Title

Tissue-specific transcription footprinting using RNA PoI DamID (RAPID) in *C. elegans*

Article summary

Gene expression is a major determinant of cell fate and physiology, yet it is notoriously difficult to characterize for the widely used model system *C. elegans*. By using a new method based on *in vivo* covalent modification of DNA by RNA polymerases, we determine genome-wide transcription patterns in single tissues of embryos or young adult animals. We show that the method is able to identify actively transcribed genes in tissues representing down to 0.2% of the somatic cells in adult animals. Additionally, this method can be fully performed in a single laboratory by using third generation sequencing methods (ONT).

Authors

Georgina Gómez-Saldivar[*,1], Jaime Osuna-Luque[*,2,3,4], Dominique A. Glauser[!,1], Sophie Jarriault[!,3] and Peter Meister[!,2]

[*] These authors contributed equally to this work.

[1] University of Fribourg, Switzerland

[2] Cell Fate and Nuclear Organization, Institute of Cell Biology, University of Bern, Switzerland

[3] Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC); Institut National de la Santé et de la Recherche Médicale (INSERM) U964/Centre National de la Recherche Scientifique (CNRS) UMR 1704/Université de Strasbourg, France.

[4] Graduate School for Cellular and Biomedical Sciences, University of Bern, Switzerland

[!] to whom correspondence should be sent.

Abstract

Differential gene expression across cell types underlies the development and cell physiology in multicellular organisms. *C. elegans* is a powerful, extensively used model to address these biological questions. A remaining bottleneck relates, however, to the difficulty to obtain comprehensive tissue-specific gene transcription data, since available methods are still challenging to execute and/or require large worm populations. Here, we introduce the RNA Pol DamID (RAPID) approach, in which the Dam methyltransferase is fused to an ubiquitous RNA polymerase subunit in order to create transcriptional footprints *via* methyl marks on the DNA of transcribed genes. To validate the method, we determined the polymerase footprints in whole animals, sorted embryonic blastomeres and in different tissues from intact young adults by driving Dam fusion expression tissue-specifically. We obtained meaningful transcriptional footprints in line with RNA-seq studies in whole animals or specific tissues. To challenge the sensitivity of RAPID and demonstrate its utility to determine novel tissue-specific transcriptional profiles, we determined the transcriptional footprints of the pair of XXX neuroendocrine cells, representing 0.2% of the somatic cell content of the animals. We identified 2362 candidate genes with putatively active transcription in XXX cells, among which the few known markers for these cells. Using transcriptional reporters for a subset of new hits, we confirmed their XXX expression for a majority of them and identified novel XXX-specific markers. Taken together, our work establishes RAPID as a valid method for the determination of polymerase footprints in specific tissues of *C. elegans* without the need for cell sorting or RNA tagging.

Introduction

Differential gene expression across cell types encompasses both key determinants and markers of the cells' identity. Cataloging these differences can provide critical insights and entry points in research aiming at elucidating the mechanisms controlling fundamental biological processes, such as organismal development and cell/tissue physiology. *C. elegans* is a widely used model animal, particularly well suited for integrative studies bridging our understanding across the molecular, cellular and organismal levels. With its transparent body, *C. elegans* was the first animal used to analyze tissue-specific transcription *in vivo* with GFP reporters (Chalfie et al. 1994), an approach still extensively used. In contrast to the versatility of the model for individual gene expression analysis, more holistic approaches such as tissue-specific transcriptomics are much more challenging. Two main strategies have been developed to analyse specific tissues/cell types in *C. elegans*. The first one relies on the animal disruption and cell isolation, or dissociation, followed by the sorting of labelled cells (Von Stetina et al. 2007; Spencer et al. 2011), or nuclei (Steiner et al. 2012; Haenni et al. 2012) prior to transcriptomics analysis. While some initial protocols worked better in embryos than larvae or older animals probably due to the tough cuticle, more recent ones were successfully used to analyze major tissues (Kaletsky et al. 2018) as well as neuronal subsets in adults (down to 6 neurons per animals; Kaletsky et al. 2016; Wang et al. 2015). The second strategy involves purification of tissue-specific enriched messenger RNAs via the pull-down of poly-A binding proteins expressed in specific cell types (Blazie et al. 2015, 2017; Hrach et al. 2020; Mattout et al. 2011; Roy et al. 2002; Pauli et al. 2006) or the targeted expression of a modified splice leader trans-spliced to the transcripts by the cellular machinery (Ma et al. 2016). However, these techniques are labour intensive, require a large number of animals as starting material, or are associated with a strong unspecific background signal. Therefore, carrying cell-type specific transcriptional profiling in few differentiated cells remains challenging.

As a complementary approach that could circumvent some of these limitations, we investigated whether we could use DNA adenine methylation identification (DamID) to footprint actively transcribed genes in specific cell types. This approach has been successfully used in *Drosophila* to determine transcribed genes in rare brain cell types (Southall et al. 2013). DamID relies on the low-level expression of a fusion between the *E. coli* Dam methyltransferase and the protein of interest, here a subunit of the RNA polymerase. Binding of the latter to DNA leads to the methylation of GATC sites in the vicinity of the binding site (Figure 1A). After DNA extraction, methylated GATCs are specifically cleaved by the restriction enzyme *Dpn*I and amplified using adapter-mediated

PCR before sequencing (Figure S1A). The adaptation of the method to *C. elegans* seemed promising since endogenous adenine methylation/demethylation is very rare in worms and not targeted to GATC motifs (Greer et al. 2015). The GATC site frequency moreover allows to expect a good spatial resolution. Indeed, *C. elegans* has 269,049 GATC sequences per haploid genome , corresponding to an average of one site for every 374 bp and a median of 210 bp (Gómez-Saldivar, Meister, and Askjaer 2016).

Here, we describe the RAPID approach (RNA Polymerase DamID) for transcriptional footprinting in specific tissues in *C. elegans* in both embryonic blastomeres and in young adults. Using a fusion between a small subunit present in all 3 RNA polymerases, we show that the technique can be used on both fluorescently sorted blastomere cells, or DNA isolated from entire young adults using a Cre/*lox* cell-type-specific expression. To test the versatility of the method, we determined the polymerase footprints in three different tissues at each stage. In young adults, these tissues represent between 10 and 0.2% of the total cells of the animal. We show that meaningful transcriptional patterns can be recovered using this technique, in line with previously used cell-sorting and RNA tagging methods. We have further explored how this technique can also be used to unravel tissue-specific markers and report here the identification of eight new reporters for adult XXX cells. Additionally, as a result of the phasing-out of older sequencers able to sequence amplicons from different sizes as produced by DamID-seq, we show that new long read sequencing methods can be used to sequence DamID libraries, which allows carrying out the experiments, from DNA extraction to amplicon sequencing in less than a week without any external sequencing facility.

Material and methods

Plasmids and transgenic strains

All plasmids and worm strains are listed in Supplementary Table 1 and 2. Tissue-specific DamID plasmids were generated using Gibson assembly. All DamID plasmids were integrated as single copies using MosSCI either on chromosome II or IV (see details in Supplementary Table 1; Frøkjaer-Jensen et al. 2008). Strains and plasmids are available upon request.

Worm growth

Worms were grown on solid NGM, seeded with OP50 bacteria for maintenance culture and genetic crosses. All worm cultures were grown at 20°C. For cell sorting, animals were grown on peptone plates seeded with HB101, synchronized twice by hypochlorite treatment. Young adults containing up to 10 embryos were recovered 66 hours after spreading synchronized L1 larvae. For DamID experiments in young adults, worms were grown on NGM seeded with Dam negative *E. coli* GM48 for at least two generations. Around 5,000 synchronized L1s were seeded onto 100 mm plates (1,000-1,200 per plate) and collected 53h later. Worms were washed extensively with M9 (at least 10 times) and distributed in aliquots of 30 µl removing the excess of liquid. Samples were snap-frozen and stored at −80 °C.

Sorting of embryonic blastomeres

Synchronized gravid adult hermaphrodites containing 8-10 eggs were treated with hypochlorite. Eggs were then incubated for 3 hours in M9 at 25°C until they reached the 1.5 fold stage. Eggs were then transferred to 500 µl egg buffer and pelleted 1 min at 2000 rpm. The supernatant was then aspirated, leaving 100 µl of buffer with the pellet. 500 U of chitinase (Sigma-Aldrich;C8241-25U) were added and the mixture resuspended and further incubated for one hour at room temperature. Chitinase was then neutralized with 800 µl Leibovitz medium. Digested embryos were then recovered by centrifugation at 3000 rpm for 5 minutes at 4°C. Embryos were then dissociated into isolated blastomeres by pipetting up and down with a P1000 pipette, up to 150 times until dissociation was complete. The cell population was then filtered with a Millex-SV syringe 5 µm filter (Millipore;SLSV025LS). 500 fluorescent cells were sorted using a BD FACSAria Fusion (8000 events/second; 85µm nozzle) in sterile PCR tubes containing 1ul of pick buffer (50mM Tris-HCl pH 8.3; 75mM KCl; 3mM MgCl2; 137mM NaCl). Collected samples were then frozen in liquid nitrogen before further processing. After DamID processing and PCR (see below), two technical replicates were pooled for sequencing library preparation.

