Transcript accumulation rates in the early *C. elegans* embryo

Priya Sivaramakrishnan[1], Cameron Watkins[1] and John Isaac Murray[1,2]

1. Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104
2. Corresponding author - jmurr@pennmedicine.upenn.edu

ABSTRACT

Dynamic changes in transcription are widespread in the developing embryo, where cell cycles are rapid and cell fate decisions sometimes need to be made quickly, before the next cell division. In the early *Caenorhabditis elegans* embryo, specification of the intestine relies on high absolute levels of transcription factors that are a part of the gut gene regulatory network. These absolute levels are likely achieved by controlled transcript accumulation rates. However, accumulation rates have not been measured globally in the worm embryo. We used single cell RNA-seq data from the early *C. elegans* embryo to estimate the accumulation rates of zygotic genes up to the 24-cell stage. We find that rapid transcript accumulation is a characteristic feature of transcription factors across different cell types and lineages. We identified genomic features associated with high transcription rates and core promoter motifs that might drive these rates. For one Very High-rate gene *ceh-51*, which is required for mesoderm development, we measured the contributions of core promoter elements to rate. We find that each of these motifs contribute modestly to the accumulation rate of *ceh-51*, suggesting a complex relationship between promoter motifs and gene structure in controlling transcript accumulation rates. These results are a step towards understanding the regulation of transcript accumulation rates during embryonic cell fate specification.

Introduction

The developmental transcriptome is highly dynamic. Rapid gene expression changes are needed to produce the right amounts of regulators at the right time in order to specify a cell's fate. Advancements in single cell RNA sequencing (sc-RNA seq) have generated atlases that characterize the transcriptomes of individual cells during development[1-4]. Coupled with metabolic labeling of nascent transcripts, sc-RNA seq has the potential to capture transcription dynamics at high temporal resolution[5-7]. However, these methods are less optimized to quantify absolute transcript counts, which requires sequencing at very high depth.

Single-cell, single-molecule transcript imaging methods are the gold-standard for measuring the absolute number of mRNA molecules[8,9]. Single molecule RNA fluorescence *in situ* hybridization (smFISH) is performed on fixed samples, where multiple oligonucleotide probes bind to the target transcripts. Single transcripts are visualized as diffraction limited spots, enabling accurate counting of RNAs[8,10]. Studies have shown modest correlations between reads from different sc-RNA seq platforms and smFISH transcript counts, indicating the promise of using sc-RNA seq for genome-wide estimates of absolute transcript levels[11-13].

smFISH in fixed embryos have highlighted the importance of precise transcript levels for developmental robustness[14-16]. In rapidly dividing embryonic cells, transcripts also have to accumulate rapidly to produce the right final levels of mRNA required for cell fate decisions. Absolute transcript accumulation rates (change in transcript levels between mother and daughter cells over the time between cell divisions), are rarely explicitly measured during development. In *Drosophila* embryos, accumulation rates of four gap genes *hunchback*, *Krüppel*, *knirps* and *giant* were estimated to be fairly similar (33 transcripts/minute per nucleus) during the 15 minute nuclear cell cycle 13 (nc13)[14]. At this time, the fly embryo is still developing as a syncytium with nuclei sharing common cytoplasm, and accumulation rates could be different for genes expressed after cellularization. Genome-wide bulk RNA seq measurements of transcript accumulation on broad time scale (multiple rounds of cell division) in the *Xenopus* embryo similarly showed that most genes, including transcription factors, had uniform accumulation rates (transcripts/hour)[17]. It is unclear if the property of most zygotic genes having approximately equal accumulation rates will extend to other species. Furthermore, it is unknown whether the accumulation rates of all zygotic genes are regulated by common factors and mechanisms.

The *Caenorhabditis elegans* embryo provides an ideal developmental system to answer these questions. The invariant lineage of the *C. elegans* embryo could reflect a high level of precision in expression levels of cell fate specification factors[18]. Absolute transcript measurements of intestinal transcription factors are consistent with this[15]. Combined smFISH and embryo staging by automated lineage analysis further showed that the intestinal specification transcription factors *end-3* and *end-1* reproducibly accumulate to very high maximum transcript levels[19]. The final transcript levels of these genes in developing gut cells appears to be important, as sub-threshold levels of *end-1/3* lead to failure of intestinal specification[15,20].

To determine if high transcript accumulation rates are common for early *C. elegans* zygotic genes, and to characterize rates genome-wide, we inferred absolute transcript changes from early embryonic sc-RNA seq data. We found that transcript accumulation rates vary between genes across cell types and lineages. We identified common gene structural features associated with rapidly transcribed genes, such as short gene length, shorter and fewer introns and reduced *trans*-splicing. We also found motifs enriched at the promoters of High-rate genes including the Initiator (INR) motif. Using an exemplar Very High-rate gene, *ceh-51*, we tested the role of these core promoter motifs in regulating rapid transcription. Our results show that rates are regulated at multiple levels, consistent with the robustness of transcript abundance and overall development of the *C. elegans* embryo.

Results

Analysis of sc-RNA seq data reveals differences in transcript accumulation rates in the early *C. elegans* embryo

To examine transcription accumulation rates beyond the E (endodermal/intestinal) lineage, we analyzed an existing sc-RNA seq dataset. The dataset includes the transcriptomes from cells up to the 24-cell stage of the *C. elegans* embryo[21] (Fig. 1A). Since the timing of division varies between lineages, the 16-cell stage in the Tintori dataset reflects the 16 cells of the AB lineage or the 24-cell stage of the embryo (Fig 1A). Transcripts per million (TPM) counts that are usually used for RNA seq analysis do not provide a robust measurement of the actual RNA counts in a cell because total RNA levels scale with cell size[22,23]. We converted TPM counts to estimates of absolute transcript numbers by correcting for cell volume and adjusted for expected total transcripts per cell (Methods). We calculated each gene's absolute change as the change in the

absolute number of transcripts in each daughter compared to the mother (Fig. 1A, Methods). We also calculated the fold change for each gene in the daughter compared to the parent cell. By combining both absolute and fold change metrics, we removed genes fluctuating around high levels that could appear as a large absolute change due to technical noise, allowing us to predict genes with high levels of *de novo* transcript accumulation.

In order to identify common features associated with different levels of accumulation rates and study rate regulation, we first categorized zygotic genes into bins based on the distribution of absolute and fold change across all genes (Fig. 1B). We classified genes having >200 estimated new transcripts as “Very High”, 50-100 new transcripts as “High”, and 1-50 new transcripts as “Low”) (Fig 1B and Methods). We confirmed that at the expected combination of absolute and fold changes, the number of genes expressed in germline cells, which should be transcriptionally silent, was low (Supplementary Fig. 1A). The majority of zygotically expressed genes showed low accumulation rate across all stages tested (Fig. 1C, Supplementary Fig. 1D). Zygotic genome activation in *C. elegans* occurs earlier than other invertebrate model systems, at the 4-cell stage[24,25]. Consistent with the timing of genome activation, most Very High and High-rate genes are expressed at the 8- and 16-cell stages (Fig. 1C,D). At these stages, most cells except the germline (P lineage) and the D lineage show dynamic changes (Fig. 1D, Supplementary Fig. 1C). Germline transcription is not activated until after the primordial germ cells Z2 and Z3 are born close to the 100-cell stage, which is not represented in this dataset[26,27]. Similarly, much of the transcription in the D blastomere, which exclusively produces body wall muscle, begins a little after the 24-cell stage and hence too late to be measured in this single-cell dataset. Overall, the AB and E lineages have the most number of genes with high absolute changes, followed by the MS and C founders (Fig. 1D and Supplementary Fig. 1C). As expected, Very High and High-rate genes clustered based on embryo stage (Fig. 1E), except the germline cells P1-P3, which clustered together (Fig. 1E).

Comparison of inferred accumulation rates with single molecule FISH

sc-RNA seq suffers from reduced efficiency in capturing the entire transcriptome of a cell and from estimating different amounts of RNA between cells[11]. To test the accuracy of our strategy used to identify High-rate genes, we used smFISH, the gold standard method for measuring absolute mRNA numbers[8]. We targeted smFISH probes to exons to estimate levels of mature transcripts in the embryo[15,19]. All six candidate High-rate genes we tested showed corresponding

high transcript counts by smFISH (Fig. 2A-C). Comparing the smFISH counts to adjusted absolute RNA counts from sc-RNA seq shows a significant positive correlation (R = 0.7, p = 0.0039) (Fig. 2B). The level of correlation is similar to previous comparisons between different scRNA seq platforms and smFISH counts in melanoma cell lines[11].

