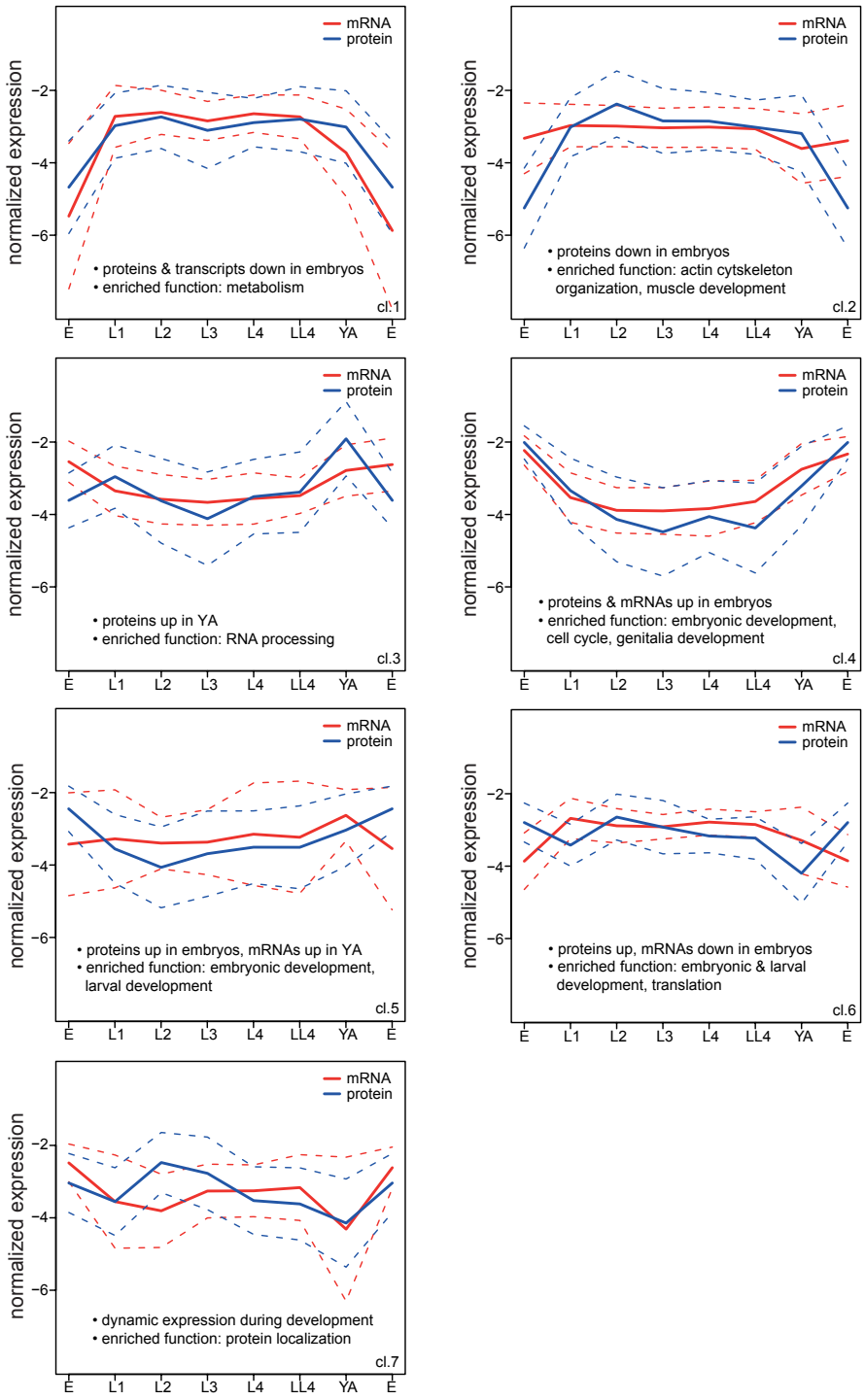


Figure S2, Related to Figure 3. Correlation of transcript fold changes across biological replicates. Fold changes of transcripts at the embryo-L1 transition computed from RPKM for all replicates (biological replicates: A, B; technical replicates of A: A1, A2; Avg. denotes the average of A1 and A2) and for the average of all samples (lower right corner) are shown. The diagonal panels display histograms of fold changes, above the diagonal the correlation between replicates is shown and below the diagonal scatter plots are depicted with a moving average (blue line) and the diagonal (red broken line). (A) *C. elegans.* (B) *C. briggsae*. (C) Comparison of the fraction of total mRNA from each stage in the reference sample of *C. briggsae* as computed by the *in silico* mixture to estimates based on microscopic analysis (n>1000). Only the L3 stage is under-represented among the contributions estimated by the regression leading to an overestimation of the other larval stages. Error bars indicate the error of the regression coefficients and the uncertainty of our estimates, respectively.


A
R² = 0.77
ρ = 0.75
mRNA expression: L1 [log2(RPKM)]
mRNA expression: Embryo [log2(RPKM)]
R² = 0.46
ρ = 0.62
protein expression: L1 [log2(Int.)]
protein expression: Embryo [log2(Int.)]
R² = 0.30
ρ = 0.41
protein FC: L1/Embryo [log2FC(L/H)]
mRNA FC: L1/Embryo [log2FC(RPKM)]
R² = 0.95
ρ = 0.94
mRNA expression: L3 [log2(RPKM)]
mRNA expression: L2 [log2(RPKM)]
R² = 0.81
ρ = 0.80
protein expression: L3 [log2(Int.)]
protein expression: L2 [log2(Int.)]
R² = 0.01
ρ = 0.03
protein FC: L3/L2 [log2FC(L/H)]
mRNA FC: L3/L2 [log2FC(RPKM)]
R² = 0.78
ρ = 0.79
mRNA expression: YA [log2(RPKM)]
mRNA expression: LL4 [log2(RPKM)]
R² = 0.70
ρ = 0.67
protein expression: YA [log2(Int.)]
protein expression: LL4 [log2(Int.)]
R² = 0.14
ρ = 0.30
protein FC: YA/LL4 [log2FC(L/H)]
mRNA FC: YA/LL4 [log2FC(RPKM)]
B
mRNA
protein
E→L1
L2→L3
LL4→YA
Frequency
log2FC

C


mRNA
protein
normalized expression
−2
−4
−6
E
L1
L2
L3
L4
LL4
YA
E
• proteins & transcripts down in embryos
• enriched function: metabolism
cl.1
• proteins down in embryos
• enriched function: actin cytskeleton organization, muscle development
cl.2
• proteins up in YA
• enriched function: RNA processing
cl.3
• proteins & mRNAs up in embryos
• enriched function: embryonic development, cell cycle, genitalia development
cl.4
• proteins up in embryos, mRNAs up in YA
• enriched function: embryonic development, larval development
cl.5
• proteins up, mRNAs down in embryos
• enriched function: embryonic & larval development, translation
cl.6
• dynamic expression during development
• enriched function: protein localization
cl.7