DamID amplification, library preparation, and sequencing

For sorted cells, frozen samples were lysed by addition of 2 ul of lysis buffer (10mM TrisAc, 10 mM MgAc, 50 mM KA, 0.67% Tween 20, 0.67% Igepal + 1mg/ml Proteinase K) and incubated during 2 hours at 60ºC before Proteinase K inactivation at 95ºC for 15 minutes. DamID amplicons were obtained as previously described (Gomez-Saldivar, Meister, and Askjaer 2016) using 30 PCR cycles.

For young adults, DamID was performed on 500 ng gDNA extracted from the animals using DNeasy Blood and Tissue Kit (Qiagen #69504). Two replicates for each stage and cell type were processed. DamID amplicons were obtained as previously described (Gomez-Saldivar, Meister, and Askjaer 2016) with 20 PCR cycles for worm-wide DamID, 22 PCR cycles for muscle DamID, 22-24 for intestine DamID and 24-26 cycles for XXX cells DamID. New Illumina patterned flow cell technology does not allow sequencing of amplicons larger than 600 bp. Both types of DamID PCR amplicons were therefore sequenced using nanopore sequencing. After AMPure XP (Beckman-Coulter) purification with 1.8 beads volume, DamID PCR amplicons were directly used for nanopore library barcoding and preparation using NBD-104 and LSK-109 kits (Oxford Nanopore). Libraries were sequenced to obtain at least 1 million reads per library on MinION using flow cells R9.4.1. We compared the performance of both sequencing approaches by sequencing the same libraries with both old paired-end HiSeq2500 Illumina flow cells and nanopore sequencers. Coverage obtained by both approaches were similar (see snapshot of mapped reads in Figure S1E). However, nanopore sequencing allowed the direct sequencing of longer amplicons (Figure S1F), likely because cluster amplification does not work well on molecules longer than 800 bp, even on old generation (HiSeq2500) flow cells. We compared DamID samples libraries generated using previously described entire animals expressing either a free GFP-Dam fusion or a perinuclear lamin fusion (Sharma et al. 2014). Statistical comparison of both techniques at the single restriction fragment level led to Pearson correlation coefficient between short PE and long reads between 0.84 and 0.86 (Figure S1G).

Bioinformatic analysis

Nanopore sequences were demultiplexed using deepbinner and basecalled using guppy 3.6 with the high accuracy model before mapping to the ce11 genome using minimap2 (Li 2018). Reads were considered as DamID amplicons when both ends mapped +/- 8 bp from a genomic GATC motif. Filtered libraries were then used to call polymerase footprinting values using the damidseq_pipeline package (O. J. Marshall and Brand 2015) using bam files (parameters: --bamfiles --extend_reads=0). The damidseq pipeline normalizes for GATC

accessibility using a free Dam (in our case Dam::GFP) profile. Gene values were extracted using polii.gene.call (https://owenjm.github.io/damidseq_pipeline/) using WS270 genes definition. Correlations between libraries and methylation footprinting values were made using *ad hoc* R scripts available upon request. Normalized footprinting tracks were visualized using IGV. For the comparison with ChIP-seq tracks, fastq files were downloaded from GEO, mapped to ce11, normalized for sequencing depth with a pseudocount of 8 and log2 normalized to the input. Final figure construction was made using IGV and Adobe Illustrator.

Array construction and microscopy

Reporter plasmids were assembled using the 3-fragment MultiSite Gateway system (ThermoFisher Scientific). Promoters of candidate genes were PCR amplified from a N2 gDNA template. Primer sequences and promoter lengths are described in Supplementary Table 3. Candidate promoters were cloned into pDONR-P4-P1R vector (Invitrogen) by BP recombinations generating slot 1 Entry vectors (Supplementary Table 2). Final expression plasmids were generated by LR recombinations between promoters (slot 1), mNeonGreen fluorescent ORF (slot 2; dg353; Hostettler et al. 2017), unc-54 3' UTR (slot 3; pMH473, gift from Marc Hammarlund) and pDEST-R4-P3. The control localization/expression plasmid [*sdf-9p::NLS::wrmScarlet*] was generated through a LR recombination reaction between dg801 (slot 1), dg651 (slot 2; Marques et al. 2019), pMH473 (slot 3), and pDEST-R4-P3. Finally, *promoter::mNeonGreen* and *sdf-9p::NLS::wrmScarlet* plasmids were co-injected into N2 young adult hermaphrodites at 20 ng/ul each, with 20 ng/ul of dg9 (*unc-122p::RFP*; red coelomocyte) as a co-injection marker (Addgene #8938; Miyabayashi et al. 1999). Two stable lines of each candidate promoter were selected (Supplementary Table 1). For imaging, young adult animals identified based on the red coelomocyte reporter expression were transferred to a 2% agarose pad containing 0.01% sodium azide. Worms were then imaged on a Zeiss Axiovert microscope with a 20x objective driven by Visiview, using a Photometrics Coolsnap Myo CCD camera with GFP and RFP settings (60 planes spaced by 2 µm), plus a middle plane section in DIC settings. Using Fiji, the acquired optical stacks were then partially z-projected to capture the worm section containing the XXX cells, identified using the *sdf-9p::NLS::wrmScarlet* signal. Final figure construction was made using Adobe Illustrator. For each promoter except *asp-9*, two independent extrachromosomal arrays were scored with at least 5 animals.

Identification of tissue-enriched genes and validation

Gene coverage of RAPID profile is listed in Supplementary Table 4 (blastomeres) and 5 (young adults) (Sheet 1-2) Tissue-enriched genes were defined as genes that are significantly expressed (FDR < 0.05) in both replicates of one tissue (Supplementary Table 4 sheet 2 and Supplementary Table 5 sheets 3-4). Representative examples of tissue-enriched genes used in Figure 2C were selected from CeNGEN project (https://cengen.shinyapps.io/SCeNGEA/) (Hammarlund et al. 2018) and visualized using IGV (https://igv.org/).

Genes enriched uniquely in one tissue were defined as genes significantly and differentially expressed relative to the other two tissues (Supplementary Table 4 sheet 3 and Supplementary Table 5 sheet 5). Analysis of the uniqueness of the genes from each tissue was performed using the JavaScript library jvenn (http://bioinfo.genotoul.fr/jvenn).

For the correct comparison between the genes identified in the muscle with RAPID and the set of genes identified with SRT (trans-Splicing Rna Tagging; Ma et al. 2016), only SL1 trans-spliced genes were selected, based on the intersection with the 10,589 mRNAs annotated as SL1-trans-spliced by modENCODE (Allen et al. 2011), generating a subset of 3477 genes. For the comparison between the genes identified in the intestine with RAPID and the set of genes identified with FANS (Fluorescence Activated Nuclear Sorting; Haenni et al. 2012), the gene list was modified eliminating the genes downregulated and keeping only the genes found as expressed and upregulated, in total 9169 genes. For the comparison between the genes identified with RAPID and the set of genes identified with FACS (Kaletsky et al. 2018) in muscle and intestine, we directly used the expressed genes list from Supplementary Table 1, which contained 7690 and 9603 genes, respectively. Chi-square ($\chi 2$) statistic tests were calculated using the GraphPad QuickCalcs Web site https://www.graphpad.com/quickcalcs/contingency1/ (accessed August, 2020).

Tissue-specific prediction analysis

Tissue expression prediction test was performed using Top 500 significant unique-enriched genes of muscle and intestine (Supplementary Table 5 sheet 6), selecting a multi-gene search within the Tissue-specific Expression Predictions for *C. elegans* program, version 1.0 (http://worm-tissue.princeton.edu/search/multi). For the tissue expression prediction test using the RAPID muscle-SL1 group, the parameters published in (Ma et al. 2016) were used.

Gene ontology analysis

Gene Ontology terms were performed on unique tissue-enriched gene lists. GO terms and q-value were obtained using gProfiler (https://biit.cs.ut.ee/gprofiler/gost) ; (version e100_eg47_p14_7733820) with g:SCS multiple testing correction method applying significance threshold of 0.05 (Supplementary Table 5 sheet 7; Raudvere et al. 2019).