Having validated high expression levels of High-rate genes by smFISH, we examined their rates further. We focused on genes involved in the specification of the MS (muscle) and E (intestine) as cell fate commitment occurs early in these cell types. We counted the total number of transcripts at each embryo stage and calculated the change in transcript number in the maximally expressing cell for each gene compared to the previous cell stage. This typically reflects the change between two cell divisions and represents the time where the gene is needed for its function in fate specification.

tbx-35 is a muscle specification factor, which begins to be expressed in the MS cell[28]. Between the birth and the division of the MS cell (17 min), ~300 transcripts of *tbx-35* are produced, giving an accumulation rate of 20 transcripts/minute (Fig. 2C). This is comparable with *end-3*, an intestinal specification gene expressed around the same time as *tbx-35* but in the E cell (23 transcripts/minute, Fig. 2D). *ceh-51*, which expressed one cell cycle later in the MS2 cells and is a target of TBX-35, accumulated at an even higher rate of 30 transcripts/minute[29] (Fig. 2C). In contrast, *elt-7* which is regulated by *end-3* in the E lineage[30], has almost half the accumulation rate of *end-3* (10 transcripts/minutes; Fig. 2D) in the E daughter cells. While *elt-7* is produced at similar final transcript levels as *end-3* and higher than *tbx-35*, the longer cell cycle times in the E daughters result in a lower accumulation rate for *elt-7*.

High resolution imaging and lineage tracing of individual cells has provided detailed measurements on the timing of early cell divisions[31]. We used this to calculate the rate of change in transcript number for all High-rate genes (High and Very High categories) from sc-RNA seq data. In several cases (C, E, MS lineages in particular), median rates for each cell were higher in parent cells compared to both daughters (Fig. 2E, mean rate of High-rate genes 2.4-fold greater in 8-cell stage *vs.* 16-cell stage).

High-rate genes include dosage-sensitive transcription factors

To understand the types of genes that are rapidly transcribed and determine if they are involved with developmental fate decisions, we tested for gene ontology categories enriched in all high (High and Very High) *vs.* all Low-rate genes. High-rate genes were approximately two-fold enriched for DNA binding activity (transcription factors) and for involvement in cell fate specification and gastrulation (Fig. 3A, Supplementary Fig. 3A,B). By contrast Low-rate genes are associated with terminal cell type-specific processes (Fig. 3A, Supplementary Fig. 3A,B).

We asked whether embryonic lethality was associated with a particular rate category by testing for enrichment of phenotypes curated by WormBase[32]. Surprisingly, the proportion of Low-rate genes with embryonic lethal phenotypes (either mutant or RNAi) was significantly greater than the High-rate genes (Fig. 3B). We hypothesized that this could be due to the masking of phenotypes due to functional redundancies in High-rate genes. This would mean that two or more of the genes would have to be mutated/removed in order to see a lethal phenotype. Consistent with this, High-rate genes are significantly more likely to have paralogs, which also suggests that these genes are more dosage sensitive (Fig. 3C, ~17-fold enrichment).

We hypothesized that since Low-rate genes are associated with embryonic lethality, they may be more likely to be evolutionarily conserved. Using the Ortholist2[33] tool, we found that a significantly greater proportion of Low-rate genes had a one-to-one human ortholog (Fig. 3D). We extended this analysis to determine the conservation of genes in both rate categories with *C. briggsae*, a close relative of *C. elegans*[34]. Although they diverged ~30 million years ago, both species are similar morphologically and behaviorally[35]. Since High-rate genes were more likely to have paralogs in *C. elegans*, we asked whether they were also more likely to have one or multiple orthologues in *C. briggsae*[36,37]. We found that Low-rate genes had significantly more one-to-one orthologs in *C. briggsae* (~1.7-fold). But, High-rate genes were more likely to have many-to-one (2.5-fold) or many-to-many (3-fold) orthologs in *C. briggsae* (Fig. 3E).

High-rate genes share gene-specific structural features and form genomic clusters

Previous studies have shown that zygotic genes in several organisms tend to be shorter and also have shorter introns compared to maternally expressed genes[38,39]. We examined gene length and determined that Very High and High-rate genes have shorter gene length than Low-rate genes (Fig. 4A, median length very high, 1kb and low, 2.5kb). Higher rate is also associated with shorter introns (Fig 4B, median length very high, 71bp and low, 198bp). The presence of introns has been

linked to higher transcription efficiencies[40,41]. Particularly, transcription was found to be faster in genes with longer first introns[42,43]. We calculated the number of gene length-normalized introns for each rate category and found a decreasing trend between rate and intron number (mean number of introns Very High 1.8 and Low 5.3, Fig. 4D). Our findings are more consistent with the intron delay hypothesis during development, which posits that genes with longer introns are expressed later in development when cell cycles are longer[44].

Around 70% of *C. elegans* genes undergo *trans*-splicing, where the 5' end of RNAs are replaced with a splice leader from a small nuclear riboprotein by the spliceosome complex[45]. We queried published transcription start site (TSS) data[46] to determine if accumulation rates were associated with *trans*-splicing and found that High and Very High-rate genes are significantly less likely to be *trans*-spliced (Fig. 4C, ~3-fold difference between Very High and Low). We also noticed that genes in the Very High and High categories tend to have fewer annotated isoforms (Fig. 4E). It is unclear whether the absence of *trans*-splicing and fewer isoforms helps achieve high-rate transcription; these mechanisms should be explored further in the future.

Position effects, or chromosomal position influence on gene expression is a well-described phenomenon[47]. To begin understanding the impact of chromatin context on transcript accumulation rates, we analyzed the distribution of genes in each rate category across chromosomes. We found a striking over-representation of High and Very High-rate genes on chromosome II (Fig. 4F). Examining the position of High-rate genes within each chromosome shows large clusters of rapidly transcribed genes on the very left arm of chromosome II (Fig. 4G, example cells E, MSx1 and ABprx). Most of these clustered genes encode either F-box genes, which are involved in ubiquitin-mediated proteolysis, or are in the *bath* family (BTB and MATH domain containing). More clusters of genes with fold change >5 are seen on chromosome II compared to other chromosomes (Fig. 4 supplement).

These analyses indicate specific genome architecture is associated with high transcript accumulation and we hypothesize that each feature will influence PolII processivity, allowing for rapid accumulation of cell fate specification genes.

The INR motif is enriched in promoters of High-rate genes

Apart from gene structure associated with high-rate transcription, *cis*-regulatory elements likely play a role in controlling accumulation. Most genes in the embryonic gut gene regulatory network in *C. elegans* are driven by promoter-proximal regions rather than distal enhancers[48,49]. To find for promoter motifs that might control rapid transcription, we first curated bona-fide TSS using three different datasets[46,50,51] (Fig. 5A). We focused on a set of well-studied core promoter motifs (INR, TATA and SP1) and early lineage-specific transcription factor motifs (POP-1, SKN-1 and MED-1), and looked for enrichment of these specific elements in High *vs.* Low-rate genes (Fig. 5B). The initiator element (INR) stood out as having significantly greater motif occurrence in all High-rate genes compared with Low-rate and maternal genes (Fig. 5C). Further, the INR motif overlapped closely with the TSS of High and Very High-rate genes, but there was less enrichment for this motif at the TSS for Low-rate genes (Fig. 5D). Transcription of several *C. elegans* genes has been previously shown to initiate at the 'TCA' core bases of the INR motif[50]. The INR motif is also not enriched at promoters of maternal genes. Binding sites for another canonical core promoter transcription factor SP1, are nearly uniformly enriched in Low, High and Very High-rate genes but not in maternal genes (Fig. 5C), suggesting that distinct core promoter motifs may be used for maternal gene expression.

Multiple promoter motifs contribute to accumulation rates

We used the Very High-rate gene *ceh-51* as a model to test the importance of promoter-proximal motifs for transcript accumulation rates. The T-Box transcription factor TBX-35 is known to regulate *ceh-51* and four binding sites for TBX-35 were previously identified in the upstream region (promoter) of *ceh-51*[29] (Fig. 6A). We used CRISPR to mutate the TBX-35 binding sites. We obtained a mutant strain with a 14bp deletion that partially removed the second and third TBX-35 site; and used smFISH to determine if the accumulation rate of *ceh-51* was affected by this deletion. In each embryo we also estimated the number of *elt-7* transcripts in the embryo as a control as *elt-7* begins to be expressed around the same time as *ceh-51* but should not be affected by promoter deletions in *ceh-51*.

Loss of these TBX-35 sites resulted in a 25% reduction in *ceh-51* accumulation from the 14-cell to the 26-cell stage, when *ceh-51* is maximally expressed (Fig. 6C). Another mutant that has two TBX-35 sites (sites 3 and 4) deleted along with an intervening SP1 site, had a very similar reduction in rate (Fig. 6C). The *C. elegans* homolog of SP1, SPTF-3, is a general transcription factor that has been previously shown to be required for proper expression of gut specification genes and the regulation of gastrulation[52]. We observed that SP1 sites are enriched in all zygotic

rate classes relative to maternal genes (Fig. 5C), further suggesting its importance specifically during embryonic development.

Based on the enrichment and close overlap of the INR motif with the TSS of High-rate genes, we expected that the INR may not only be required for rate control but also for overall expression of High-rate genes. To test this, we deleted the core motif of the INR element and surprisingly found only a 22% reduction in the *ceh-51* accumulation rate (Fig. 6C). There is a weak consensus TATA site upstream of the INR element that may take on the role of transcription initiation in the absence of the initiator element (Fig. 6A), or an unidentified alternative TSS.