D

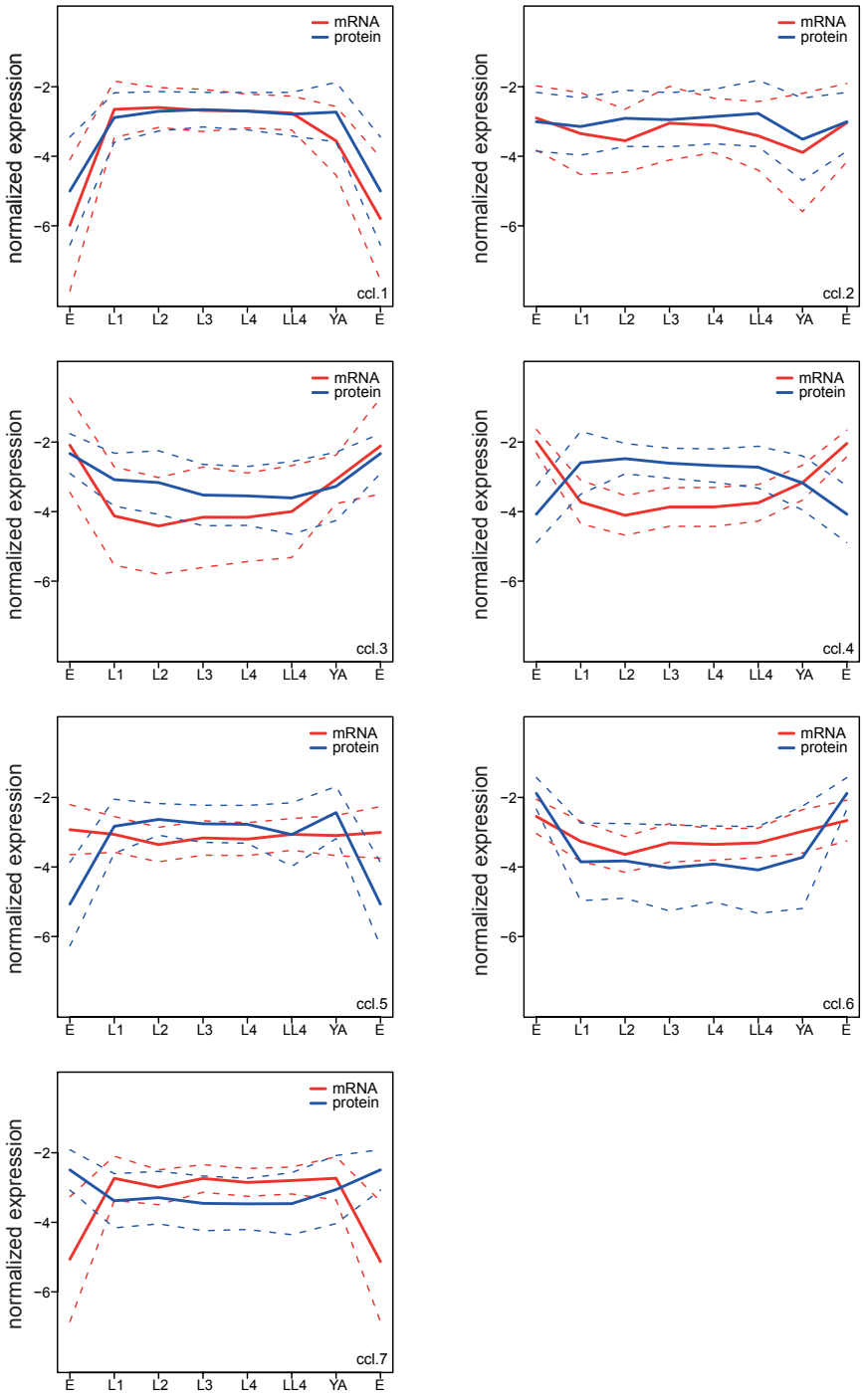
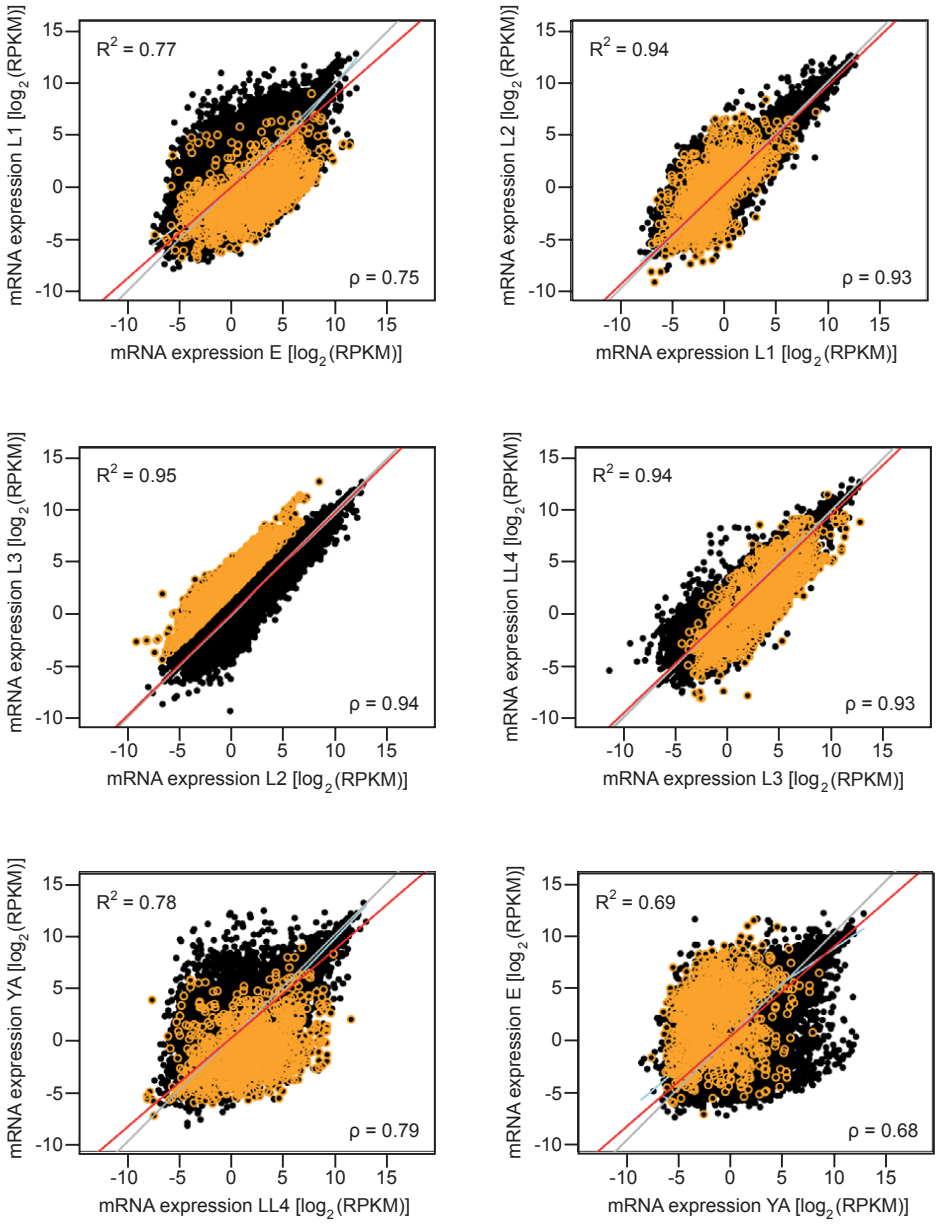
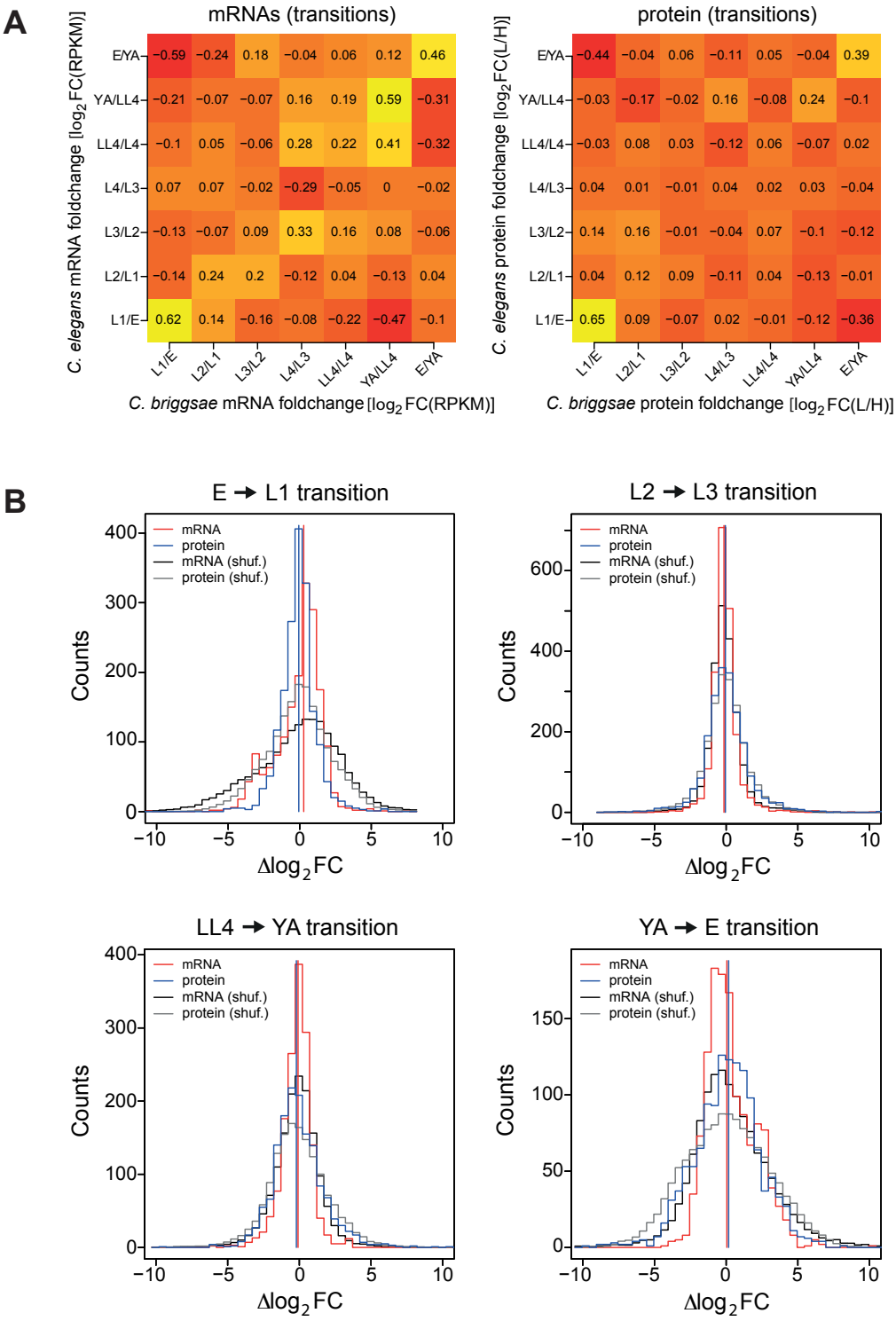

mRNA
protein
normalized expression
−2
−4
−6
E
L1
L2
L3
L4
LL4
YA
E
ccl.1
ccl.2
ccl.3
ccl.4
ccl.5
ccl.6
ccl.7