Candidate selection for reporter analysis

Candidate genes were selected for follow-up promoter analyses with the dual goal of demonstrating that they are indeed transcribed in XXX cells and of identifying markers with expression restrained to XXX cells. Since XXX cells represent only 0.2% of the *C. elegans* cellular content, we reasoned that XXX-specific markers should produce very little or no signal in the whole animal samples. Therefore, out of the 862 XXX-enriched genes (with a significant RAPID signal in XXX, but not in muscle or in intestine samples), we further removed genes detected in at least one worm-wide sample. We obtained a refined list of 275 XXX-marker gene candidates, which still included the known XXX cell markers *daf-9* and *eak-4* (Supplementary Table 5, sheet 8). From this candidate list, we selected *dhs-17* as an uncharacterized gene with a plausible link to XXX cell function, plus 11 random candidates. In order to limit caveats related to operons, we excluded a candidate if another transcript was located less than 200 bp upstream. The XXX-specific RAPID signals (RNA pol occupancy values) in the 12-gene subset (average=1.35, SD=0.59) were similar to those in the starting 275-gene set (average=1.47, SD=0.64; p=.16 by Student's t-test).

Data availability

Embryonic and adult DamID-seq data is available under the GEO accession codes [data submitted to GEO, waiting for accession code].

Results and discussion

RNA polymerase footprinting using RAPID

The RAPID approach relies on the expression of trace levels of the *E. coli* DNA adenine methyltransferase (Dam) fused to a subunit of the RNA polymerases, and the analysis of its occupancy to evaluate the transcriptional state of genes. We first tested whether we could use AMA-1, the largest, catalytic subunit of RNA polymerase II, as reported in Targeted DamID, TaDa, done in neuronal lineages in *Drosophila* (Southall et al. 2013). However, the Dam signal was too weak, possibly due to the fact that the expression level of this very large fusion protein under transcriptional control of the uninduced *hsp-16.2* promoter was too low (not shown), or that the localization of the Dam domain within the *C. elegans* polymerase complex was not favorable to access its DNA substrate. We therefore replaced AMA-1 with RPB-6, the *C. elegans* homolog of the RNA polymerase subunit F present in all 3 RNA polymerases (personal communication M. Barkoulas; Katsanos et al., in preparation). RPB-6 together with RPB-1 and RPB-2, form a 'clamp' that retains DNA near the active center of PolII (Cramer et al. 2000), stabilizing the transcription on the DNA template. In contrast to AMA-1 attempts, we obtained a strong DamID PCR signal in animals carrying a single copy transgene of the *rpb-6::Dam* fusion, which had not been subjected to heat shock and therefore expressed only low levels of the Dam fusion in the whole animal. After sequencing and signal normalization with GFP::Dam data to control for overall chromatin accessibility, we could detect general patterns of methylation that were consistent with polymerase-dependent methylation (Figure 1C) and a good correlation at the gene level (Figure 1D). We compared the RAPID profiles with recent ChIP-seq datasets for RNA polymerase II downscaled to GATC fragment resolution (Garrido-Lecca, Saldi, and Blumenthal 2016; Miki, Carl, and Großhans 2017; Kalinava et al. 2017). Visually, the patterns appeared similar (see examples of profile in Figure 1C and D). Even if the nature of the data is not strictly identical, we performed a comparative analysis between the datasets for all genes, downscaling the resolution of the ChIP-seq data to GATC fragment (Figure 1B). RAPID and RNA polymerase ChiP-seq show a significant correlation (R=0.41-0.48). Overall, inter-replicate correlations were higher between the RAPID replicates (R=0.81) and between the RNA polymerase II ChIP experiments (R=0.63-0.97). The difference between the RAPID and RNA Pol II ChIP profiles most likely results from the different approach to recover DNA and from the fact that RAPID in adults represent a cumulative transcriptional activity picture, while the RNA pol II ChIP represents a polymerase occupancy snapshot at the young adult stage.

We also examined the profile in genes transcribed by Pol I and Pol III. As expected from the differential inclusion of AMA-1 and RPB-6 subunits in the different polymerases, RAPID with RPB-6 also labelled genes transcribed by RNA Pol I and III. RAPID showed high enrichments on the rRNA genes at the end of chromosome I (Figure S2A), as well as a majority of previously characterized snoRNAs transcribed by RNA polymerase III (46 out of 57; Ikegami and Lieb 2013). snoRNA are sometimes located in introns of PolII-transcribed genes, making it difficult to differentiate whether the RAPID signal originates from the snoRNA transcription or from the overlapping gene. Nevertheless, 24 out of 44 RAPID-positive snoRNA genes could be unambiguously assigned to PolIII transcription as there was no overlap (Figure S2B). In contrast, the corresponding signal was markedly weaker in the young adult RNA pol II ChIP-seq datasets and very high for the RNA PolIII subunit RPC-1 (performed in embryos, Figure S2B; Ikegami and Lieb 2013). Taken together, these data show that the RAPID method is suitable to reveal RNA polymerase footprints, serving as indirect indications for transcriptional activity made by the three different RNA polymerases.

RAPID in tissue-specific sorted embryonic cells.

Since RAPID could label transcribed genes in whole animals, our next goal was to adapt the method for tissue-specific analysis. We first explored whether RAPID could be used to footprint transcription in specific cell types of the embryo purified using fluorescent sorting. Using strains constitutively expressing Dam fusions, we sorted 1000 cells for 3 different tissues of various abundance: body wall muscle, intestine and the rectal Y cell (80, 20 and 1 cell per embryo, respectively; Figure S3A, Supplementary Table 4). RAPID could produce footprints for all samples (Figure S3C), and the reproducibility of the RAPID signal at the gene level between replicas was good for the intestinal sample (*r=0.78*) yet lower for muscle and Y cell samples (*r=0.46* and *0.24*, respectively; Figure S4A). RAPID appears however suitable to identify transcription footprints of cell type specific genes, such as the intestinal *cpr-6* (cathepsin B; Pauli et al. 2006) and *elo-2* (fatty acid elongase ELOVL; Han et al. 2017), or Glutathione S-Transferase *gst-4* expressed in muscle, as well as the muscle-specific WDR1 homolog *unc-78* (Hasegawa et al. 2007; Figure S3C; Mohri and Ono 2003). Comparison of the footprints obtained for all three tissues allowed us to define 395 Y-specific genes (Figure S3B), and among them, shallow Y-specific RAPID signal was visible for *ceh-6* and *sox-2*, two transcription factors previously shown to be involved in Y-to-PDA transdifferentiation initiation (Figure S3C; Kagias et al. 2012). In contrast, we observed some unexpected RAPID footprints for genes which we expected to be expressed in a

tissue-specific manner. For example, the myosin *myo-3* was observed in all 3 tissues (Figure S3D). Conversely, some genes which are known to be expressed in tissue-specific manner did not show high RAPID values, *e.g.* the intestine-specific *asp-1* and *spp-5* genes or the muscle-specific *pat-10* one (Figure S3D). These differences might have a biological significance: most transcription patterns have been defined in older animals, yet those genes might be expressed more broadly, or not at all, in embryonic blastomeres. Alternatively, the lack of a robustly established transcription pattern might hinder faithful footprinting. Indeed, RAPID requires the enzymatic action of the methyltransferase to build up the signal, while footprints are removed at each replication round in rapidly dividing blastomeres. Together with the stochasticity of single methylation events, this might decrease the precision of the profiles. In agreement with this, intestinal cells, most of which are born and remain postmitotic early, display high reproducibility in the RAPID signal, while late-born muscle and the Y cell are less reproducible (Figure S3C,D). In conclusion, RAPID allows transcriptional footprinting of sorted cells, provided cells have been born long enough for the Dam fusions to methylate reproducibly their target genes.

Tissue-specific expression of RAPID in young adult animals.