Embryonic survival is unaffected in INR and TBX-35 site mutants (Supplementary Fig. 6B). Thus, it appears that promoter element control of accumulation rates has redundancies such that not a single motif but a combination of *cis*-regulatory elements maybe required for the rapid accumulation of *ceh-51*.

Discussion

sc-RNA seq as a useful tool to predict absolute transcript levels

Our comparisons with smFISH counts suggest that the estimation of absolute transcript accumulation rates from scRNA-seq data is fairly robust. Similar correlations between smFISH and scRNA-seq, with errors typically in the +/- 2-fold range, were seen previously in human cell lines, suggesting they may be generalizable across species and cell types[11]. However, the correlation between imaging and sequencing counts during development could potentially be improved by more precise staging and collection of cells for sequencing, as the correspondence between the stages analyzed by scRNA-seq and smFISH can only be inferred within a few minutes using nuclear counts and a small number of marker genes. Recently developed high-throughput FISH methods such as seqFISH+ or MERFISH, which can image up to 10,000 genes, report strong correlations between single molecule transcript counts and RNA seq[12,53]. However, imaging methods still require the design of expensive probe sets. Thus, comparing smFISH counts and sc-RNA seq reads or unique molecular identifiers (UMIs) in multiple model systems should help develop common ‘correct factors’ for absolute transcript inference from sequencing methods.

Cell size and transcription rates

An important normalization for sc-RNA seq counts to estimate absolute mRNA numbers is to account for cell volume. Transcription rates are also known to scale with cell size[23,54]. The *C. elegans* embryo develops within a constant-size chitinase egg shell and cells within the eggshell reduce in size as development progresses. We see higher transcript accumulation rates overall in the 8-cell stage, where cells are larger compared to the 16-cell stage, suggesting that transcription rates in the *C. elegans* embryo are also driven by cell size (Fig. 2D). Future work should focus on identifying mechanisms involved in this scaling.

Coordination between synthesis and decay rates

Absolute transcript levels are representative of both transcript production and degradation. In cell culture systems, during the response to stimuli such as lipopolysaccharides or hypoxia, changes in transcript synthesis contribute more for overall changes in RNA levels than do changes in degradation[55,56]. Our current analysis does not take into account the rate of RNA degradation, and thus our accumulation rates are by definition underestimates of the *de novo* transcription rates, making their accumulation all the more impressive. Many of the regulators we analyze here are transiently expressed, being largely degraded within one or two cell cycles after the time of maximal expression. Thus, these transcripts are in fact rapidly degraded, although the timing of degradation could be limited to these later stages. Similarly, Low-rate genes could be transcribed at high rates if their degradation acts to keep their transcripts levels low. How the interplay between transcription and degradation influences transcript abundance in the *C. elegans* embryo will be an interesting area for future investigation.

Boundaries on the maximum achievable transcription rates

What are the theoretical limits on transcription rates? Several factors contribute to this, including the available number of PolII and other transcription factor molecules, the rates of initiation and elongation, gene length and pause sequences. For short genes such as *end-3* (~1.3kb primary transcript), at the commonly assumed elongation rate of 1.5kb/min[57,58], the entire gene length would need to be completely occupied by PolII complexes during maximum expression. smFISH analysis of early zygotic genes in the *Drosophila* embryo have also measured high rates of

transcription, nearing the steric limits of PolII, suggesting that rates approaching theoretical maximum could be a necessary feature of rapidly dividing embryos[14,59]. Elongation rates have not been directly measured in the *C. elegans* embryo although elongation speeds as high as 6kb/min are seen in mouse embryonic stem cells[43]. It will be interesting to determine whether initiation and elongation rates are coordinately high for genes with high accumulation rates in the worm embryo.

Redundancy and conservation of high-rate transcription in nematode species

We find that High-rate genes are transcription factors that are more likely to have paralogs that could potentially have redundant functions (Fig. 3C). These results are consistent with our hypothesis that rapid transcription is required for transcription factors involved in cell fate decisions. One explanation for the tendency of these transcription factors to have co-expressed paralogs is that the required functional dosage of transcripts is not easily achievable from a single gene, with combined transcripts from paralogous genes required to reach the appropriate threshold level. The significant enrichment of High-rate genes in the many-to-many orthology group suggests that the need for duplicate genes to achieve rapid transcription might be conserved in *C. briggsae*, leading to independent gene duplications (Fig. 3E). We also find that clusters of High-rate genes (for example on chromosome II) belong to the same gene family, which could have also risen out of duplication (Fig. 4F). Genes involved in gut specification, such as the *meds* and *ends*, tend to be present as multiple paralogs in the Elegans supergroup and these genes also exhibit synteny across species[48]. Whether the spatial organization of High-rate genes is required for their expression is an important question to be addressed in the future.

We have characterized transcription accumulation rates in a developing multi-cellular organism and find that the regulation of rates is highly complex. We hypothesize that high rates are required to achieve precision in final transcript levels, which drives fate specification. How this precision will translate to protein levels is unclear. Quantitative tools to measure absolute protein levels in single cells will help us determine how high transcription rates are related to protein levels, and how this ultimately regulates developmental robustness.

Methods

Accumulation rates from sc-RNA seq data: All analysis was performed in R. Gene names from the Tintori et al. dataset were transferred to WS260 reference. Median Transcript per million (TPM) counts for each gene in each was calculated, cells annotated as "tossed" were not used. Median counts were corrected for cell volume and number of transcripts/cell (8000000 adjustment, which fits the expected range of mRNA concentration based on the slope of 0.95 calculated from the fitted line in Figure 2B), to get absolute transcript amounts. Absolute change was calculated as daughter minus parent; and fold change between daughter and parent used a pseudocount of 10. Genes with absolute transcript amounts > 100 in the P0 cell were considered to be maternal genes.

Rate categorization: Distributions of absolute change (AC) and fold change (FC) (Fig.1B) were used to bin rate categories. AC >200, FC > 10 was considered Very High-rate, AC between 50-200, FC between 5-10 as High-rate and AC between 1-50 and between FC 1-5 as Low-rate. If the AC or FC was Very High but the other fell into the High category, it was classified as High. Similarly, if AC and FC were High or Very High, AC > 50 or FC > 5, but FC was Low (1-5), those genes were classified as Low-rate. For gene structure and motif analysis, only High and Very High genes that were never Low-rate in any cell were included in the analysis.

Gene Ontology and conservation with *C. briggsae*: GO was performed using PANTHER tools (http://geneontology.org/). OrthoFinder was used to examine conservation between *C. elegans* and *C. briggsae*[37]. Longest transcript isoforms were used to identify Orthogroups, which were then categorized as having one-to-one, one-to-many, many-to-one or many-to-many orthologous genes between *C. elegans* and *C. brissgae*.

Paralogous genes and embryonic lethal phenotype: Paralog list was obtained from Tintori et al.[21] and Wormbase SimpleMine was used to identify allele and RNAi mutant phenotypes (https://wormbase.org//tools/mine/simplemine.cgi).

Motif tools: The Find Individual Motif Occurrences (FIMO) tools in the Meme suite was used for motif enrichment using a p-value cut-off of $1E^{-4}$ and the R seqLogo package to plot motifs.

Worm stains and growth: Worms were grown on standard NGM plates on OP50 bacteria. For smFISH, worms were grown on large, enriched peptone plates seeded with NA22 bacteria. All growth was at 20°C unless otherwise stated.

Single molecule RNA FISH (smFISH): smFISH was performed as described in Nair et al., 2013[19]. Briefly, adult worms were treated with alkaline bleach to release embryos, which were washed with M9 buffer. Embryos were fixed with 4% formaldehyde (in PBS) and permeabilized by immersing the tube in dry ice and ethanol bath for 1 minute. After a 20-minute incubation on ice, embryos were washed with PBS and stored in 70% ethanol at 4°C. For hybridization with FISH probes, first the ethanol was washed off the embryos with wash buffer (10% formamide in 2X SSC with 0.1% Triton X-100). FISH probes were used at 2nM concentration in 100µl of hybridization buffer (0.1% dextran sulphate, 0.1% formamide in 2X SSC buffer). Hybridization was performed overnight at 37°C. Before imaging, embryos were stained with Hoechst stain for 30 minutes and mounted on a cover slip in 2X SSC. smFISH images were taken with a standard epifluorescence scope.

smFISH transcript counts: TIFF stack images were processed using Matlab. Embryos were segmented and a Laplacian of Gaussian filter was applied to isolate single transcripts. Transcript counts were thresholded manually to obtain spot counts across all Z stacks. Each channel for different probe sets were separately thresholded. Hoechst stain was used to determine the number of nuclei and cell stage.

CRISPR deletions: Promoter deletions were performed by CRISPR using short single-stranded guide RNAs (as described in Dokshin et al., 2018[60]). Guides and tracrRNA were obtained from Integrated DNA Technologies (IDT). *S. pyogenes* Cas9 was purchased from the QB3 MacroLab, UC Berkley. Co-injection marker sur-5::GFP. Deletions were confirmed by sequencing.