Figure S3, Related to Figure 4. Correlation of transcript and protein abundances between developmental stages and clusters of correlated gene expression. (A) Scatter plots of log_2-expression at subsequent developmental stages for the embryo-L1 (upper panel), the L2-L3 (middle panel) and the LL4-young adult transition (lower panel). From left to right each panel shows the correlation of transcript expression, the correlation of protein expression measured by normalized intensities, and the correlation between transcript and protein fold changes (measured by SILAC ratios) at the respective transition. The gray line represents the diagonal, the red and the blue line correspond to a linear regression and a moving average, respectively. Spearman's correlation coefficient is indicated in the lower right corner (ρ). Genes preferentially expressed in the germline are marked in orange. (B) Histogram of transcript and protein log_2-fold changes in *C. elegans* at the embryo-L1 (left), the L2-L3 (middle) and the LL4-young adult transition (right). (C-D) Hierarchical clustering was performed based on mutual gene distances quantified by Pearson's correlation coefficient of normalized expression profiles (see Supplemental Experimental Procedures). The solid red and blue lines indicate the average normalized expression of transcripts and proteins, respectively. The broken lines indicate the standard deviation of expression within each cluster. Units on the y-axis are normalized expression values computed by normalizing for each gene the sum of SILAC ratios or RPKM, respectively, across all stages to one. (C) In *C. elegans*, seven clusters with clearly distinct expression profiles were recovered (cl.1, …, cl.7). Details of these clusters are given in Table S3A. (D) In *C. briggsae*, seven distinct clusters were obtained (ccl.1, …, ccl.7). Details of the clusters are given in Table S3B and the overlap of these clusters with clusters cl.1, ..., cl.7 is shown in Table S3C.

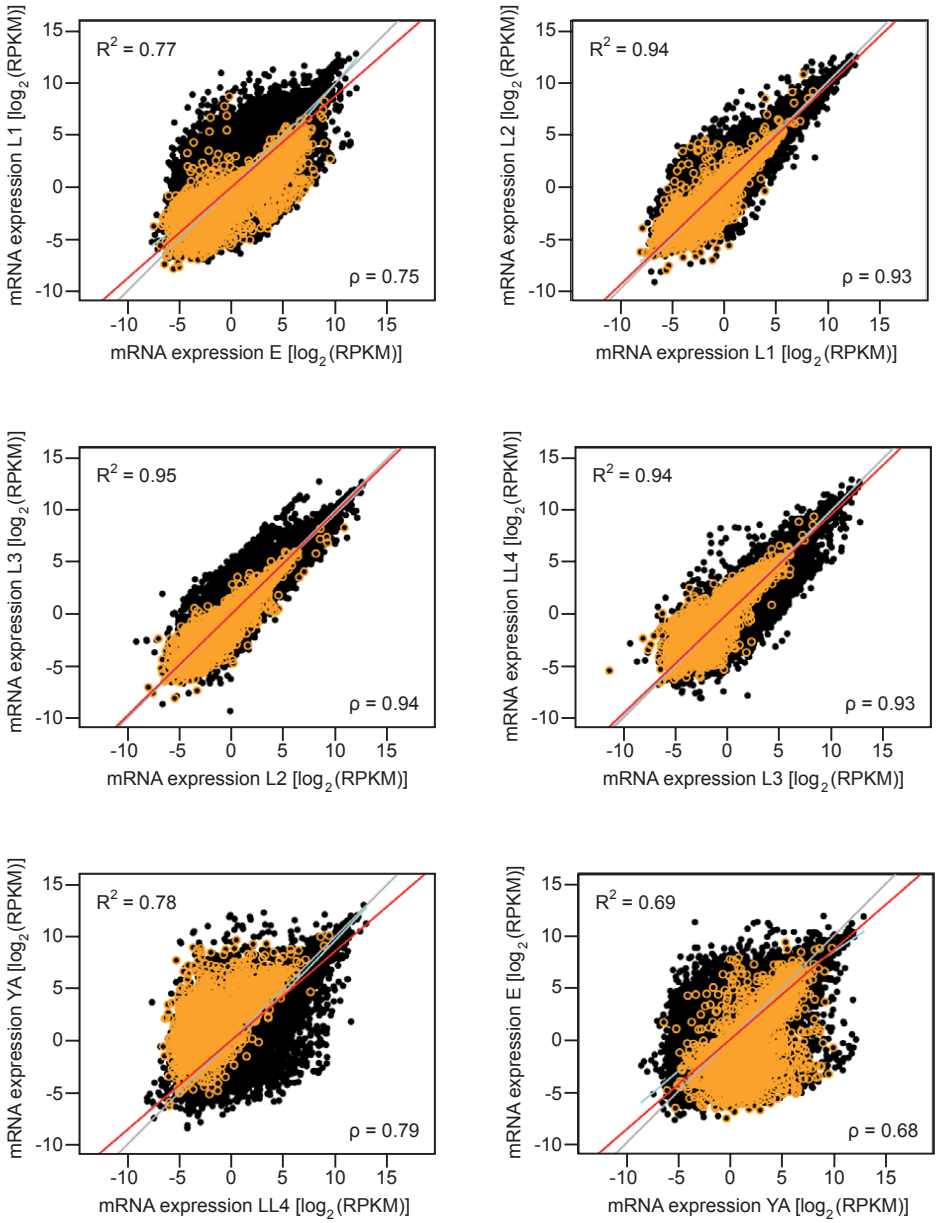

mRNA expression L1 [log2(RPKM)]
R2 = 0.77
ρ = 0.75
mRNA expression E [log2(RPKM)]
mRNA expression L2 [log2(RPKM)]
R2 = 0.94
ρ = 0.93
mRNA expression L1 [log2(RPKM)]
mRNA expression L3 [log2(RPKM)]
R2 = 0.95
ρ = 0.94
mRNA expression L2 [log2(RPKM)]
mRNA expression LL4 [log2(RPKM)]
R2 = 0.94
ρ = 0.93
mRNA expression L3 [log2(RPKM)]
mRNA expression YA [log2(RPKM)]
R2 = 0.78
ρ = 0.79
mRNA expression LL4 [log2(RPKM)]
mRNA expression E [log2(RPKM)]
R2 = 0.69
ρ = 0.68
mRNA expression YA [log2(RPKM)]
-10
-5
0
5
10
15
all genes
dynamically expressed genes

C. elegans

Figure S4, Related to Figure 5. Identification of genes with differential transcript expression across larval development. Genes with differential expression across larval stages were identified as a group of outliers at the L2-L3 transition. This group was separated from the bulk of points by requiring at 4-fold up-regulation at L3 versus L2. These genes are highlighted in orange in all scatter plots and display dynamic expression across development.