Our next goal was to adapt the RAPID approach for the analysis of tissue-specific transcription profiles in adult worms. To circumvent the technical difficulties associated with mature animal dissociation and cell sorting, we sought to implement a transgenic approach for the cell-specific targeting of the RPB-6::Dam fusion. Several approaches have been developed to target Dam fusion expression to specific tissues, using either bicistronic constructs attenuating Dam translation (Southall et al. 2013), tamoxifen-induced nuclear translocation (Pindyurin et al. 2016; Redolfi et al. 2019) or tissue-specific recombination cassettes (Muñoz-Jiménez et al. 2017). We used the latter strategy using an insertion of a floxed mCherry ORF located between the *hsp-16.2* promoter and the Dam fusion (Figure 2A). Tissue-specific expression of the Cre recombinase leads to low-level expression of the Dam fusion (Ruijtenberg and van den Heuvel 2015). We used 3 different Cre drivers expressed in muscle, intestine and a pair of head neuroendocrine cells called XXX, using *myo-3*, *elt-2* and *sdf-9* promoters, respectively. In contrast to previous strategies used in the nematode to determine cell-type specific transcription, which require tens of thousands of worms, RAPID requires only 4’000 animals as starting material for DNA extraction. DamID signal across genes was highly correlated between replicates, and less correlated between different Cre drivers, suggesting tissue-specific DamID footprinting (Figure S4B). We called enriched expressed genes with an FDR of <0.05 for both replicates of each tissue, resulting

in 2052, 2296 and 2362 genes with high RAPID footprint in intestine, muscle and XXX cells, respectively. Out of these, 1011, 747 and 862 were defined as tissue-specific in these three cell types, while only a minority (536) harbored high RAPID footprints in all sampled tissues (Figure 2B, Supplementary Table 5).

We validated these results focusing on individual genes with well-characterized tissue-specific expression (Figure 2C). The intestine-specific prograstricsin aspartyl protease homolog *asp-1* and the saposin-like channel defense *spp-5* were exclusively footprinted in intestinal samples (Roeder et al. 2010; Tcherepanova et al. 2000). *myo-3*, the muscle myosin and *pat-10*, a troponin, both exclusively expressed in muscle tissues, harbored high RAPID footprints in muscle samples (Miller, Stockdale, and Karn 1986; Kagawa, Takuwa, and Sakube 1997). Finally, the XXX-specific genes *daf-9* (encoding a P450 cytochrome homolog) and *eak-4* (encoding a membrane protein) were exclusively marked with RAPID in XXX cells (Gerisch et al. 2001; Hu, Xu, and Ruvkun 2006).

To expand the validation of our RAPID footprint gene lists beyond selected tissular markers, we next engaged two more comprehensive approaches. These approaches focused on muscle and intestine, since no large-scale gene expression data are available in XXX cells. First, we compared the sets of genes with detectable RAPID footprints in body wall muscle and intestine, with sets of genes detected in these tissues using genome-wide approaches (RNA-seq of FACS-sorted cells by Kaletsky et al. 2018) ; of SRT messengers (Ma et al., 2016); or of FACS-sorted nuclei (Haenni et al., 2012). A majority of genes with detectable RAPID footprints were also identified with tissue-specific transcriptomics (Figure 3A,B, 77% for muscle, 72% for intestine). Second, we focused on a narrower set of genes with high RAPID footprint present exclusively in one of the tissues. These gene lists were analyzed with the tissue predictor tool with a multi-gene query (http://worm-tissue.princeton.edu/; Chikina et al. 2009). Muscle-only and intestine-only RAPID footprinted genes gave a high prediction score for muscle and intestine, respectively (Figure 3C,D). In contrast, their scores were much lower for non-targeted tissues (including neurons, hypodermis and germline).

Taken together, the results of our comparative analyses with previous gene expression studies suggest a valid identification of tissue-specific polymerase transcriptional activity with RAPID that largely matches the transcriptional output in these tissues. RAPID thereby allows reproducible profiling of transcription footprints in single tissues using low input material.

RAPID identifies genes expressed in XXX cells

XXX cells are neuroendocrine cells derived from hypodermal embryonic progenitors. They express enzymes required for steroid hormone synthesis such as DAF-9 and are implicated in the control of dauer formation during development (Ohkura et al. 2003; Gerisch et al. 2001). As we still have limited information on the set of genes expressed in XXX cells, as well as on the role of those cells in adults, our XXX-cell transcribed gene dataset could represent a valuable resource for future studies. We thus conducted additional experiments to confirm the validity of RAPID profiles to infer gene transcription in those cells and identify new markers.

A GO term analysis highlighted many terms related to secretory and neural functions in the gene set specific for XXX cells, while they were absent from the muscle- and intestine-enriched gene sets (Supplementary Table 5 sheet 7). These results are expected for neuroendocrine cells, with functional properties very similar to those of neurons.

To further confirm the validity of RAPID profiles to infer gene transcription, we selected 12 genes with high RAPID signal in XXX cells compared to intestine, muscle and whole animals, by comparing their profiles in the 4 tissues (Figure S5). We then created transcriptional reporters by cloning their putative promoters in front of a mNeonGreen fluorescent reporter (Hostettler et al. 2017). Transgenic lines bearing extrachromosomal arrays were then created by injection, together with an XXX-specific red reporter driven by the *sdf-9* promoter, which we previously used for XXX-specific RAPID. Out of the 12 promoters, eight drove expression in the XXX cells at various levels (*asic-1, dhs-17, F14H8.2, C06A1.3, asp-9, B0034.5, F52E1.5,* and *nlp-15*, Figure 4), one was impossible to score (*T22E5.6*), as it also drove high expression in the pharynx located very close to the XXX cells, and three did not show detectable expression in XXX cells (in adults; *mab-9, twk-39* and *lgc-36*). The absence of expression in the latter group could indicate that these genes represent false positives in the RAPID analysis, that the transcription was active only at earlier developmental timepoints, or that the arbitrarily defined promoters failed to reflect the endogenous transcriptional activity. Interestingly, in half of the reporters expressed in XXX cells (4/8), we observed only limited or no signal in other cell types. Reporters for *dhs-17* or *F14H8.2* were predominantly expressed in XXX cells and sequence homology suggests plausible cell-specific roles. Indeed, *dhs-17* encodes an uncharacterized member of the short-chain dehydrogenases/reductases family (SDR), which is potentially involved in steroid metabolism (Y. Zhang et al. 2013) and may therefore contribute to this specific neuroendocrine function of XXX cells. Our RAPID analysis also identified *F14H8.2*, a

paralog of *eak-4* selectively expressed in XXX cells and regulating dauer formation (Hu, Xu, and Ruvkun 2006).

Taken together, the results of our reporter gene analysis indicate that the RAPID approach applied to only two cells per animal successfully identifies actively transcribed genes in these cells, including genes with relatively low expression and potential novel cell-specific markers.

Limitations and perspectives

Determining the transcriptional profile of specific cell types has been a major hurdle for *C. elegans* researchers, as most research is carried out on intact, living animals. In addition, research on specific cell types, which can represent as little as one cell in the whole animal, is often impaired by the paucity of markers for cell identity and cell function. With RAPID, we present a simple, fast and cost-effective technique to identify transcribed genes by the footprint of DNA-engaged polymerases in individual tissues or cells, requiring only the knowledge of a tissue- or cell-specific promoter. RAPID involves the enzymatic modification of DNA, a major advantage as only cells in which the Dam fusion is expressed will produce the DamID signal. Cell-type-specific RAPID hence makes use of DNA extracted from entire animals as a template for DamID, thereby avoiding cell or RNA purification steps and reducing the size of the starting population to a few thousands animals instead of tens of thousands of animals required for cell or RNA purification. However, RAPID is a stochastic average of the methylation by polymerases since the cell's last DNA replication (which erases the DamID signal). Unsurprisingly, we observed a better signal to noise ratio in older animals in which the transcriptional program has been established longer, thereby reducing the stochasticity of the methylations (Figure S4).

RAPID could be further improved by timing the expression of the Dam fusions, for instance using degron-tagged degradation (L. Zhang et al. 2015). We tested auxin-mediated Dam degradation; however, degradation is not complete and the remaining low levels of Dam fusions are sufficient to create a RAPID signal (data not shown). Alternatively, degron-tagged Cre recombinases might provide a more reliable way to time the expression of the Dam. In addition to static tissue-specific transcriptional profiles, a potential future application of RAPID would be to examine dynamic transcriptional changes occuring in response to environmental changes, as successfully achieved in *Drosophila* (Widmer et al. 2018). In a post-mitotic context, the approach would be limited to transcriptional up-regulation events since, in the absence of cell-replication-independent demethylation mechanisms, reduction of polymerase occupancy would not be reflected in a decreased Dam signal. Beyond transcription, cell-type-specific DamID similar to RAPID can be applied

to a variety of chromatin factors to delineate cell-type specific genome-wide binding profiles in *C. elegans*, ranging from histone marks (Villaseñor et al. 2020; Owen J. Marshall and Brand 2017) to transcription factors (Filion et al. 2010; Schuster et al. 2010) or chromatin accessibility (Aughey et al. 2018).

In conclusion, our work indicates that RAPID is a valid approach to analyse transcriptional activity in specific *C. elegans* cell types. Because of its ease of implementation and relatively high sensitivity, we anticipate that it will represent a powerful complement to cell-sorting-based approaches, particularly suited for situations where the amount of starting material is limiting or the fraction of targeted cells is low.