Acknowledgements

We are extremely grateful to Arjun Raj and members of the Raj lab for their generosity in letting us use their microscopes, smFISH troubleshooting, advice on image analysis and overall, for being fantastic neighbors. We would also like to thank the members for the Murray Lab (past and present) and the Penn Worm Group for their insightful discussions and comments.

References

1. Packer, J. S. *et al.* A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science* **365**, doi:10.1126/science.aax1971 (2019).
2. Farrell, J. A. *et al.* Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, doi:10.1126/science.aar3131 (2018).
3. He, P. *et al.* The changing mouse embryo transcriptome at whole tissue and single-cell resolution. *Nature* **583**, 760-767, doi:10.1038/s41586-020-2536-x (2020).
4. Briggs, J. A. *et al.* The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* **360**, doi:10.1126/science.aar5780 (2018).
5. Qiu, Q. *et al.* Massively parallel and time-resolved RNA sequencing in single cells with scNT-seq. *Nat Methods* **17**, 991-1001, doi:10.1038/s41592-020-0935-4 (2020).
6. Herzog, V. A. *et al.* Thiol-linked alkylation of RNA to assess expression dynamics. *Nat Methods* **14**, 1198-1204, doi:10.1038/nmeth.4435 (2017).
7. Schwalb, B. *et al.* TT-seq maps the human transient transcriptome. *Science* **352**, 1225-1228, doi:10.1126/science.aad9841 (2016).
8. Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**, 877-879, doi:10.1038/nmeth.1253 (2008).
9. Femino, A. M., Fay, F. S., Fogarty, K. & Singer, R. H. Visualization of single RNA transcripts in situ. *Science* **280**, 585-590, doi:10.1126/science.280.5363.585 (1998).
10. Lyubimova, A. *et al.* Single-molecule mRNA detection and counting in mammalian tissue. *Nat Protoc* **8**, 1743-1758, doi:10.1038/nprot.2013.109 (2013).
11. Torre, E. *et al.* Rare Cell Detection by Single-Cell RNA Sequencing as Guided by Single-Molecule RNA FISH. *Cell Syst* **6**, 171-179.e175, doi:10.1016/j.cels.2018.01.014 (2018).
12. Eng, C. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235-239, doi:10.1038/s41586-019-1049-y (2019).
13. Zhang, M. J., Ntranos, V. & Tse, D. Determining sequencing depth in a single-cell RNA-seq experiment. *Nat Commun* **11**, 774, doi:10.1038/s41467-020-14482-y (2020).
14. Little, S. C., Tikhonov, M. & Gregor, T. Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell* **154**, 789-800, doi:10.1016/j.cell.2013.07.025 (2013).
15. Raj, A., Rifkin, S. A., Andersen, E. & van Oudenaarden, A. Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913-918, doi:10.1038/nature08781 (2010).

16. Ochiai, H., Sugawara, T., Sakuma, T. & Yamamoto, T. Stochastic promoter activation affects Nanog expression variability in mouse embryonic stem cells. *Sci Rep* **4**, 7125, doi:10.1038/srep07125 (2014).
17. Owens, N. D. L. *et al.* Measuring Absolute RNA Copy Numbers at High Temporal Resolution Reveals Transcriptome Kinetics in Development. *Cell Rep* **14**, 632-647, doi:10.1016/j.celrep.2015.12.050 (2016).
18. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev Biol* **100**, 64-119, doi:10.1016/0012-1606(83)90201-4 (1983).
19. Nair, G., Walton, T., Murray, J. I. & Raj, A. Gene transcription is coordinated with, but not dependent on, cell divisions during C. elegans embryonic fate specification. *Development* **140**, 3385-3394, doi:10.1242/dev.098012 (2013).
20. Choi, H., Broitman-Maduro, G. & Maduro, M. F. Partially compromised specification causes stochastic effects on gut development in C. elegans. *Dev Biol* **427**, 49-60, doi:10.1016/j.ydbio.2017.05.007 (2017).
21. Tintori, S. C., Osborne Nishimura, E., Golden, P., Lieb, J. D. & Goldstein, B. A Transcriptional Lineage of the Early C. elegans Embryo. *Dev Cell* **38**, 430-444, doi:10.1016/j.devcel.2016.07.025 (2016).
22. Padovan-Merhar, O. *et al.* Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Mol Cell* **58**, 339-352, doi:10.1016/j.molcel.2015.03.005 (2015).
23. Sun, X. M. *et al.* Size-Dependent Increase in RNA Polymerase II Initiation Rates Mediates Gene Expression Scaling with Cell Size. *Curr Biol* **30**, 1217-1230.e1217, doi:10.1016/j.cub.2020.01.053 (2020).
24. Edgar, L. G., Wolf, N. & Wood, W. B. Early transcription in Caenorhabditis elegans embryos. *Development* **120**, 443-451 (1994).
25. Seydoux, G. & Fire, A. Soma-germline asymmetry in the distributions of embryonic RNAs in Caenorhabditis elegans. *Development* **120**, 2823-2834 (1994).
26. Spencer, W. C. *et al.* A spatial and temporal map of C. elegans gene expression. *Genome Res* **21**, 325-341, doi:10.1101/gr.114595.110 (2011).
27. Wang, J. T. & Seydoux, G. Germ cell specification. *Adv Exp Med Biol* **757**, 17-39, doi:10.1007/978-1-4614-4015-4_2 (2013).

28. Broitman-Maduro, G., Lin, K. T., Hung, W. W. & Maduro, M. F. Specification of the C. elegans MS blastomere by the T-box factor TBX-35. *Development* **133**, 3097-3106, doi:10.1242/dev.02475 (2006).
29. Broitman-Maduro, G. *et al.* The NK-2 class homeodomain factor CEH-51 and the T-box factor TBX-35 have overlapping function in C. elegans mesoderm development. *Development* **136**, 2735-2746, doi:10.1242/dev.038307 (2009).
30. Dineen, A., Osborne Nishimura, E., Goszczynski, B., Rothman, J. H. & McGhee, J. D. Quantitating transcription factor redundancy: The relative roles of the ELT-2 and ELT-7 GATA factors in the C. elegans endoderm. *Dev Biol* **435**, 150-161, doi:10.1016/j.ydbio.2017.12.023 (2018).
31. Richards, J. L., Zacharias, A. L., Walton, T., Burdick, J. T. & Murray, J. I. A quantitative model of normal Caenorhabditis elegans embryogenesis and its disruption after stress. *Dev Biol* **374**, 12-23, doi:10.1016/j.ydbio.2012.11.034 (2013).
32. Wormbase release WS282. http://www.wormbase.org (2021).
33. Kim, W., Underwood, R. S., Greenwald, I. & Shaye, D. D. OrthoList 2: A New Comparative Genomic Analysis of Human and. *Genetics* **210**, 445-461, doi:10.1534/genetics.118.301307 (2018).
34. Stein, L. D. *et al.* The genome sequence of Caenorhabditis briggsae: a platform for comparative genomics. *PLoS Biol* **1**, E45, doi:10.1371/journal.pbio.0000045 (2003).
35. Cutter, A. D., Dey, A. & Murray, R. L. Evolution of the Caenorhabditis elegans genome. *Mol Biol Evol* **26**, 1199-1234, doi:10.1093/molbev/msp048 (2009).
36. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* **16**, 157, doi:10.1186/s13059-015-0721-2 (2015).
37. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* **20**, 238, doi:10.1186/s13059-019-1832-y (2019).
38. Heyn, P. *et al.* The earliest transcribed zygotic genes are short, newly evolved, and different across species. *Cell Rep* **6**, 285-292, doi:10.1016/j.celrep.2013.12.030 (2014).
39. Heyn, P., Kalinka, A. T., Tomancak, P. & Neugebauer, K. M. Introns and gene expression: cellular constraints, transcriptional regulation, and evolutionary consequences. *Bioessays* **37**, 148-154, doi:10.1002/bies.201400138 (2015).
40. Jeffares, D. C., Penkett, C. J. & Bähler, J. Rapidly regulated genes are intron poor. *Trends Genet* **24**, 375-378, doi:10.1016/j.tig.2008.05.006 (2008).