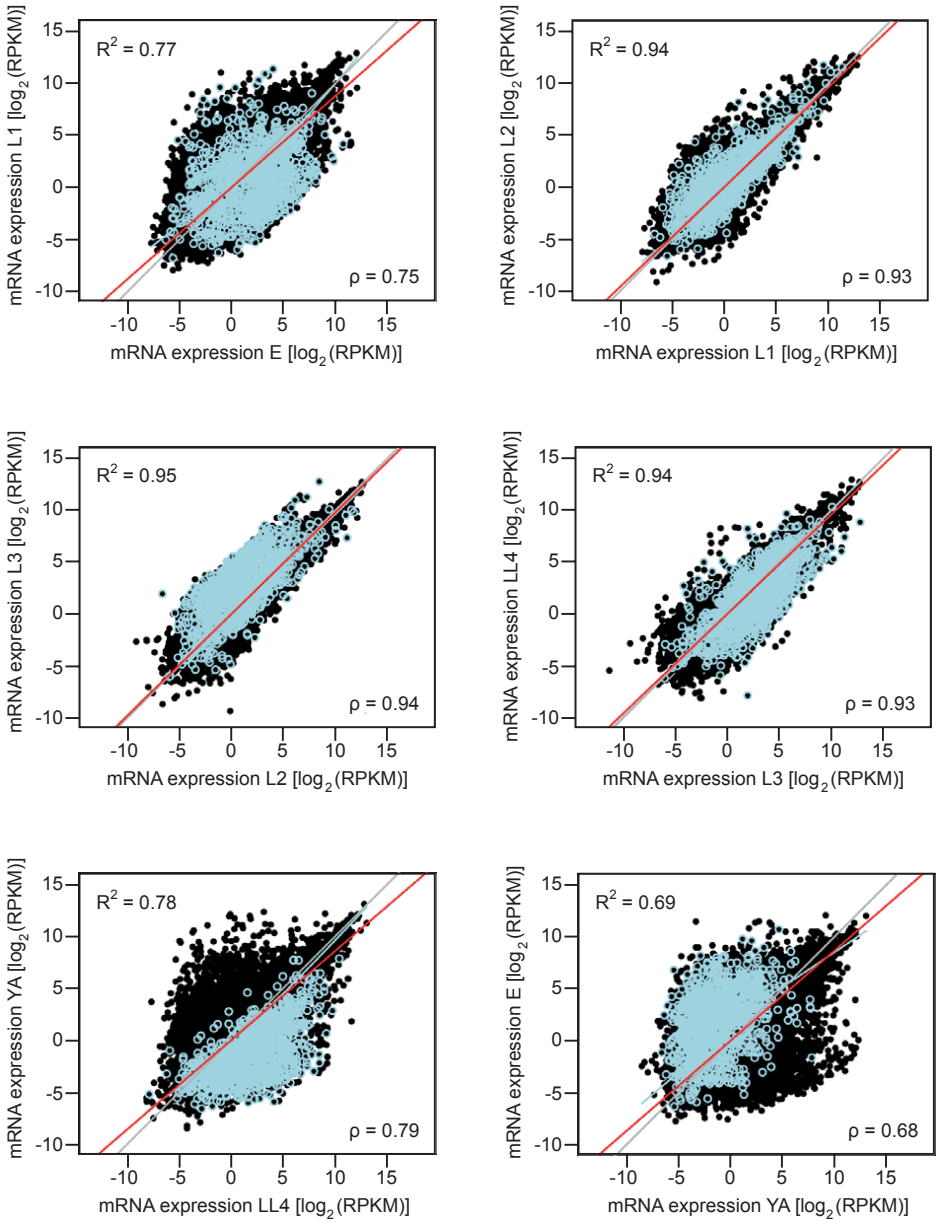

A
mRNAs (transitions)
protein (transitions)
C. elegans mRNA foldchange [log2FC(RPKM)]
C. elegans protein foldchange [log2FC(L/H)]
E/YA −0.59 −0.24 0.18 −0.04 0.06 0.12 0.46
YA/LL4 −0.21 −0.07 −0.07 0.16 0.19 0.59 −0.31
LL4/L4 −0.1 0.05 −0.06 0.28 0.22 0.41 −0.32
L4/L3 0.07 0.07 −0.02 −0.29 −0.05 0 −0.02
L3/L2 −0.13 −0.07 0.09 0.33 0.16 0.08 −0.06
L2/L1 −0.14 0.24 0.2 −0.12 0.04 −0.13 0.04
L1/E 0.62 0.14 −0.16 −0.08 −0.22 −0.47 −0.1
E/YA −0.44 −0.04 0.06 −0.11 0.05 −0.04 0.39
YA/LL4 −0.03 −0.17 −0.02 0.16 −0.08 0.24 −0.1
LL4/L4 −0.03 0.08 0.03 −0.12 0.06 −0.07 0.02
L4/L3 0.04 0.01 −0.01 0.04 0.02 0.03 −0.04
L3/L2 0.14 0.16 −0.01 −0.04 0.07 −0.1 −0.12
L2/L1 0.04 0.12 0.09 −0.11 0.04 −0.13 −0.01
L1/E 0.65 0.09 −0.07 0.02 −0.01 −0.12 −0.36
L1/E L2/L1 L3/L2 L4/L3 LL4/L4 YA/LL4 E/YA
C. briggsae mRNA foldchange [log2FC(RPKM)]
C. briggsae protein foldchange [log2FC(L/H)]
B
E → L1 transition
L2 → L3 transition
LL4 → YA transition
YA → E transition
mRNA
protein
mRNA (shuf.)
protein (shuf.)
Counts
Δlog2FC

C


mRNA expression L1 [log2(RPKM)]
mRNA expression E [log2(RPKM)]
R2 = 0.77
ρ = 0.75
mRNA expression L2 [log2(RPKM)]
mRNA expression L1 [log2(RPKM)]
R2 = 0.94
ρ = 0.93
mRNA expression L3 [log2(RPKM)]
mRNA expression L2 [log2(RPKM)]
R2 = 0.95
ρ = 0.94
mRNA expression LL4 [log2(RPKM)]
mRNA expression L3 [log2(RPKM)]
R2 = 0.94
ρ = 0.93
mRNA expression YA [log2(RPKM)]
mRNA expression LL4 [log2(RPKM)]
R2 = 0.78
ρ = 0.79
mRNA expression E [log2(RPKM)]
mRNA expression YA [log2(RPKM)]
R2 = 0.69
ρ = 0.68
all genes
germ line specific genes