Acknowledgments

We are grateful to Laurence Bulliard, Cihan Enci, Lisa Schild and Jennifer Semple for expert technical support; Marc Hammarlund, Piali Sengupta and Sanders van der Heuvel for the gift of material; Michalis Barkoulas for sharing preliminary data; some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). The study was supported by the Swiss National Science Foundation (IZCNZ0_174703/SBFI_C16.0013, BSSGI0_155764 and PP00P3_150681 to D.A.G.; ZCSZ0-174641/SBFI_C15.0076, PP00P3_159320 and 31003A_176226 to PM), the ERC CoG PlastiCell grant #648960 to SJ and a BMBS COST Action (BM1408).

Figure legends

Figure 1. RAPID: RNA Pol II DamID tissue-specific scheme.

A. *Dam::rpb-6* bound to all three RNA polymerase types modifies proximal GATC motifs. DamID analysis provides RNA polymerase footprints *in vivo*.

B. Gene-level correlations in young adults whole animals of RAPID and RNA polymerase Chromatin immunoprecipitation-seq using different antibodies (a: (Garrido-Lecca, Saldi, and Blumenthal 2016); b: (Kalinava et al. 2017); c: (Miki, Carl, and Großhans 2017)).

C, D. Profiles at large scale (C) and gene scale (housekeeping genes, D) of RAPID and the different RNA polymerase II ChIP-seq studies cited in B.

Figure 2. Tissue-specific expression profiles using RAPID by Cre/*lox* recombination in young adult animals.

A. Experimental system for tissue-specific expression of Dam fusions and RAPID analysis.

B. Venn diagram of overlap between expressed genes identified using RAPID in 3 different tissues.

C. RAPID profiles for previously characterized genes expressed in a tissue-specific manner in intestine, muscle, and the XXX neuroendocrine pair.

Figure 3. Validation of RAPID muscular and intestinal transcriptomes.

A. Overlap of genes footprinted using RAPID with muscle-expressed genes detected by FACS (left; (Kaletsky et al. 2018) and SRT (right; Ma et al. 2016). For the comparison to SRT data, the subset of SL1-trans spliced genes detected using RAPID was used.

B. Overlap of genes footprinted using RAPID with intestine-expressed genes detected by FACS (Kaletsky et al. 2018) and FANS (Haenni et al. 2012). χ^2 test for all comparisons, p-value < 0.0001.

C-D. Tissue enrichment prediction of tissue-specific RAPID transcriptomes. The top 500 hits of unique tissue-enriched genes for muscle (C) and intestine (D) were analyzed by the online program Tissue-specific Expression Predictions for *C. elegans* (http://worm-tissue.princeton.edu/search). Tissue enrichment scores (mean and SEM) showed that the top genes of each tissue identify their corresponding original tissue.

Figure 4. Promoters of genes identified as expressed in XXX cells using RAPID drive fluorescent reporter expression in XXX cells. Left panel, partial z-projections of adult heads showing expression of *sdf-9p::mScarlet* in XXX cells. Middle panel, partial z-projections of adult heads showing the expression pattern of *mNeonGreen* reporter under transcriptional

control of the promoters of indicated genes. Right panel, merge of fluorescent channels within DIC images validate the novel expression of mentioned genes in XXX cells. Bar: 10 µm.

Supplementary figures

Figure S1. DamID Sequencing using Illumina or Nanopore technologies.

A. Principle of methylated GATC identification using DamID. Methylated GATC motifs are cut using the methylation-sensitive *Dpn*I restriction enzyme, before **B.** adapter ligation and PCR amplification using a primer hybridizing in the adapter. **C.** For Illumina sequencing, classical DNA library preparation is used, involving Y adapter ligation and limited PCR amplification. DamID amplicons can only be sequenced on non-patterned Illumina flow cells as their variable length is detrimental to cluster formation and polony amplification. To solve the latter issue, amplicons from the first PCR round were directly sequenced using ligation-mediated library preparation and ONT long read sequencing. **D.** While Illumina sequencing highlights the beginning of the amplicons (and the end in case of paired end sequencing), ONT long read sequencing captures the entire length of the amplicon. **E.** Snapshot of the same library sequenced with Illumina short read sequencing (paired end, upper panel) and ONT long read sequencing (bottom panel). **F.** Comparison of amplicons length obtained from both sequencing techniques for the same library (blue, Illumina; orange, ONT). The smaller sizes of Illumina-sequenced amplicons is due to the loss of these amplicons during polony amplification in Illumina sequencing, as the average size of the amplicons determined using automated electrophoresis before loading onto the flow cell correspond to the ONT-sequenced one (data now shown). **G.** Pearson correlation at the single GATC fragment resolution for 4 different libraries of 2 different Dam fusions (with either GFP or LMN-1) sequenced with either Illumina (short PE) or ONT (Long) sequencers.

Figure S2. RAPID allows identification of PolI and PolIII transcribed genes.

A. RAPID footprinting of the rDNA locus on the right telomere of chromosome I, together with RNA polymerase II ChIP-seq (8WG16 ChIP-seq from young adult animals from (Miki, Carl, and Großhans 2017). **B.** RAPID signal at polIII snoRNA genes, with no PolII overlapping signal (24 out of 57 genes; 8WG16 ChIP-seq as in A). PolIII-specific genes were taken from (Ikegami and Lieb 2013), ChIP-seq for RPC-1 (RNA polymerase III subunit A) performed in embryos. The 8WG16 antibody produces a low residual signal on the snoRNA genes. **C.** RAPID signal at polIII snoRNA genes with overlapping PolII ChIP-seq signal (20

out of 57 genes; data as in A and B). **D.** snoRNA genes with no RAPID signal (13 out of 57 genes; data as in A and B). Classifications were done twice independently by visual inspection of the tracks.

Figure S3. RAPID using sorted embryonic blastomeres

A. Embryonic blastomeres from a strain expressing constitutively Dam constructs (either fused to GFP or RPB-6) are sorted based on the expression of fluorescent markers using fluorescence-activated cell sorting. For each construct and each tissue, 1'000 cells are pooled before performing DamID. **B.** Venn diagram showing overlap between RAPID footprinted genes (FDR > 0.05). **C, D.** Individual profiles for genes specific for the different tissues tested (intestine, muscle and Y cell).

Figure S4. Pearson correlation coefficient between RAPID libraries at the single gene level resolution for the different tissues tested.

A. Correlation of RAPID footprinting signal between replicates performed in different tissues in embryonic blastomeres sorted using fluorescent markers. **B.** Correlation of RAPID footprinting signal between replicates performed in different tissues using Cre-mediated recombination for tissue-specific expression of the Dam fusions.

Figure S5. RAPID profile of candidate genes with XXX-specific expression.

RAPID footprinting of candidate uncharacterized genes expressed in a tissue-specific manner in XXX cells but not in muscle, intestine and enriched worm-wide are shown. Candidate XXX neuronal genes are ordered according to the RAPID footprinting level.

Supplementary information

Supplementary Table 1

Worm strains collection

Strain name	Source	Genotype
BN195	Towbin 2012	*bqSi195[hsp-16.41p:: dam::myc::lmn-1); unc-119(+)] II*
BN196	Towbin 2012	*bqSi196[hsp-16.41p::gfp::myc::dam 3'UTR; unc-119(+)] II*
IS3675	This study	*fpSi2[hsp-16.2p::lox::degron::dam::rpb-6::unc-54 3'UTR; unc-119(+)] II*
MS438	CGC	*Irls25[elt-2p::nls::gfp::lacz; rol-6(su1006)] V*
PD4251	CGC	*ccIs4251[myo-3p::gfp::lacz::nls;myo-3p::mito gfp; dpy-20(+)] I*
IS2922	This study	*fpIs76[egl-5p(6.5kb)::Δpes-10::mcherry; Bluescript] I* *fpIs92[hlh-16p::GFP::hlh-16(exon1)::hlh-16 3'UTR;pRF4]*
IS3467	This study	*fpIs76[egl-5p(6.5kb)::Δpes-10::mcherry; Bluescript] I* *fpIs92[hlh-16p::GFP::hlh-16(exon 1))::hlh-163'UTR;pRF4]* *bqSi196[hsp-16.41p::gfp::myc::dam; unc-119(+)] II*
IS3685	This study	*fpIs76[egl-5p(6.5kb)::Δpes-10::mcherry; Bluescript] I* *fpIs92[hlh-16p::GFP::hlh-16(exon 1))::hlh-163'UTR;pRF4]* *fpSi2[hsp-16.2p::lox::degron::dam::rpb-6::unc-54 3'UTR; unc-119(+)] II*
IS3651	This study	*ccIs4251[myo-3p::gfp::lacz::nls;myo-3p::mito gfp; dpy-20(+)] I* *bqSi196[hsp-16.41p::gfp::myc::dam; unc-119(+)] II*
IS3684	This study	*ccIs4251[myo-3p::gfp::lacz::nls;myo-3p::mito gfp; dpy-20(+)] I* *fpSi2[hsp-16.2p::lox::degron::dam::rpb-6::unc-54 3'UTR; unc-119(+)] II*