41. Brinster, R. L., Allen, J. M., Behringer, R. R., Gelinas, R. E. & Palmiter, R. D. Introns increase transcriptional efficiency in transgenic mice. *Proc Natl Acad Sci U S A* **85**, 836-840, doi:10.1073/pnas.85.3.836 (1988).
42. Fukaya, T., Lim, B. & Levine, M. Rapid Rates of Pol II Elongation in the Drosophila Embryo. *Curr Biol* **27**, 1387-1391, doi:10.1016/j.cub.2017.03.069 (2017).
43. Jonkers, I., Kwak, H. & Lis, J. T. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* **3**, e02407, doi:10.7554/eLife.02407 (2014).
44. Swinburne, I. A. & Silver, P. A. Intron delays and transcriptional timing during development. *Dev Cell* **14**, 324-330, doi:10.1016/j.devcel.2008.02.002 (2008).
45. Allen, M. A., Hillier, L. W., Waterston, R. H. & Blumenthal, T. A global analysis of C. elegans trans-splicing. *Genome Res* **21**, 255-264, doi:10.1101/gr.113811.110 (2011).
46. Saito, T. L. *et al.* The transcription start site landscape of C. elegans. *Genome Res* **23**, 1348-1361, doi:10.1101/gr.151571.112 (2013).
47. Wilson, C., Bellen, H. J. & Gehring, W. J. Position effects on eukaryotic gene expression. *Annu Rev Cell Biol* **6**, 679-714, doi:10.1146/annurev.cb.06.110190.003335 (1990).
48. Maduro, M. F. Evolutionary Dynamics of the SKN-1 → MED → END-1,3 Regulatory Gene Cascade in. *G3 (Bethesda)* **10**, 333-356, doi:10.1534/g3.119.400724 (2020).
49. Grishkevich, V., Hashimshony, T. & Yanai, I. Core promoter T-blocks correlate with gene expression levels in C. elegans. *Genome Res* **21**, 707-717, doi:10.1101/gr.113381.110 (2011).
50. Chen, R. A. *et al.* The landscape of RNA polymerase II transcription initiation in C. elegans reveals promoter and enhancer architectures. *Genome Res* **23**, 1339-1347, doi:10.1101/gr.153668.112 (2013).
51. Kruesi, W. S., Core, L. J., Waters, C. T., Lis, J. T. & Meyer, B. J. Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. *Elife* **2**, e00808, doi:10.7554/eLife.00808 (2013).
52. Sullivan-Brown, J. L. *et al.* Identifying Regulators of Morphogenesis Common to Vertebrate Neural Tube Closure and Caenorhabditis elegans Gastrulation. *Genetics* **202**, 123-139, doi:10.1534/genetics.115.183137 (2016).
53. Xia, C., Fan, J., Emanuel, G., Hao, J. & Zhuang, X. Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene

expression. *Proc Natl Acad Sci U S A* **116**, 19490-19499, doi:10.1073/pnas.1912459116 (2019).

54. Metzl-Raz, E., Kafri, M., Yaakov, G. & Barkai, N. Gene Transcription as a Limiting Factor in Protein Production and Cell Growth. *G3 (Bethesda)* **10**, 3229-3242, doi:10.1534/g3.120.401303 (2020).
55. Rabani, M. *et al.* High-resolution sequencing and modeling identifies distinct dynamic RNA regulatory strategies. *Cell* **159**, 1698-1710, doi:10.1016/j.cell.2014.11.015 (2014).
56. Tiana, M. *et al.* Metabolic labeling of RNA uncovers the contribution of transcription and decay rates on hypoxia-induced changes in RNA levels. *RNA* **26**, 1006-1022, doi:10.1261/rna.072611.119 (2020).
57. Jonkers, I. & Lis, J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 167-177, doi:10.1038/nrm3953 (2015).
58. Kwak, H. & Lis, J. T. Control of transcriptional elongation. *Annu Rev Genet* **47**, 483-508, doi:10.1146/annurev-genet-110711-155440 (2013).
59. Boettiger, A. N. & Levine, M. Rapid transcription fosters coordinate snail expression in the Drosophila embryo. *Cell Rep* **3**, 8-15, doi:10.1016/j.celrep.2012.12.015 (2013).
60. Dokshin, G. A., Ghanta, K. S., Piscopo, K. M. & Mello, C. C. Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in. *Genetics* **210**, 781-787, doi:10.1534/genetics.118.301532 (2018).

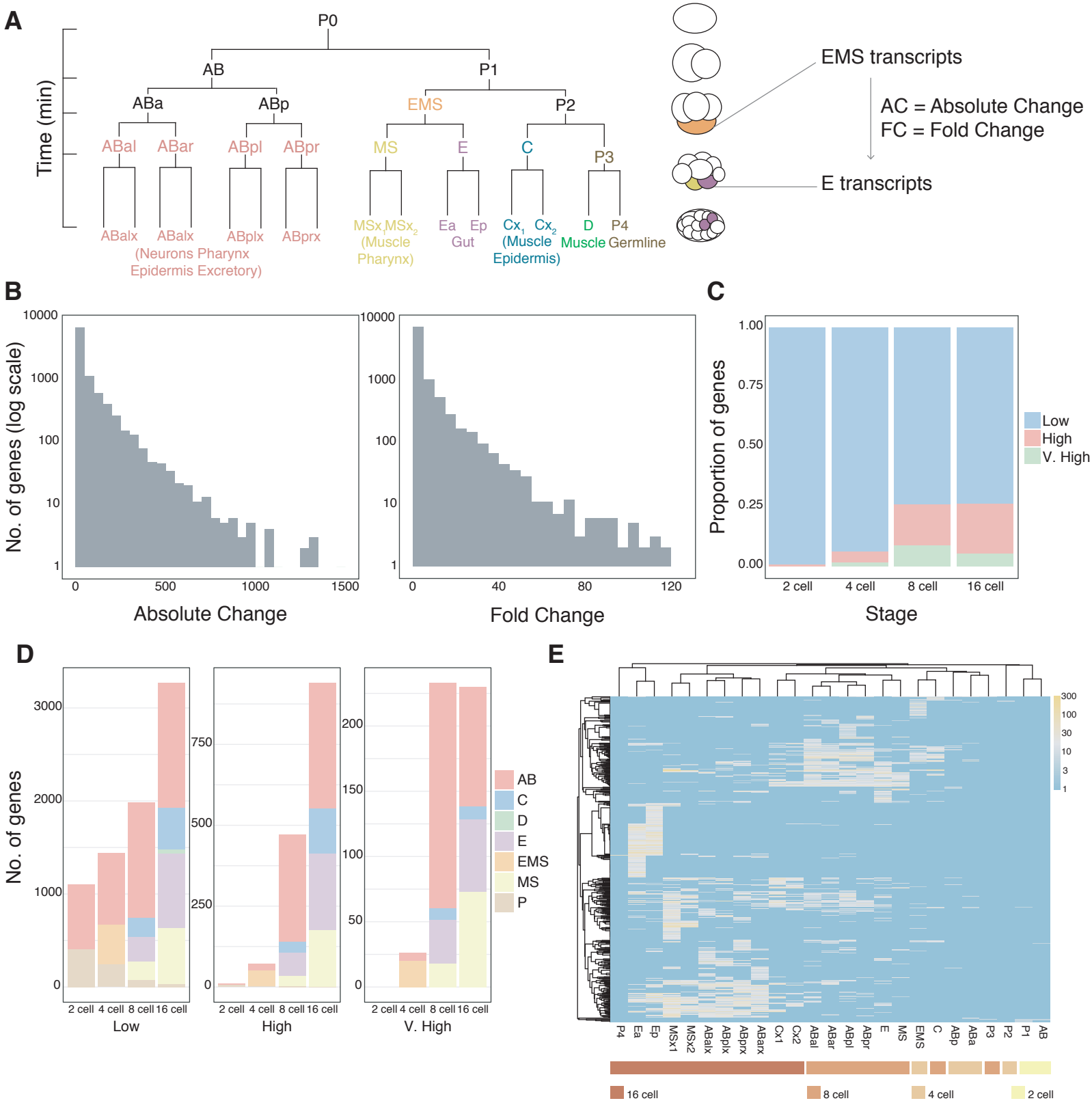
A
P0
AB
P1
ABa
ABp
EMS
P2
ABal
ABar
ABpl
ABpr
MS
E
C
P3
ABalx
ABalx
ABplx
ABprx
MSx1 MSx2
Ea Ep
Cx1 Cx2
D P4
(Neurons Pharynx
Epidermis Excretory)
(Muscle
Pharynx)
Gut
(Muscle
Epidermis)
Muscle Germline
Time (min)
EMS transcripts
AC = Absolute Change
FC = Fold Change
E transcripts
B
No. of genes (log scale)
Absolute Change
Fold Change
C
Proportion of genes
Stage
2 cell
4 cell
8 cell
16 cell
Low
High
V. High
D
No. of genes
Low
High
V. High
AB
C
D
E
EMS
MS
P
E
16 cell
8 cell
4 cell
2 cell

Figure 1

Figure 1. Using single-cell RNA sequencing data to estimate absolute transcription rates. **A.** Lineage tree of the *C. elegans* embryo upto the 24-cell stage (AB16) showing all the founder cells (AB, EMS, MS, E, C, D and P) and simplified list of the tissue types that they produce. Accumulation rates were determined by change in the number of transcripts in daughters compared to the parent cell (see Methods). **B.** Histogram distribution of absolute change (AC) and fold change (FC). **C.** Fraction of the transcriptome at each embryonic stage based on rate category - very high (V. High), High and Low. **D.** Number of genes in each rate category by embryo stage and main founder lineages. **E.** Heat map of all high-rate genes (very high and high) using average method of clustering with Pearson distances. Colored bars show embryo stage.