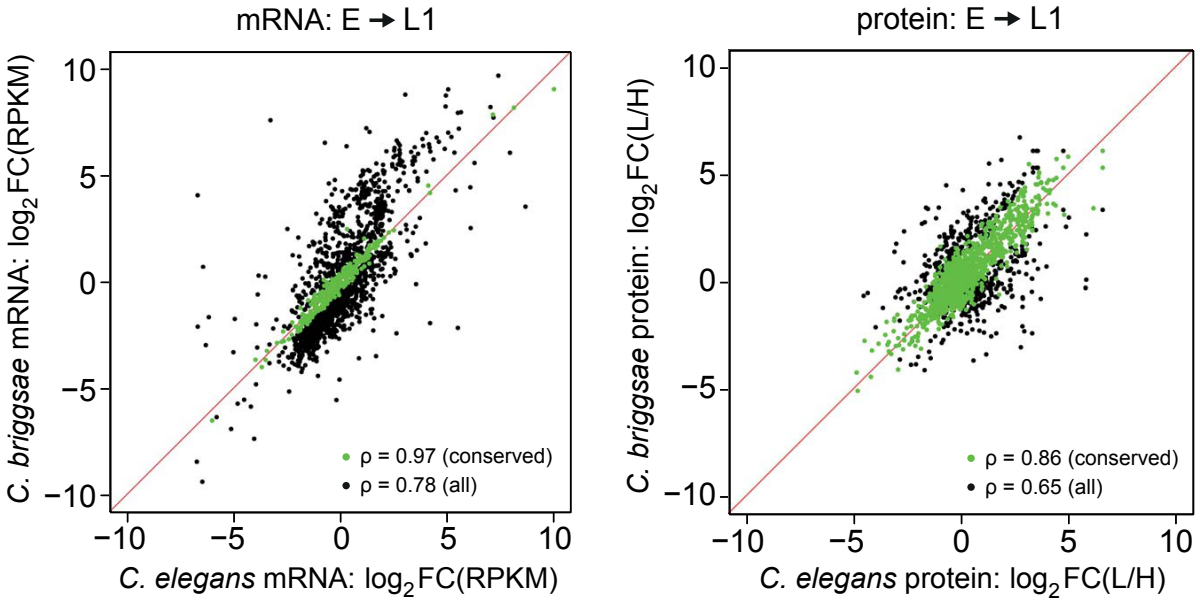
C. elegans

D

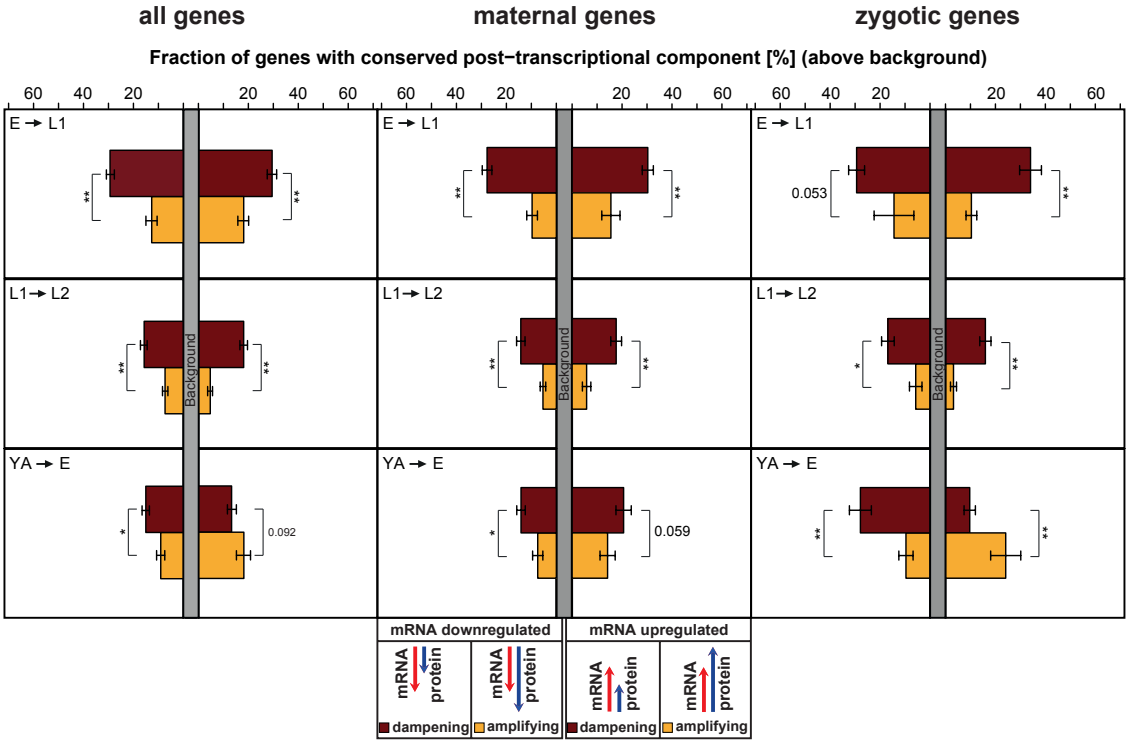

mRNA expression L1 [log2(RPKM)]
mRNA expression E [log2(RPKM)]
R2 = 0.77
ρ = 0.75
mRNA expression L2 [log2(RPKM)]
mRNA expression L1 [log2(RPKM)]
R2 = 0.94
ρ = 0.93
mRNA expression L3 [log2(RPKM)]
mRNA expression L2 [log2(RPKM)]
R2 = 0.95
ρ = 0.94
mRNA expression LL4 [log2(RPKM)]
mRNA expression L3 [log2(RPKM)]
R2 = 0.94
ρ = 0.93
mRNA expression YA [log2(RPKM)]
mRNA expression LL4 [log2(RPKM)]
R2 = 0.78
ρ = 0.79
mRNA expression E [log2(RPKM)]
mRNA expression YA [log2(RPKM)]
R2 = 0.69
ρ = 0.68
all genes
somatic genes
(downregulated in germ line)

C. elegans

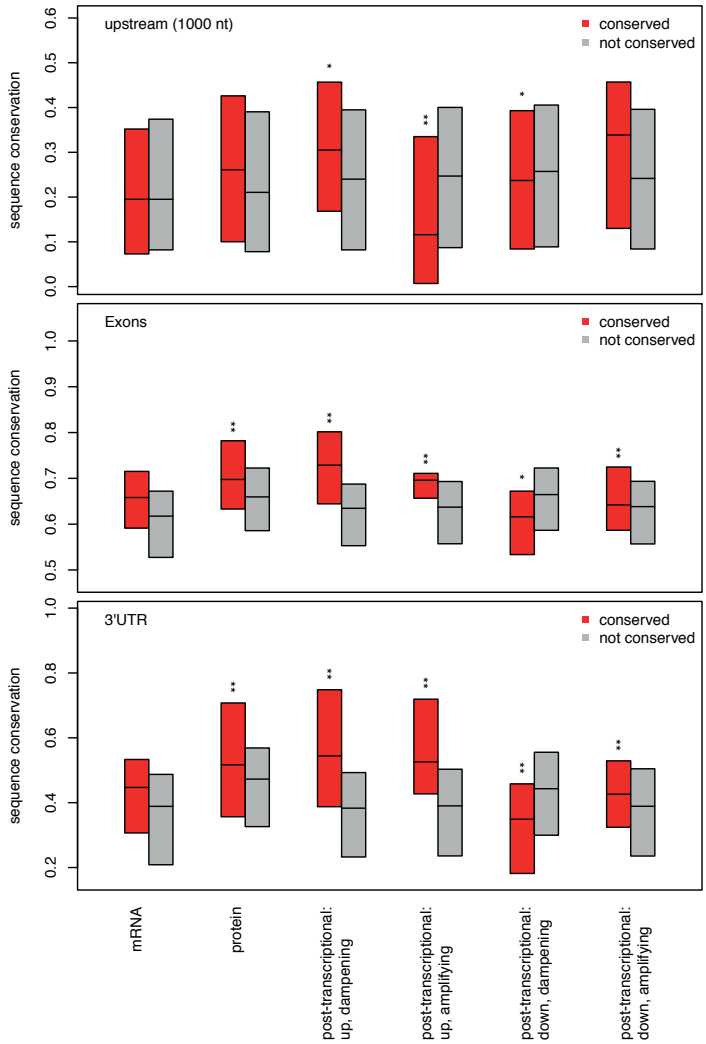
E


mRNA: E → L1
protein: E → L1
C. briggsae mRNA: log2FC(RPKM)
C. elegans mRNA: log2FC(RPKM)
C. briggsae protein: log2FC(L/H)
C. elegans protein: log2FC(L/H)
ρ = 0.97 (conserved)
ρ = 0.78 (all)
ρ = 0.86 (conserved)
ρ = 0.65 (all)

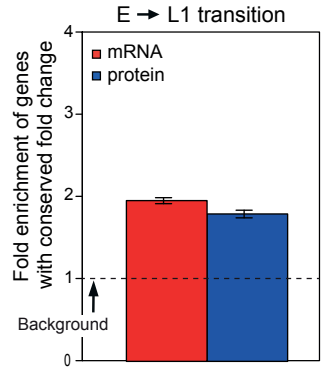
F


all genes
maternal genes
zygotic genes
Fraction of genes with conserved post−transcriptional component [%] (above background)
E → L1
L1→ L2
YA → E
Background
0.053
0.092
0.059
mRNA downregulated
mRNA upregulated
mRNA
protein
dampening
amplifying

G


upstream (1000 nt)
Exons
3'UTR
conserved
not conserved
sequence conservation
mRNA
protein
post-transcriptional: up, dampening
post-transcriptional: up, amplifying
post-transcriptional: down, dampening
post-transcriptional: down, amplifying

H


E → L1 transition
mRNA
protein
Fold enrichment of genes with conserved fold change
Background

I

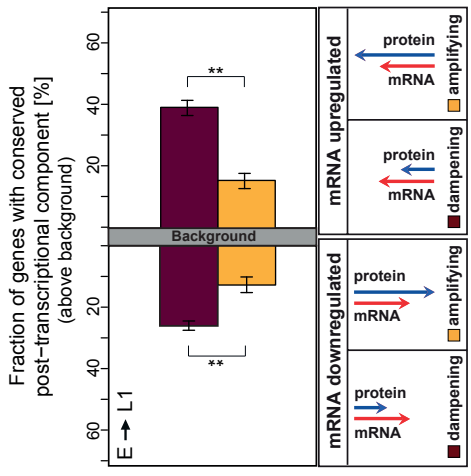
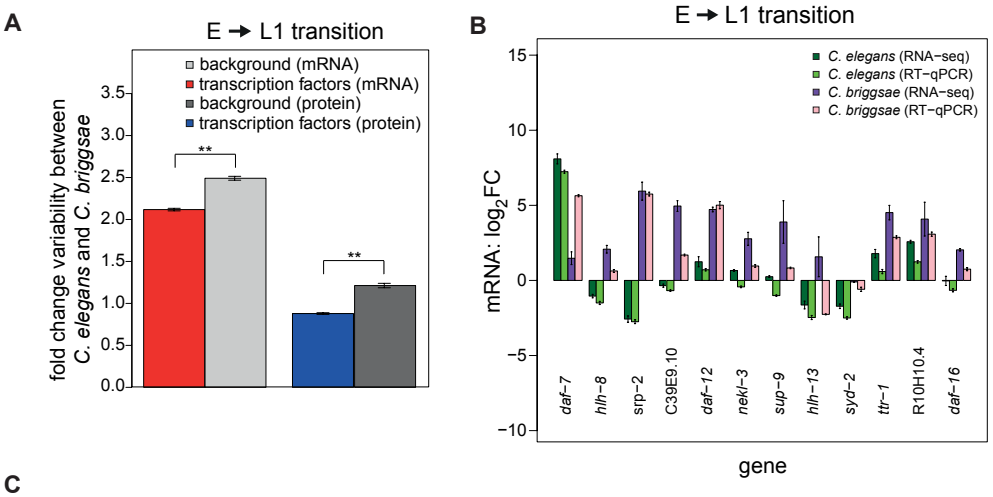
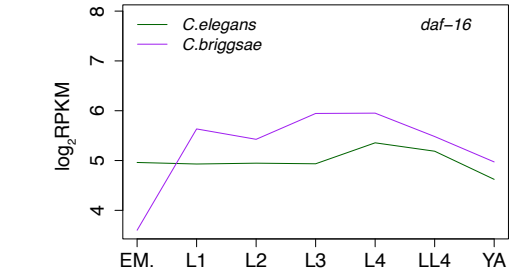