IS3653	This study	*bqSi196[hsp-16.41p::gfp::myc::dam; unc-119(+)] II* *iirIs25[elt-2p::nls::GFP::lacZ + rol-6(su1006)] V*
IS3687	This study	*fpSi2[hsp-16.2::lox::degron::dam::rpb-6::unc-54 3'UTR; unc-119(+)] II* *irIs25[elt-2p::nls::gfp::lacZ + rol-6(su1006)] V*
SV1439	CGC	*heSi142[elt-2p::elg-13nls::Cre::tbb-2 3'UTR] X*
SV1461	CGC	*heSi148[myo-3p::elg-13nls::Cre::tbb-2 3'UTR] X*
DAG827	This study	*domSi827[sdf-9p::Cre::tbb-2 3'UTR; unc-119(+)] X*
PMW641	This study	*ubsSi40[hsp-16.2::lox:mcherry::lox::degron::gfp::dam::unc-54 3'UTR; unc-119(+)] IV*
DAG958	This study	*domSi958[hsp-16.2p::lox::mCherry::lox::degron::dam::rpb-6::unc54 3'UTR; unc-119(+)] II*
IS3568	This study	*heSi142[elt-2p::elg-13nls::Cre::tbb-2 3'UTR] X;* *domSi958[unc-119(+)hsp-16.2p::lox::mCherry::lox::degron::dam::rpb-6::unc54 3'UTR] II*
PMW748	This study	*heSi148[myo-3p::elg-13nls::Cre::tbb-2 3'UTR] X ;* *domSi958[hsp-16.2p::lox::mCherry::lox::degron::dam::rpb-6::unc54 3'UTR; unc-119(+)] II*
DAG879	This study	*domSi827[sdf-9p::Cre::tbb-2 3'UTR; unc-119(+)] X* *ubsSi40[hsp-16.41::lox:mcherry::lox::degron::gfp::dam::unc-54 3'UTR; unc-119(+)] IV*
DAG976	This study	*heSi148[myo-3p::elg-13nls::Cre::tbb-2 3'UTR] ttTi14042 X* *domSi958[hsp-16.2p::lox::mCherry::lox::degron::dam::rpb-6::unc54 3'UTR; unc-119(+)] II*
DAG1042	This study	*domSi827[sdf-9p::Cre::tbb-2 3'UTR; unc-119(+)] X* *domSi958[hsp-16.2p::lox::mCherry::lox::degron::dam::rpb-6::unc54 3'UTR; unc-119(+)] II*

DAG1043	This study	*heSi142[elt-2p::elg-13nls::Cre::tbb-2 3'UTR] X ubsSi40[hsp-16.41::lox:mcherry::lox::degron::gfp::dam::unc-54 3'UTR; unc-119(+)] IV*
DAG1256-1257	This study	*domEx1256-1257[asic-1p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1258-1259	This study	*domEx1258-1259[dhs-17p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1260-1261	This study	*domEx1260-1261[mab-9p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1262-1263	This study	*domEx1262-1263[twk-39p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1264-1265	This study	*domEx1264-1265[lgc-36p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1266-1267	This study	*domEx1266-1267[F14H8.2p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1268-1269	This study	*domEx1268-1269[C06A1.3p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1270	This study	*domEx1270[asp-9p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1271-1272	This study	*domEx1271-1272[B0034.5p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1273-1274	This study	*domEx1273-1274[F52E1.5::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
DAG1275-1276	This study	*domEx1275-1276[T22E5.6p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*

DAG1277-1278	This study	*domEx1277-1278[nlp-15p::mNeonGreen::unc-54 3' UTR, sdf-9p::egl-13NLS::wrmScarlet::unc-54 UTR, unc-122p::RFP]*
N2	CGC	Wild type Bristol N2

Supplementary Table 2

Plasmid collection

Plasmid name	Source	Plasmid information	Backbone
#372	This study	*hsp-16.41p::degron::dam::ama-1::unc-54 3'UTR*	pCFJ151
#509	This study	*hsp-16.2p::degron::dam::rpb-6::unc54 3'UTR*	pCFJ151
#417	This study	*hsp-16.2p::lox::mCherry::lox::degron::dam::ama-1::unc-54 3' UTR*	pCFJ151
#418	This study	*hsp-16.2p::lox::mCherry::lox::degron::gfp::dam::unc-54 3' UTR*	pCFJ151
#434	This study	*hsp-16.2p::lox::mCherry::lox::degron::dam::rbp-2::unc-54*	pCFJ151
pSJ4646/ #503	This study	*hsp-16.2p::lox::mCherry::lox::degron::dam::rpb-6::unc-54 3'UTR*	pCFJ151
pSR32	Sander Van den Heuvel Lab	*pie-1p::Cre*	pCFJ355
dg9	Miyabayashi 1999	*unc-122p::rfp*	
dg353	Hostettler 2017	*mNeonGreen* Entry (slot2)	pDON221
dg651	Marques 2019	*egl-13nls::wrmScarlet* Entry (slot2)	pDON221
dg800	This study	*sdf-9p::Cre::tbb-2 3'UTR*	pCFJ355
dg801	This study	*sdf-9p* Entry (slot1)	pDONR-P4-P1R
dg802	This study	*asic-1p* Entry (slot1)	pDONR-P4-P1R

dg803	This study	*dhs-17p* Entry (slot1)	pDONR-P4-P1R
dg804	This study	*mab-9p* Entry (slot1)	pDONR-P4-P1R
dg805	This study	*twk-39p* Entry (slot1)	pDONR-P4-P1R
dg806	This study	*lgc-36p* Entry (slot1)	pDONR-P4-P1R
dg807	This study	*F14H8.2p* Entry (slot1)	pDONR-P4-P1R
dg808	This study	*C06A1.3p* Entry (slot1)	pDONR-P4-P1R
dg809	This study	*asp-9p* Entry (slot1)	pDONR-P4-P1R
dg810	This study	*B0034.5p* Entry (slot1)	pDONR-P4-P1R
dg811	This study	*F52E1.5p* Entry (slot1)	pDONR-P4-P1R
dg812	This study	*T22E5.6p* Entry (slot1)	pDONR-P4-P1R
dg813	This study	*nlp-15p* Entry (slot1)	pDONR-P4-P1R
dg814	This study	*sdf-9p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg815	This study	*sdf-9p::egl-13NLS::wrmScarlet::unc-54 3' UTR*	pDEST-R4-P3
dg816	This study	*asic-1p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg817	This study	*dhs-17p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg818	This study	*mab-9p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3

dg819	This study	*twk-39p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg820	This study	*lgc-36p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg821	This study	*F14H8.2p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg822	This study	*C06A1.3p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg823	This study	*asp-9p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg824	This study	*B0034.5p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg825	This study	*F52E1.5p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg826	This study	*T22E5.6p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
dg827	This study	*nlp-15p::mNeonGreen::unc-54 3' UTR*	pDEST-R4-P3
pMH473	Gift from Marc Hammarlund	*unc-54 3' UTR* Entry (slot3)	pDONR- P2R-P3

Supplementary Table 3

Primers for cloning of promoters used for the transcriptional reporter analysis of XXX cell-transcribed candidate genes.

All primers start with Gateway cloning sequences, followed with the indicated sequence. (forward: `ggggacaactttgtatagaaaagtt`; reverse: `ggggactgcttttttgtacaaacttg`)

Gene name	Forward primer	Reverse primer	Size (bp)
sdf-9	*gtcgacttgtcaatgtcgc*	*ccttcattttgaaaataatatatctagtcttag*	3497
asic-1	*gacacacttgtacgggatacctt*	*tttggtggcctgaaatttgatcaa*	1992
dhs-17	*gataacacattgaccctctccac*	*ttctacaaattacagacgagacaactg*	1969
mab-9	*gccatggaacccgaatttttgcat*	*cttatggaataagaagctccggga*	2026
twk-39	*ttgtcgggtttgtcagtttcct*	*tttcagtcgtctatttgactgcca*	2006
lgc-36	*gcatcaaggaactgctttcaca*	*tggaaattcgatgaagcatttgtgtg*	560
F14H8.2	*agagttcaaaaaggtgttcccaa*	*agagttcaaaaaggtgttcccaa*	1779
C06A1.3	*ttgtctggcgtgtatttggtac*	*atgaagacgttgctaaatcctctttc*	660
asp-9	*tgacctattttgaactgctg*	*tcagctcaactatcaaaatttaatcac*	1668
B0034.5	*taatgcgtgataatgatgagac*	*taatgcgtgataatgatgagac*	1992
F52E1.5	*tccaattttatgcgttgagt*	*ggcaggttatttaaacggt*	1580
T22E5.6	*tgcaagttaatttttcgttc*	*tgtgaaaagctgaaatacag*	2037
nlp-15	*tggtcaagtacataccaggggtatg*	*cgttatgttgtgtctctttcg*	2186

Supplementary Table 4 (excel sheet)

Averaged gene-level values for RNA polymerase signal (polii), GATC number and false discovery rate (FDR) for all blastomere libraries.