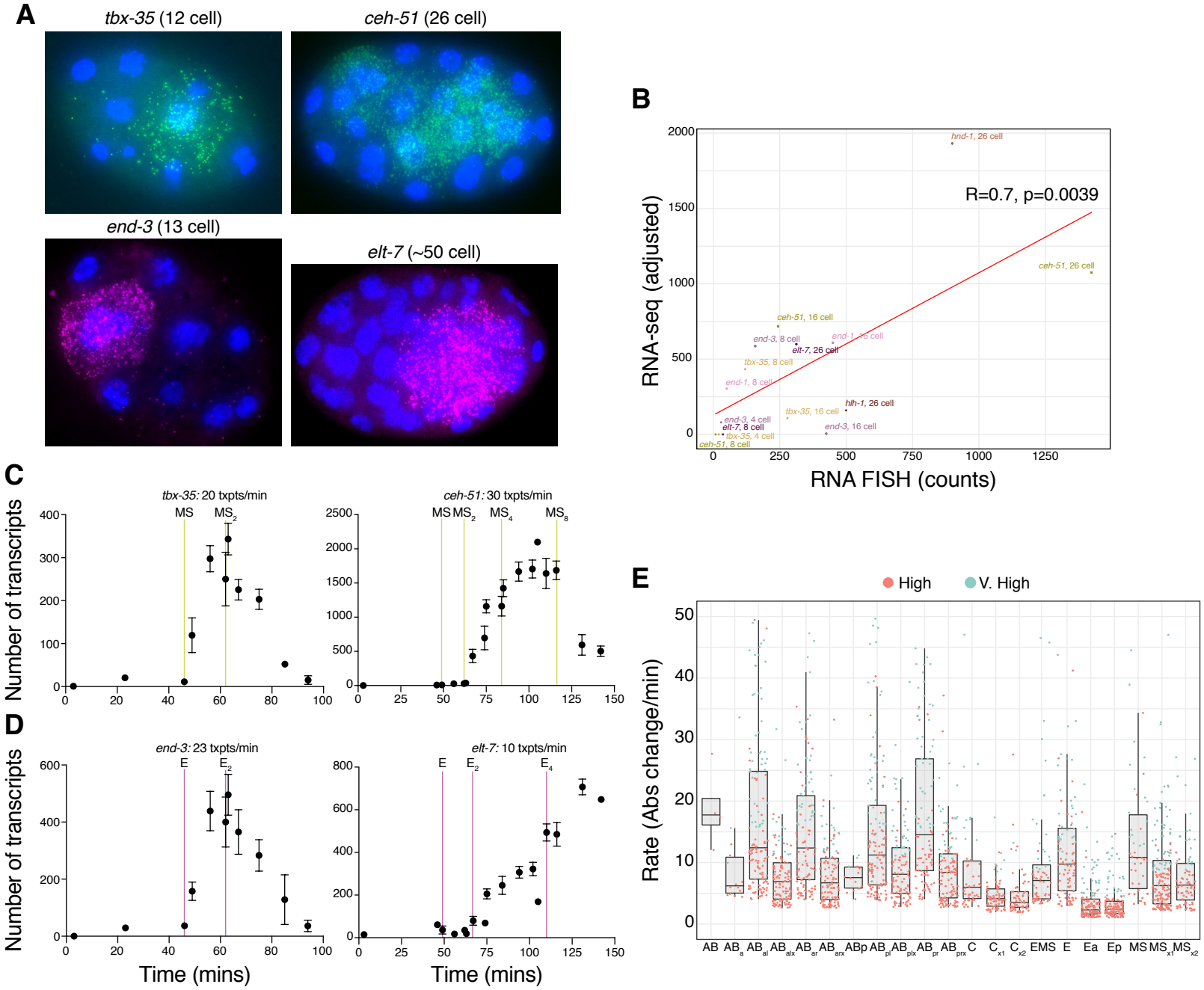
A
tbx-35 (12 cell)
ceh-51 (26 cell)
end-3 (13 cell)
elt-7 (~50 cell)
B
RNA-seq (adjusted)
RNA FISH (counts)
R=0.7, p=0.0039
hnd-1, 26 cell
ceh-51, 26 cell
ceh-51, 16 cell
end-3, 8 cell
end-1, 16 cell
elt-7, 26 cell
tbx-35, 8 cell
end-1, 8 cell
hlh-1, 26 cell
tbx-35, 16 cell
end-3, 4 cell
elt-7, 8 cell
tbx-35, 4 cell
end-3, 16 cell
ceh-51, 8 cell
C
tbx-35: 20 txpts/min
ceh-51: 30 txpts/min
Number of transcripts
D
end-3: 23 txpts/min
elt-7: 10 txpts/min
Time (mins)
E
High
V. High
Rate (Abs change/min)
AB
ABp
EMS
E
Ea
Ep
MS

Figure 1

Figure 2. Comparison of inferred rates from sc-RNA seq with smFISH.

A. Representative single molecule FISH (smFISH) images (maximum intensity projections of Z stacks) of genes indicated at specific embryo stages. **B.** Correlation between inferred absolute transcript numbers from sc-RNA seq and smFISH (Pearson correlation). **C.** smFISH counts of muscle transcription factors *tbx-35* (left) and *ceh-51* (right) over developmental time were used to calculate transcript accumulation rates between the two cell divisions of maximal expression. **D.** Same as C for *end-3* (left) and *elt-7* (right) involved in intestinal specification. **E.** Transcript accumulation rates (absolute change/min) for all high-rate genes (V. High and high) across cell types in the AB, C, EMS, E and MS lineages.

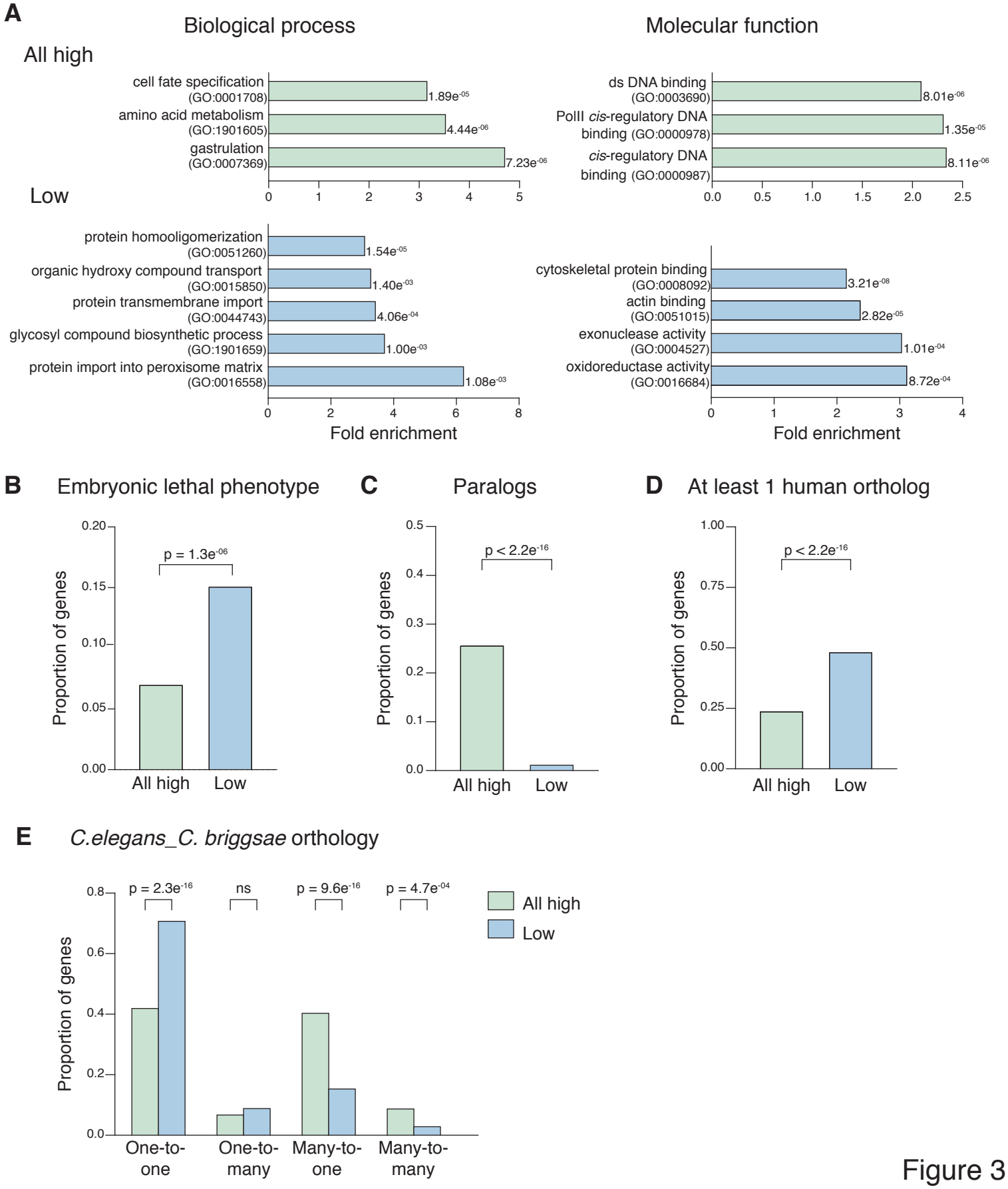
A
Biological process
Molecular function
All high
cell fate specification (GO:0001708)
amino acid metabolism (GO:1901605)
gastrulation (GO:0007369)
1.89e-05
4.44e-06
7.23e-06
ds DNA binding (GO:0003690)
PolII cis-regulatory DNA binding (GO:0000978)
cis-regulatory DNA binding (GO:0000987)
8.01e-06
1.35e-05
8.11e-06
Low
protein homooligomerization (GO:0051260)
organic hydroxy compound transport (GO:0015850)
protein transmembrane import (GO:0044743)
glycosyl compound biosynthetic process (GO:1901659)
protein import into peroxisome matrix (GO:0016558)
1.54e-05
1.40e-03
4.06e-04
1.00e-03
1.08e-03
cytoskeletal protein binding (GO:0008092)
actin binding (GO:0051015)
exonuclease activity (GO:0004527)
oxidoreductase activity (GO:0016684)
3.21e-08
2.82e-05
1.01e-04
8.72e-04
Fold enrichment
Fold enrichment
B
Embryonic lethal phenotype
Proportion of genes
p = 1.3e-06
All high
Low
C
Paralogs
Proportion of genes
p < 2.2e-16
All high
Low
D
At least 1 human ortholog
Proportion of genes
p < 2.2e-16
All high
Low
E
C.elegans_C. briggsae orthology
Proportion of genes
p = 2.3e-16
ns
p = 9.6e-16
p = 4.7e-04
All high
Low
One-to-one
One-to-many
Many-to-one
Many-to-many