Fraction of genes with conserved post-transcriptional component [%] (above background)
E → L1
Background
mRNA upregulated
mRNA downregulated
protein
mRNA
amplifying
dampening

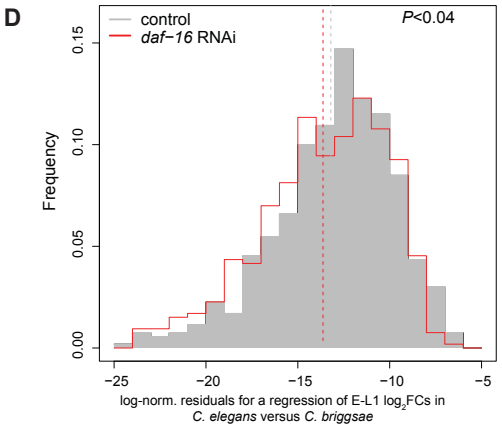
Figure S5, Related to Figure 6. Evolutionary comparison of developmental transcript and protein fold changes. (A) The correlation between log_2-fold changes in *C. elegans* and *C. briggsae* is shown for transcripts (left) and proteins (right). (B) To quantify conservation of transcript and protein expression, the inter-species difference in log_2-fold changes of transcripts (red) and proteins (blue) is plotted as a histogram and compared to a background distribution of log_2-fold changes computed after shuffling the orthology assignments of genes between the two species. Histograms are shown from left to right for the embryo-L1 and the L2-L3 transition in the upper panel, and from left to right for the LL4-young adult transition and the young adult-embryo transition in the lower panel. The highest level of conservation was observed at the embryo-L1 transition, only very limited conservation at the L2-L3 and intermediate conservation at the LL4-young adult transition. (C-D) Genes up- and down-regulated in the germline versus the somatic cell lineages were identified by comparing transcript expression in wild type and *glp-4* (bn2ts) mutant worms lacking the germline. In the scatter plots comparing transcript expression at subsequent developmental stages, germline specific genes are highlighted in orange (C) and somatic genes, i. e. genes down-regulated in the germline are marked in light-blue (D). Germline expressed genes overlap with genes up-regulated at the LL4-young adult transition and somatic genes correspond to genes down-regulated at this transition. The comparison further reveals that germline specific genes are up- and somatic genes are down-regulated in young adults versus embryos, indicating that the contribution of the germline to the transcriptome increases upon this developmental transition. (E) Log_2-fold changes at the embryo-L1 transition were averaged across replicates. Genes with overlapping error intervals derived from standard deviation of transcript or protein fold changes were considered conserved on transcript or protein level, respectively. The scatter plot compares fold changes of all genes (black) with conserved fold changes (green) for transcripts (left) and proteins (right). In either case, the correlation of conserved versus non-conserved genes increases substantially. Spearman's correlation coefficient (ρ) is shown in the lower right corner of the plots. (F) Fraction of *C. elegans* genes with a positive (negative) log_2-fold difference between protein and transcript fold changes (*post-transcriptional component*) that exhibit a positive (negative) log_2-fold difference in *C. briggsae*. This quantity reflects the conservation of the post-transcriptional component. Background conservation as

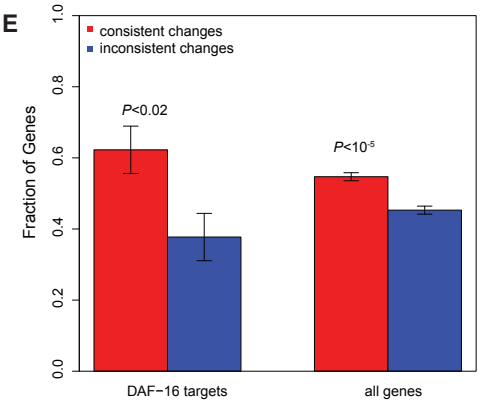
expected for unrelated species was obtained by shuffling orthology relations between genes in *C. elegans* and *C. briggsae* and has been subtracted. Conservation is separately shown for dampening and amplifying post-transcriptional changes in case of up- or down-regulation of transcripts at the embryo-L1, L1-L2 and the young adult-embryo transition. The ensemble of all genes (left) was split up into two groups of genes enriched in maternal (middle) and in zygotically transcribed genes (right) (see Supplemental Experimental Procedures). At the young adult-embryo transition, maternal and zygotic genes display complementary conservation patterns of the post-transcriptional component for up-regulated transcripts. While for maternal genes, as for all other transitions, dampening of protein fold changes is preferentially conserved, zygotic genes exhibit lower conservation of dampening, but, in contrast to all other transitions, enhanced conservation of post-transcriptional regulation that amplifies protein fold changes. Hence, upon zygotic activation post-transcriptional regulation that buffers transcript fold changes of freshly transcribed genes is only weakly conserved and is thus not as crucial as for maternal genes. Error bars are based on random counting statistics. Notably, this distinct conservation pattern for zygotic versus maternal genes is only observed for genes transcriptionally up-regulated at the young adult-embryo transition. (G) Conservation of regulatory sequence at the embryo-to-L1 transition. We measured sequence conservation from three-way alignments of *C. elegans*, *C.briggsae* and *C. remanei* downloaded from the UCSC genome browser (Dreszer et al., 2012). We extracted alignments for promoter sequence (1kb upstream), for 3'UTR based on a recently published annotation (Mangone et al., 2010) and for all exons. We only considered mRNAs and proteins, which change at least twofold at this transition in *C. elegans* and assessed conservation of up- or down-regulation in *C. briggsae.* We then compared sequence conservation between genes with conserved and non-conserved fold changes. We also included genes with a conserved post-transcriptional component in each of the four classes (up- or down-regulated mRNA with dampening or amplifying regulation) in this comparison. The three panels display conservation of promoter sequence, all exons, and of the 3'UTR. Boxes represent interquartile range and the median is indicated as black line. Sequence conservation is shown for genes with conserved (red) and non-conserved (grey) fold changes. (H, I) Reproducibility of fold change conservation measured with recently published *C.*

briggsae gene models (Uyar et al., 2012): 13,355 one-to-one orthologs to *C. elegans* genes were identified by reciprocal best protein blast hits. Quantified fold changes for these genes in *C. briggsae* were compared to *C. elegans* and evolutionary conservation was highly similar to measurements based on our own gene models. (H) Comparison of mRNA and protein fold changes at the Embryo-L1 transition as in Figure 6D. (I) Evolutionary conservation of the post-transcriptional component as in Figure 6E. (*P<0.05, **P<0.001)


A
E → L1 transition
background (mRNA)
transcription factors (mRNA)
background (protein)
transcription factors (protein)
fold change variability between
C. elegans and C. briggsae
B
E → L1 transition
C. elegans (RNA−seq)
C. elegans (RT−qPCR)
C. briggsae (RNA−seq)
C. briggsae (RT−qPCR)
mRNA: log2FC
daf−7
hlh−8
srp−2
C39E9.10
daf−12
nekl−3
sup−9
hlh−13
syd−2
ttr−1
R10H10.4
daf−16
gene


C
C.elegans
C.briggsae
daf−16
log2RPKM
EM.
L1
L2
L3
L4
LL4
YA


D
control
daf−16 RNAi
P<0.04
Frequency
log-norm. residuals for a regression of E-L1 log2FCs in
C. elegans versus C. briggsae


E
consistent changes
inconsistent changes
P<0.02
P<10−5
Fraction of Genes
DAF−16 targets
all genes

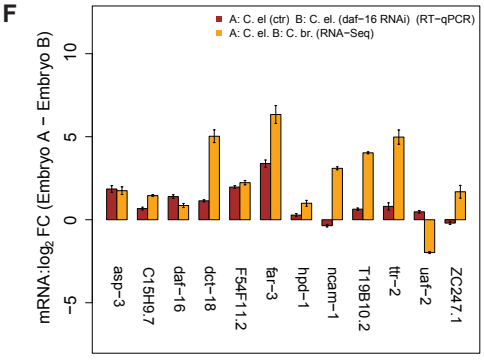

F
A: C. el (ctr) B: C. el. (daf−16 RNAi) (RT−qPCR)
A: C. el. B: C. br. (RNA−Seq)
mRNA:log2 FC (Embryo A − Embryo B)
asp−3
C15H9.7
daf−16
dct−18
F54F11.2
far−3
hpd−1
ncam−1
T19B10.2
ttr−2
uaf−2
ZC247.1