Supplementary Table 5 (excel sheet)

Transcribed genes analysis determined by RAPID. Genes for each tissue are listed.

Enriched and uniques genes, and Gene ontology attributes are tabulated in different sheets.

References

Allen, Mary Ann, Ladeana W. Hillier, Robert H. Waterston, and Thomas Blumenthal. 2011. "A Global Analysis of C. Elegans Trans-Splicing." *Genome Research* 21 (2): 255–64.

Aughey, Gabriel N., Alicia Estacio Gomez, Jamie Thomson, Hang Yin, and Tony D. Southall. 2018. "CATaDa Reveals Global Remodelling of Chromatin Accessibility during Stem Cell Differentiation in Vivo." *eLife* 7 (February). https://doi.org/10.7554/eLife.32341.

Blazie, Stephen M., Cody Babb, Henry Wilky, Alan Rawls, Jin G. Park, and Marco Mangone. 2015. "Comparative RNA-Seq Analysis Reveals Pervasive Tissue-Specific Alternative Polyadenylation in Caenorhabditis Elegans Intestine and Muscles." *BMC Biology* 13 (January): 4.

Blazie, Stephen M., Heather C. Geissel, Henry Wilky, Rajan Joshi, Jason Newbern, and Marco Mangone. 2017. "Alternative Polyadenylation Directs Tissue-Specific miRNA Targeting inCaenorhabditis elegansSomatic Tissues." *Genetics*. https://doi.org/10.1534/genetics.116.196774.

Chalfie, M., Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher. 1994. "Green Fluorescent Protein as a Marker for Gene Expression." *Science* 263 (5148): 802–5.

Chikina, Maria D., Curtis Huttenhower, Coleen T. Murphy, and Olga G. Troyanskaya. 2009. "Global Prediction of Tissue-Specific Gene Expression and Context-Dependent Gene Networks in Caenorhabditis Elegans." *PLoS Computational Biology* 5 (6): e1000417.

Cramer, P., D. A. Bushnell, J. Fu, A. L. Gnatt, B. Maier-Davis, N. E. Thompson, R. R. Burgess, A. M. Edwards, P. R. David, and R. D. Kornberg. 2000. "Architecture of RNA Polymerase II and Implications for the Transcription Mechanism." *Science* 288 (5466): 640–49.

Filion, G. J., J. G. van Bemmel, U. Braunschweig, W. Talhout, J. Kind, L. D. Ward, W. Brugman, et al. 2010. "Systematic Protein Location Mapping Reveals Five Principal Chromatin Types in Drosophila Cells." *Cell* 143 (2): 212–24.

Frøkjaer-Jensen, Christian, M. Wayne Davis, Christopher E. Hopkins, Blake J. Newman, Jason M. Thummel, Søren-Peter Olesen, Morten Grunnet, and Erik M. Jorgensen. 2008. "Single-Copy Insertion of Transgenes in Caenorhabditis Elegans." *Nature Genetics* 40 (11): 1375–83.

Garrido-Lecca, Alfonso, Tassa Saldi, and Thomas Blumenthal. 2016. "Localization of RNAPII and 3' End Formation Factor CstF Subunits on C. Elegans Genes and Operons." *Transcription* 7 (3): 96–110.

Gerisch, B., C. Weitzel, C. Kober-Eisermann, V. Rottiers, and A. Antebi. 2001. "A Hormonal Signaling Pathway Influencing C. Elegans Metabolism, Reproductive Development, and Life Span." *Developmental Cell* 1 (6): 841–51.

Gómez-Saldivar, Georgina, Peter Meister, and Peter Askjaer. 2016. "DamID Analysis of Nuclear Organization in Caenorhabditis Elegans." *Methods in Molecular Biology* 1411: 341–58.

Gomez-Saldivar, G., P. Meister, and P. Askjaer. 2016. "DamID of Nuclear Envelope Proteins in C. Elegans." In *The Nuclear Envelope, Methods and Protocols*, edited by S. Shackleton, P. Collas, and E. C. Schirmer. Springer.

Greer, Eric Lieberman, Mario Andres Blanco, Lei Gu, Erdem Sendinc, Jianzhao Liu, David Aristizábal-Corrales, Chih-Hung Hsu, L. Aravind, Chuan He, and Yang Shi. 2015. "DNA Methylation on N6-Adenine in C. Elegans." *Cell* 161 (4): 868–78.

Haenni, S., Z. Ji, M. Hoque, N. Rust, H. Sharpe, R. Eberhard, C. Browne, et al. 2012. "Analysis of C. Elegans Intestinal Gene Expression and Polyadenylation by

Fluorescence-Activated Nuclei Sorting and 3'-End-Seq." *Nucleic Acids Research* 40 (13): 6304–18.
Hammarlund, Marc, Oliver Hobert, David M. Miller, and Nenad Sestan. 2018. "The CeNGEN Project: The Complete Gene Expression Map of an Entire Nervous System." *Neuron*. https://doi.org/10.1016/j.neuron.2018.07.042.
Han, Shuo, Elizabeth A. Schroeder, Carlos G. Silva-García, Katja Hebestreit, William B. Mair, and Anne Brunet. 2017. "Mono-Unsaturated Fatty Acids Link H3K4me3 Modifiers to C. Elegans Lifespan." *Nature* 544 (7649): 185–90.
Hasegawa, Koichi, Satsuki Miwa, Tomoko Tajima, Kaname Tsutsumiuchi, Hajime Taniguchi, and Johji Miwa. 2007. "A Rapid and Inexpensive Method to Screen for Common Foods That Reduce the Action of Acrylamide, a Harmful Substance in Food." *Toxicology Letters* 175 (1-3): 82–88.
Hostettler, L., L. Grundy, S. Käser-Pébernard, C. Wicky, W. R. Schafer, and D. A. Glauser. 2017. "The Bright Fluorescent Protein mNeonGreen Facilitates Protein Expression Analysis in Vivo. G3 (Bethesda) 7: 607--615."
Hrach, Heather C., Shannon O'Brien, Hannah S. Steber, Jason Newbern, Alan Rawls, and Marco Mangone. 2020. "Transcriptome Changes during the Initiation and Progression of Duchenne Muscular Dystrophy in Caenorhabditis Elegans." *Human Molecular Genetics*. https://doi.org/10.1093/hmg/ddaa055.
Hu, Patrick J., Jinling Xu, and Gary Ruvkun. 2006. "Two Membrane-Associated Tyrosine Phosphatase Homologs Potentiate C. Elegans AKT-1/PKB Signaling." *PLoS Genetics* 2 (7): e99.
Ikegami, K., and J. D. Lieb. 2013. "Integral Nuclear Pore Proteins Bind to Pol III-Transcribed Genes and Are Required for Pol III Transcript Processing in C. Elegans." *Molecular Cell* 51 (6): 840–49.
Kagawa, H., K. Takuwa, and Y. Sakube. 1997. "Mutations and Expressions of the Tropomyosin Gene and the Troponin C Gene of Caenorhabditis Elegans." *Cell Structure and Function* 22 (1): 213–18.
Kagias, K., A. Ahier, N. Fischer, and S. Jarriault. 2012. "Members of the NODE (Nanog and Oct4-Associated Deacetylase) Complex and SOX-2 Promote the Initiation of a Natural Cellular Reprogramming Event in Vivo." *Proceedings of the National Academy of Sciences of the United States of America* 109 (17): 6596–6601.
Kaletsky, Rachel, Vanisha Lakhina, Rachel Arey, April Williams, Jessica Landis, Jasmine Ashraf, and Coleen T. Murphy. 2016. "The C. Elegans Adult Neuronal IIS/FOXO Transcriptome Reveals Adult Phenotype Regulators." *Nature* 529 (7584): 92–96.
Kaletsky, Rachel, Victoria Yao, April Williams, Alexi M. Runnels, Alicja Tadych, Shiyi Zhou, Olga G. Troyanskaya, and Coleen T. Murphy. 2018. "Transcriptome Analysis of Adult Caenorhabditis Elegans Cells Reveals Tissue-Specific Gene and Isoform Expression." *PLoS Genetics* 14 (8): e1007559.
Kalinava, Natallia, Julie Zhouli Ni, Kimberly Peterman, Esteban Chen, and Sam Guoping Gu. 2017. "Decoupling the Downstream Effects of Germline Nuclear RNAi Reveals That H3K9me3 Is Dispensable for Heritable RNAi and the Maintenance of Endogenous siRNA-Mediated Transcriptional Silencing in Caenorhabditis Elegans." *Epigenetics & Chromatin* 10 (1): 6.
Li, Heng. 2018. "Minimap2: Pairwise Alignment for Nucleotide Sequences." *Bioinformatics* 34 (18): 3094–3100.
Marques, Filipe, Gabriella Saro, Andrei-Stefan Lia, Richard J. Poole, Laurent Falquet, and Dominique A. Glauser. 2019. "Identification of Avoidance Genes through Neural Pathway-Specific Forward Optogenetics." *PLoS Genetics* 15 (12): e1008509.
Marshall, O. J., and A. H. Brand. 2015. "Damidseq_pipeline: An Automated Pipeline for Processing DamID Sequencing Datasets." *Bioinformatics* 31 (20): 3371–73.