Figure 3

Figure 3. Function and evolutionary conservation of high vs. low-rate genes.

A. Gene ontology categories that show enrichment within all high (V. High and high combined) and low-rate genes, from PANTHER analysis. **B.** Proportion of all high and low-rate genes with embryonic lethal phenotypes. **C.** Proportion of all high and low-rate genes that have a paralog. **D.** Proportion of all high and low-rate genes with at least one human ortholog **E.** Orthogroups identified between *C. elegans* and *C. briggsae* in all high vs. low-rate genes. All reported *p*-values from Fisher's exact test.

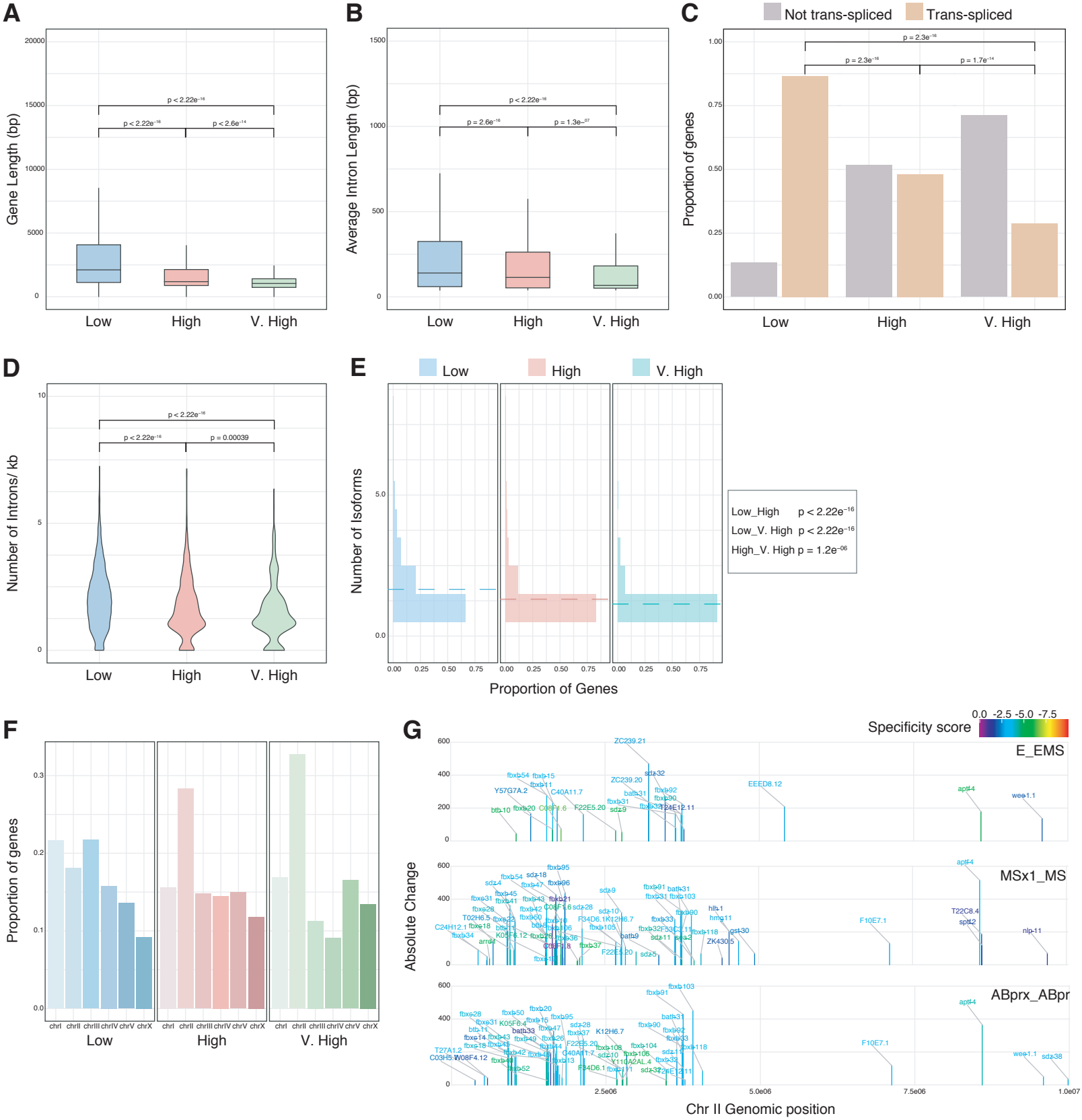
A
Gene Length (bp)
p < 2.22e-16
p < 2.22e-16
p < 2.6e-14
Low
High
V. High
B
Average Intron Length (bp)
p < 2.22e-16
p = 2.6e-16
p = 1.3e-07
Low
High
V. High
C
Not trans-spliced
Trans-spliced
Proportion of genes
p = 2.3e-16
p = 2.3e-16
p = 1.7e-14
Low
High
V. High
D
Number of Introns/ kb
p < 2.22e-16
p < 2.22e-16
p = 0.00039
Low
High
V. High
E
Low
High
V. High
Number of Isoforms
Low_High p < 2.22e-16
Low_V. High p < 2.22e-16
High_V. High p = 1.2e-06
Proportion of Genes
F
Proportion of genes
chrI chrII chrIII chrIV chrV chrX
Low
High
V. High
G
Specificity score
0.0 -2.5 -5.0 -7.5
E_EMS
MSx1_MS
ABprx_ABpr
Absolute Change
Chr II Genomic position

Figure 4

Figure 4. Genetic architecture and features associated with high transcript accumulation. **A,B**. Gene and intron lengths across different rate categories. **C**. Proportion of genes that are trans-spliced or not in each rate group using data from Saito et al., 2013. **D.** Length normalized intron number in each rate category. **E**. Distribution of isoform number, based on proportion of genes in each rate category. Dotted line shows mean. All reported significance from Wilcox test, adjusted *p*-values. **F**. Proportion of genes found on chromosomes I-V and X. **G**. Karyogram of genes with FC >5 on chromosome II showing absolute change for each gene across genomic position in the indicated cells compared to their parent. Specificity score indicates how specific the expression is to the particular cell (purple is specific to that cell, red is broadly expressed in all cells).