Figure S6, Related to Figure 7. Identification of genes with increased or reduced conservation of embryo-L1 fold changes. (A) Cross-species variability of embryo-L1 transcript and protein fold changes for transcription factors and non-transcription factor genes. Cross-species variability of fold changes is quantified by the standard deviation of the $\log_2$-ratio of fold changes in *C. elegans* and *C. briggsae.* Error bars indicate the standard error of the mean and were inferred by bootstrapping**.** (B) Validation of RNA-seq based quantification (measured in $\log_2$RPKM by RT-qPCR (measured in $-\Delta C_t$) for 12 genes with transcripts taken from four sets of differentially evolved genes in *C. elegans*. Samples measured by RT-qPCR were individual biological replicate samples from samples measured by RNA-seq. These genes are switched on or off in *C. elegans* and vice versa in *C. briggsae* either in embryos or L1-larvae. We required a minimum (maximum) expression of 4 (0.5) RPKM to consider a gene switched on or off, respectively. Four genes (*syd-2, ttr-1, daf-7, daf-12*) are known targets of DAF-16 (same as in Figure 7A). Thus, the *daf-16* mRNA level is also shown although expression differences in embryo and larvae are less pronounced than for the other genes, which obey the selection criterion. The plot displays the comparison of transcript fold changes at the embryo-L1 transition derived from $\log_2$RPKM and $-\Delta C_t$ at the embryonic and the L1 stage. Differential regulation is supported by RNA-seq and qPCR in all cases, except for *hlh-13*. Error bars represent the standard deviation across replicates ($n \geq 3$). (C) Expression of *daf-16* mRNA during *C. elegans* and *C. briggsae* development. In contrast to homogenous expression throughout *C. elegans* development, *daf-16* expression is low in *C. briggsae* embryos and strongly up-regulated at the first larval stage. Expression was aggregated across all isoforms. (D-F) Knockdown of RNAi reduces differences of embryo-L1 protein fold changes between *C. elegans* and *C. briggsae.* An independent biological replicate was performed to reproduce the observation presented in Fig. 7C-F (D) Distribution of the logarithm of the squared residuals normalized by the total sum of squares from a regression of embryo-L1 $\log_2$-fold changes between *C. elegans* and *C. briggsae*. Upon knockdown of *daf-16* in *C. elegans* embryos the distribution (red line) is shifted compared to a control transfection (grey) (Wilcoxon rank sum test $P < 0.04$). (E) Genes with differences in embryo-L1 $\log_2$-fold changes between *daf-16* knockdown and control *C. elegans* embryos that exhibit consistent (red) or inconsistent (blue) differences between *C. elegans* and *C. briggsae*. Data are shown for

all genes and for the subset of predicted *daf-16* targets (Murphy et al., 2003). A hypergeometric test was performed to assess an enrichment of consistent fold changes. Error bars are based on random counting statistics. (F) Log_2-fold changes of embryonic expression for *daf-16* and 11 targets between *daf-16* RNAi and control samples (as determined by RT-qPCR) and between *C. elegans* and *C. briggsae* (as determined by RNA-seq) are shown. Error bars represent the standard deviation across replicates ($n \geq 3$).

Supplemental Tables

A

Sample	Replicate	Caenorhabditis species	Number of reads	Reads mapped in pairs [%]	Reads mapped unpaired [%]	Total reads mapped [%]
Embryo	A1	*C. elegans*	27,076,598	77.5	6.7	84.1
Embryo	A2	*C. elegans*	92,542,536	77.9	3.7	81.6
Embryo	B	*C. elegans*	102,404,346	78.2	4.3	82.5
L1	A1	*C. elegans*	27,094,012	69.9	6.5	76.4
L1	A2	*C. elegans*	92,426,108	75.9	3.8	79.7
L1	B	*C. elegans*	96,203,398	75.7	3.9	79.5
L2	A	*C. elegans*	27,078,348	72.4	6.7	79.2
L3	A	*C. elegans*	28,404,368	71.8	6.4	78.2
L4	A	*C. elegans*	32,439,676	60.0	8.6	68.6
LL4	A	*C. elegans*	32,096,248	69.5	7.4	76.9
YA	A	*C. elegans*	27,520,378	87.9	3.4	91.2
Ref	A1	*C. elegans*	26,026,042	78.3	6.3	84.6
Ref	A2	*C. elegans*	30,625,738	86.9	3.4	90.2
Embryo	A1	*C. briggsae*	29,968,702	50.3	11.3	61.6
Embryo	A2	*C. briggsae*	15,515,744	82.6	2.5	85.1
Embryo	B	*C. briggsae*	4,211,172	80.2	2.4	82.6
L1	A1	*C. briggsae*	40,040,886	49.3	10.4	59.7
L1	A2	*C. briggsae*	20,240,052	84.6	3.0	87.5
L1	B	*C. briggsae*	25,439,560	73.8	2.7	76.6
L2	A	*C. briggsae*	41,850,214	38.5	12.4	50.9
L3	A	*C. briggsae*	47,467,568	46.1	12.1	58.2
L4	A	*C. briggsae*	42,920,508	58.7	13.0	71.6
LL4	A	*C. briggsae*	40,790,594	58.7	13.1	71.8
YA	A	*C. briggsae*	33,204,198	83.1	2.7	85.8
Ref	A1	*C. briggsae*	39,483,552	58.0	12.7	70.7
Ref	A2	*C. briggsae*	30,740,032	83.9	2.8	86.6

Table S1, Related to Figure 3. Read statistics and mRNA and protein expression data. (A) Sample A was sequenced for all stages and the reference with technical replicates for embryos, L1 larvae and the reference (A1, A2). For embryos and L1 larvae biological replicates (B) were sequenced.

A

Cluster	Number of genes	Mean coefficient of concordance: mRNA	Mean coefficient of concordance: protein	RNAi: phenotype [%]	RNAi: lethal phenotype [%]
cl.1	239	0.67**	0.33**	28	16
cl.2	497	0.41	0.35**	41	26
cl.3	240	0.59**	0.13	44	25
cl.4	269	0.74**	0.23	60**	48**
cl.5	237	0.31	0.06	56*	43**
cl.6	186	0.53**	0.05	61**	49**
cl.7	293	0.51	0.16	56*	34*

B

Cluster	Number of genes	RNAi: phenotype [%]	RNAi: lethal phenotype [%]
ccl.1	478	37	20
ccl.2	239	46	27
ccl.3	382	59**	46**
ccl.4	179	53	38
ccl.5	340	43	29
ccl.6	141	51	38
ccl.7	200	58**	46**

C

CLUSTER	ccl.1	ccl.2	ccl.3
cl.1	478/239/148/Inf/61.9%	239/239/15/0/6.3%	382/239/4/0/1.7%
cl.2	478/497/203/Inf/42.5%	239/497/53/0.04/22.2%	382/497/49/0/12.8%
cl.3	478/240/22/0/9.2%	239/240/24/0.06/10%	382/240/74/5.27/30.8%
cl.4	478/269/9/0/3.3%	239/269/18/0/7.5%	382/269/148/Inf/55%
cl.5	478/237/22/0/9.3%	239/237/31/0.48/13.1%	382/237/52/0.75/21.9%
cl.6	478/186/42/0.12/22.6%	239/186/22/0.26/11.8%	382/186/11/0/5.9%
cl.7	478/293/44/0/15%	239/293/78/13.05/32.6%	382/293/53/0.11/18.1%
CLUSTER	ccl.4	ccl.5	ccl.6
cl.1	179/239/4/0/2.2%	340/239/25/0/10.5%	141/239/4/0/2.8%
cl.2	179/497/28/0/15.6%	340/497/106/2.17/31.2%	141/497/31/0.07/22%
cl.3	179/240/46/6.99/25.7%	340/240/57/2.36/23.8%	141/240/18/0.4/12.8%
cl.4	179/269/44/4.65/24.6%	340/269/27/0/10%	141/269/21/0.49/14.9%
cl.5	179/237/18/0.09/10.1%	340/237/59/2.96/24.9%	141/237/29/2.88/20.6%
cl.6	179/186/6/0/3.4%	340/186/32/0.27/17.2%	141/186/7/0.01/5%
cl.7	179/293/33/1.03/18.4%	340/293/50/0.23/17.1%	141/293/32/2.26/22.7%
CLUSTER	ccl.7		
cl.1	200/239/46/5.49/23%		
cl.2	200/497/39/0.01/19.5%		
cl.3	200/240/7/0/3.5%		
cl.4	200/269/5/0/2.5%		
cl.5	200/237/29/0.85/14.5%		
cl.6	200/186/70/Inf/37.6%		
cl.7	200/293/13/0/6.5%		