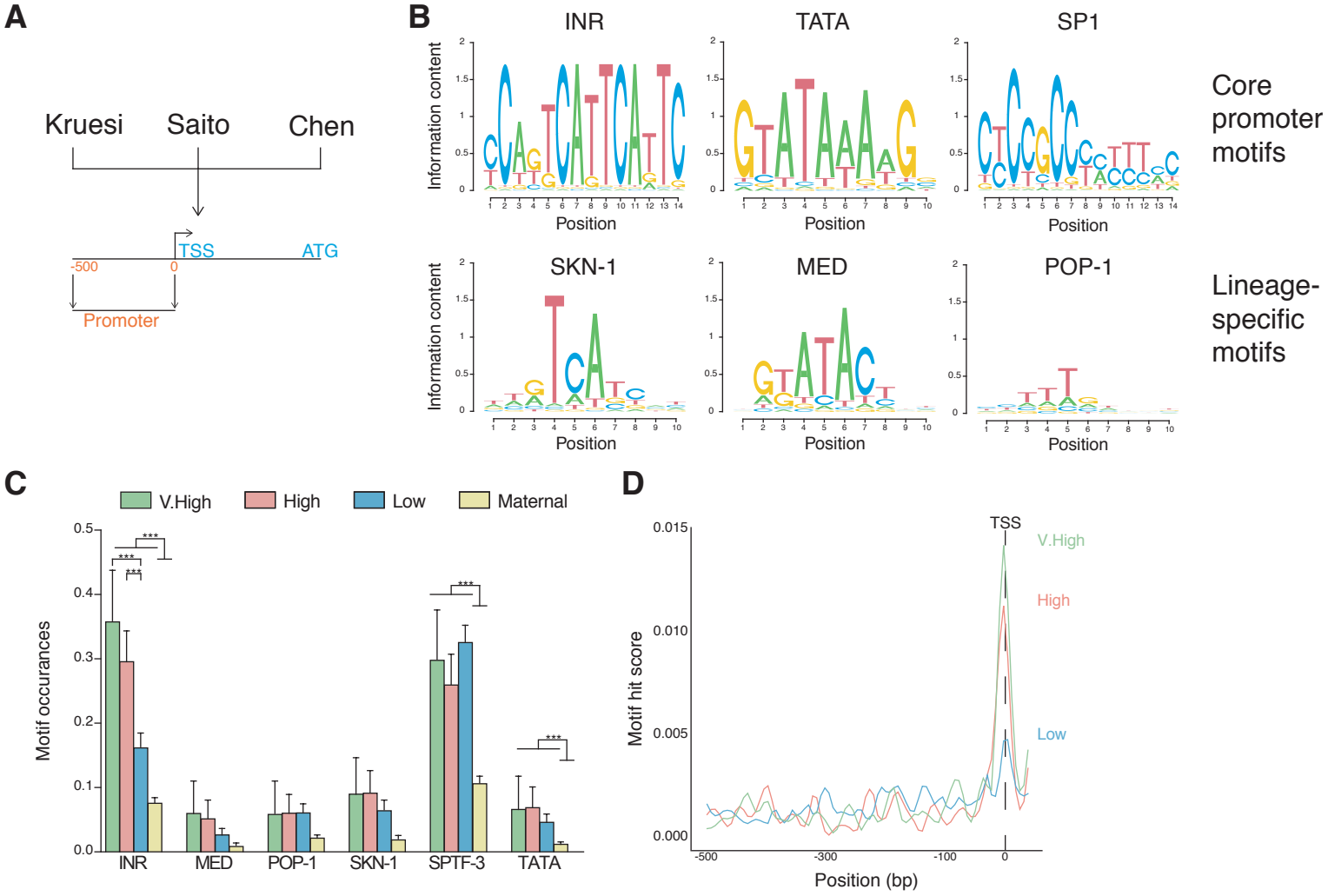
A
Kruesi
Saito
Chen
TSS
ATG
-500
0
Promoter
B
INR
TATA
SP1
Information content
Position
Core promoter motifs
SKN-1
MED
POP-1
Lineage-specific motifs
C
V.High
High
Low
Maternal
Motif occurances
INR
MED
POP-1
SKN-1
SPTF-3
TATA
D
TSS
V.High
High
Low
Motif hit score
Position (bp)

Figure 5

Figure 5. Promoter motifs associated with rapid transcription.

A. Method for combining transcription start site (TSS) data from Saito, Chen and Kruesi datasets. Liftover from ce10 to ce11 genome assemblies performed when necessary. **B**. PWMs of core promoter and lineage-specific motifs. INR and SP1 motifs from the MEME suite (https://meme-suite.org/meme/), TATA motif from Chen et al., 2013, Lineage-specific TF motifs from CIS-BP (http://cisbp.ccbr.utoronto.ca/) **C**. Motif occurrences for the 6 elements in B across different rate categories and maternal genes, 95% confidence intervals as error bars, *p*-values from test for equal proportions (prop.test), FDR adjusted *p*-values. **D**. INR motif distribution across the promoter (500bp upstream) in V. High, high and low-rate genes showing overlap of INR with the TSS.

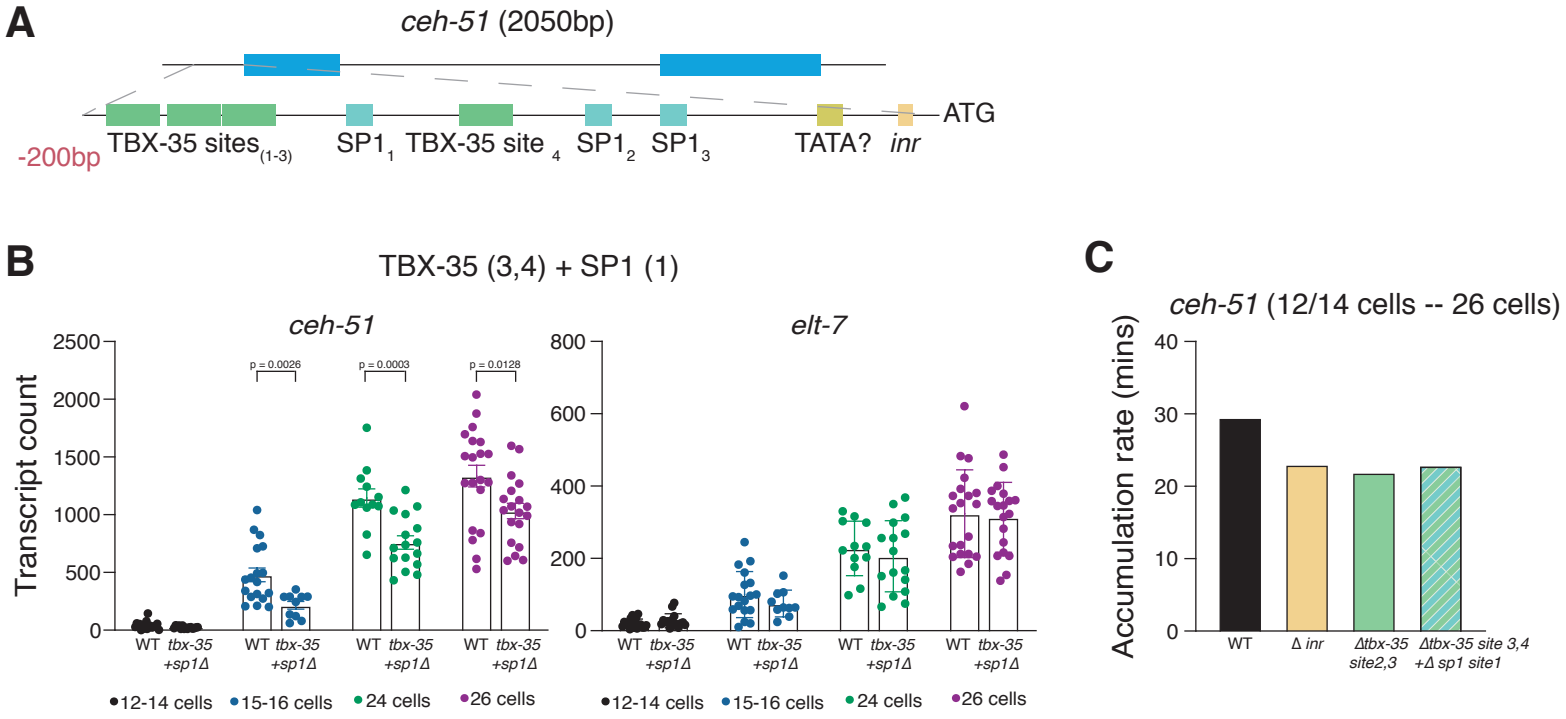
A
ceh-51 (2050bp)
ATG
-200bp
TBX-35 sites(1-3)
SP1 1
TBX-35 site 4
SP1 2
SP1 3
TATA?
inr
B
TBX-35 (3,4) + SP1 (1)
ceh-51
elt-7
Transcript count
p = 0.0026
p = 0.0003
p = 0.0128
WT
tbx-35 +sp1Δ
12-14 cells
15-16 cells
24 cells
26 cells
C
ceh-51 (12/14 cells -- 26 cells)
Accumulation rate (mins)
WT
Δ inr
Δtbx-35 site2,3
Δtbx-35 site 3,4 +Δ sp1 site1

Figure 6

Figure 6. Contribution of promoter motifs to regulation of *ceh-51* transcript accumulation rate.

A. 200bp upstream region (from translation start ATG) of *ceh-51*. **B.** smFISH exon counts of *ceh-51* and *elt-7* at the indicated stages in the promoter mutant where TBX-35 sites 3 and 4 and SP1 site 1 are deleted, n = at least 3 biological replicates, *p*-values from Mann-Whitney test. **C.** Accumulation rate of *ceh-51* (from the 12/14 cell stage - 26 cell stage) in the indicated mutants, data from at least 3 biological replicates.