D

ID	cl.1	cl.2	cl.3	cl.4	cl.5	cl.6	cl.7
Phenotype	66/28%/0	205/41%/0	105/44%/0	162/60%/5.1	133/56%/2.4	113/61%/3.9	163/56%/2.6
Lethal	38/16%/0	128/26%/0	61/25%/0	130/48%/8	102/43%/3.5	91/49%/6	101/34%/0.5
embryonic lethal	35/15%/0	117/24%/0	58/24%/0	127/47%/8.8	97/41%/3.4	89/48%/6.5	93/32%/0.4
slow growth	23/10%/0	70/14%/0	44/18%/0.9	42/16%/0.3	51/22%/2.2	29/16%/0.3	57/19%/1.5
larval arrest	16/7%/0	52/10%/0	31/13%/0.2	33/12%/0.1	51/22%/3.7	51/27%/7.1	39/13%/0.2
maternal sterile	10/4%/0	42/8%/0	26/11%/0.1	49/18%/2.3	42/18%/1.8	56/30%/10.3	39/13%/0.3
locomotion abnormal	5/2%/0	48/10%/2.2	9/4%/0	20/7%/0.4	12/5%/0	6/3%/0	41/14%/5.7
larval lethal	8/3%/0	34/7%/1.2	14/6%/0.5	15/6%/0.4	11/5%/0.2	5/3%/0	22/8%/1.4
organism morphology abnormal	4/2%/0	23/5%/1	5/2%/0	14/5%/1.2	5/2%/0	6/3%/0.3	18/6%/2
unclassified	4/2%/0	16/3%/0.3	7/3%/0.2	18/7%/2.9	5/2%/0.1	3/2%/0	15/5%/1.5
sterile progeny	1/0%/0	13/3%/0.2	10/4%/1	13/5%/1.5	10/4%/1	2/1%/0	12/4%/1
protruding vulva	2/1%/0	13/3%/0.7	5/2%/0.3	16/6%/4.6	3/1%/0.1	0/0%/0	7/2%/0.5
aldicarb resistant	1/0%/0	10/2%/0.9	4/2%/0.5	2/1%/0.1	1/0%/0	3/2%/0.5	9/3%/2
shortened life span	1/0%/0.1	9/2%/0.8	2/1%/0.2	7/3%/1.5	0/0%/0	1/1%/0.1	9/3%/2.2
reduced brood size	1/0%/0.1	6/1%/0.4	7/3%/2.3	3/1%/0.4	3/1%/0.5	0/0%/0	3/1%/0.3
Dumpy	3/1%/0.5	5/1%/0.3	0/0%/0	1/0%/0.1	2/1%/0.3	1/1%/0.2	10/3%/3.9
Sterile	10/4%/0	47/9%/0	27/11%/0	53/20%/2.7	42/18%/1.5	57/31%/9.8	42/14%/0.4
molt defect	0/0%/0	3/1%/0.1	1/0%/0.1	4/1%/0.9	2/1%/0.4	1/1%/0.2	8/3%/3
fat content reduced	1/0%/0.2	8/2%/1.5	4/2%/1.1	1/0%/0.1	0/0%/0	1/1%/0.2	5/2%/1.2
extended life span	5/2%/2	5/1%/0.7	1/0%/0.2	0/0%/0	2/1%/0.5	4/2%/1.8	1/0%/0.1
mitotic spindle abnormal early embryogenesis	0/0%/0.1	1/0%/0.1	1/0%/0.3	5/2%/2.5	1/0%/0.3	0/0%/0.2	4/1%/1.6
exploded through vulva	2/1%/0.7	3/1%/0.4	0/0%/0.1	2/1%/0.6	0/0%/0.1	0/0%/0.2	6/2%/3.1
egg laying abnormal	0/0%/0.1	5/1%/1.2	1/0%/0.3	1/0%/0.3	0/0%/0.1	0/0%/0.2	5/2%/2.3

Table S3, Related to Figure 4. Expression clusters in *C. elegans* and *C. briggsae*. (A) The table indicates the number of genes in each of the seven *C. elegans* clusters (Figure S3C), the mean coefficient of concordance between *C. elegans* and *C. briggsae* for transcript and protein expression profiles, and the percentages of genes with phenotypes and lethal phenotypes. (B) Statistics for the seven *C. briggsae* clusters (Figure S3D). The table indicates the number of genes in each cluster and the fraction of genes with phenotypes and lethal phenotypes. (*$P < 0.01$, ** $P < 0.001$). (C) Overlap between the seven expression clusters in *C. elegans* (cl.1, ..., cl.7) and the *C. elegans*

orthologs of genes in each of the seven *C. briggsae* expression clusters (ccl.1, ..., ccl.7). An entry at row i and column j contains the following numbers: number of genes in ccl.i/number of genes in cl.j/ number of genes in cl.i and ccl.j/ $-\log_{10}(P)$ for significantly enhanced overlap of cl.i and ccl.j/ percentage of overlap based on the number of genes in the smaller cluster of cl.i and ccl.j. Entries with significant overlap are highlighted in orange. (D) The table summarizes frequency and significance of enrichment of the 21 most abundant annotated RNAi phenotypes for each cluster. Each entry comprises the number of genes with the respective phenotype in each cluster, the percentage relative to the number of genes in the cluster, and the $-\log_{10}(P)$ of enrichment.

Supplemental References

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics 77, 71–94.

Breslauer, K.J., Frank, R., Blöcker, H., and Marky, L.A. (1986). Predicting DNA duplex stability from the base sequence. Proceedings of the National Academy of Sciences of the United States of America 83, 3746–3750.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology 26, 1367–1372.

Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V, and Mann, M. (2009). A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. Nature Protocols 4, 698–705.

Dieckmann-Schuppert, A., and Schnittler, H.J. (1997). A simple assay for quantification of protein in tissue sections, cell cultures, and cell homogenates, and of protein immobilized on solid surfaces. Cell and Tissue Research 288, 119–126.

Dreszer, T.R., Karolchik, D., Zweig, A.S., Hinrichs, A.S., Raney, B.J., Kuhn, R.M., Meyer, L.R., Wong, M., Sloan, C.A., Rosenbloom, K.R., et al. (2012). The UCSC Genome Browser database: extensions and updates 2011. Nucleic Acids Research 40, D918–23.

Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nature Methods 4, 207–214.

Gerstein, M.B., Lu, Z.J., Van Nostrand, E.L., Cheng, C., Arshinoff, B.I., Liu, T., Yip, K.Y., Robilotto, R., Rechtsteiner, A., Ikegami, K., et al. (2010). Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. Science (New York, N.Y.) 330, 1775–1787.

Gotoh, O. (2008). Direct mapping and alignment of protein sequences onto genomic sequence. Bioinformatics (Oxford, England) 24, 2438–2444.

Hanke, S., Besir, H., Oesterhelt, D., and Mann, M. (2008). Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. Journal of Proteome Research 7, 1118–1130.

Huang, D.W., Sherman, B.T., Zheng, X., Yang, J., Imamichi, T., Stephens, R., and Lempicki, R.A. (2009). Extracting biological meaning from large gene lists with DAVID. Current Protocols in Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [et Al.] *Chapter 13*, Unit 13.11.

Jungkamp, A.-C., Stoeckius, M., Mecenas, D., Grün, D., Mastrobuoni, G., Kempa, S., and Rajewsky, N. (2011). In vivo and transcriptome-wide identification of RNA binding protein target sites. Molecular Cell 44, 828–840.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Research 12, 996–1006.

Lewis, J.A., and Fleming, J.T. (1995). Basic culture methods. Methods in Cell Biology 48, 3–29.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England) 26, 589–595.

Mangone, M., Manoharan, A.P., Thierry-Mieg, D., Thierry-Mieg, J., Han, T., Mackowiak, S.D., Mis, E., Zegar, C., Gutwein, M.R., Khivansara, V., et al. (2010). The landscape of C. elegans 3'UTRs. Science (New York, N.Y.) 329, 432–435.

Murphy, C.T., McCarroll, S.A., Bargmann, C.I., Fraser, A., Kamath, R.S., Ahringer, J., Li, H., and Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. Nature 424, 277–283.

Pepke, S., Wold, B., and Mortazavi, A. (2009). Computation for ChIP-seq and RNA-seq studies. Nature Methods 6, S22–32.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Analytical Chemistry 75, 663–670.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nature Protocols 3, 1101–1108.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V, and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nature Protocols 1, 2856–2860.

Sulston, J.E., and Brenner, S. (1974). The DNA of Caenorhabditis elegans. Genetics 77, 95–104.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics (Oxford, England) 25, 1105–1111.

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

Uyar, B., Chu, J.S.C., Vergara, I.A., Chua, S.Y., Jones, M.R., Wong, T., Baillie, D.L., and Chen, N. (2012). RNA-seq analysis of the C. briggsae transcriptome. Genome Research 22, 1567–1580.