Marshall, Owen J., and Andrea H. Brand. 2017. “Chromatin State Changes during Neural Development Revealed by in Vivo Cell-Type Specific Profiling.” *Nature Communications* 8 (1): 2271.

Mattout, Anna, Brietta L. Pike, Benjamin D. Towbin, Erin M. Bank, Adriana Gonzalez-Sandoval, Michael B. Stadler, Peter Meister, Yosef Gruenbaum, and Susan M. Gasser. 2011. “An EDMD Mutation in C. Elegans Lamin Blocks Muscle-Specific Gene Relocation and Compromises Muscle Integrity.” *Current Biology: CB* 21 (19): 1603–14.

Ma, Xiaopeng, Ge Zhan, Monica C. Sleumer, Siyu Chen, Weihong Liu, Michael Q. Zhang, and Xiao Liu. 2016. “Analysis of C. Elegans Muscle Transcriptome Using Trans-Splicing-Based RNA Tagging (SRT).” *Nucleic Acids Research* 44 (21): e156.

Miki, Takashi S., Sarah H. Carl, and Helge Großhans. 2017. “Two Distinct Transcription Termination Modes Dictated by Promoters.” *Genes & Development* 31 (18): 1870–79.

Miller, D. M., F. E. Stockdale, and J. Karn. 1986. “Immunological Identification of the Genes Encoding the Four Myosin Heavy Chain Isoforms of Caenorhabditis Elegans.” *Proceedings of the National Academy of Sciences of the United States of America* 83 (8): 2305–9.

Miyabayashi, T., M. T. Palfreyman, A. E. Sluder, F. Slack, and P. Sengupta. 1999. “Expression and Function of Members of a Divergent Nuclear Receptor Family in Caenorhabditis Elegans.” *Developmental Biology* 215 (2): 314–31.

Mohri, Kurato, and Shoichiro Ono. 2003. “Actin Filament Disassembling Activity of Caenorhabditis Elegans Actin-Interacting Protein 1 (UNC-78) Is Dependent on Filament Binding by a Specific ADF/cofilin Isoform.” *Journal of Cell Science* 116 (Pt 20): 4107–18.

Muñoz-Jiménez, Celia, Cristina Ayuso, Agnieszka Dobrzynska, Antonio Torres-Mendéz, Patricia de la Cruz Ruiz, and Peter Askjaer. 2017. “An Efficient FLP-Based Toolkit for Spatiotemporal Control of Gene Expression in Caenorhabditis Elegans.” *Genetics* 206 (4): 1763–78.

Ohkura, Kiyotaka, Norio Suzuki, Takeshi Ishihara, and Isao Katsura. 2003. “SDF-9, a Protein Tyrosine Phosphatase-like Molecule, Regulates the L3/dauer Developmental Decision through Hormonal Signaling in C. Elegans.” *Development* 130 (14): 3237–48.

Pauli, F., Y. Liu, Y. A. Kim, P. J. Chen, and S. K. Kim. 2006. “Chromosomal Clustering and GATA Transcriptional Regulation of Intestine-Expressed Genes in C. Elegans.” *Development* 133 (2): 287–95.

Pindyurin, Alexey V., Ludo Pagie, Elena N. Kozhevnikova, Joris van Arensbergen, and Bas van Steensel. 2016. “Inducible DamID Systems for Genomic Mapping of Chromatin Proteins in Drosophila.” *Nucleic Acids Research* 44 (12): 5646–57.

Raudvere, Uku, Liis Kolberg, Ivan Kuzmin, Tambet Arak, Priit Adler, Hedi Peterson, and Jaak Vilo. 2019. “g:Profiler: A Web Server for Functional Enrichment Analysis and Conversions of Gene Lists (2019 Update).” *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkz369.

Redolfi, Josef, Yinxiu Zhan, Christian Valdes-Quezada, Mariya Kryzhanovska, Isabel Guerreiro, Vytautas Iesmantavicius, Tim Pollex, et al. 2019. “DamC Reveals Principles of Chromatin Folding in Vivo without Crosslinking and Ligation.” *Nature Structural & Molecular Biology*. https://doi.org/10.1038/s41594-019-0231-0.

Roeder, Thomas, Mareike Stanisak, Christoph Gelhaus, Iris Bruchhaus, Joachim Grötzinger, and Matthias Leippe. 2010. “Caenopores Are Antimicrobial Peptides in the Nematode Caenorhabditis Elegans Instrumental in Nutrition and Immunity.” *Developmental and Comparative Immunology* 34 (2): 203–9.

Roy, Peter J., Joshua M. Stuart, Jim Lund, and Stuart K. Kim. 2002. “Chromosomal Clustering of Muscle-Expressed Genes in Caenorhabditis Elegans.” *Nature* 418 (6901): 975–79.

Ruijtenberg, S., and S. van den Heuvel. 2015. “G1/S Inhibitors and the SWI/SNF Complex Control Cell-Cycle Exit during Muscle Differentiation.” *Cell* 162 (2): 300–313.

Schuster, E., J. J. McElwee, J. M. Tullet, R. Doonan, F. Matthijssens, J. S. Reece-Hoyes, I. A. Hope, J. R. Vanfleteren, J. M. Thornton, and D. Gems. 2010. “DamID in C. Elegans Reveals Longevity-Associated Targets of DAF-16/FoxO.” *Molecular Systems Biology* 6: 399.

Sharma, Rahul, Daniel Jost, Jop Kind, Georgina Gómez-Saldivar, Bas van Steensel, Peter Askjaer, Cédric Vaillant, and Peter Meister. 2014. “Differential Spatial and Structural Organization of the X Chromosome Underlies Dosage Compensation in C. Elegans.” *Genes & Development* 28 (23): 2591–96.

Southall, T. D., K. S. Gold, B. Egger, C. M. Davidson, E. E. Caygill, O. J. Marshall, and A. H. Brand. 2013. “Cell-Type-Specific Profiling of Gene Expression and Chromatin Binding without Cell Isolation: Assaying RNA Pol II Occupancy in Neural Stem Cells.” *Developmental Cell* 26 (1): 101–12.

Spencer, W. Clay, Georg Zeller, Joseph D. Watson, Stefan R. Henz, Kathie L. Watkins, Rebecca D. McWhirter, Sarah Petersen, et al. 2011. “A Spatial and Temporal Map of C. Elegans Gene Expression.” *Genome Research* 21 (2): 325–41.

Steiner, F. A., P. B. Talbert, S. Kasinathan, R. B. Deal, and S. Henikoff. 2012. “Cell-Type-Specific Nuclei Purification from Whole Animals for Genome-Wide Expression and Chromatin Profiling.” *Genome Research* 22 (4): 766–77.

Tcherepanova, I., L. Bhattacharyya, C. S. Rubin, and J. H. Freedman. 2000. “Aspartic Proteases from the Nematode Caenorhabditis Elegans. Structural Organization and Developmental and Cell-Specific Expression of Asp-1.” *The Journal of Biological Chemistry* 275 (34): 26359–69.

Villaseñor, Rodrigo, Ramon Pfaendler, Christina Ambrosi, Stefan Butz, Sara Giuliani, Elana Bryan, Thomas W. Sheahan, et al. 2020. “ChromID Identifies the Protein Interactome at Chromatin Marks.” *Nature Biotechnology*. https://doi.org/10.1038/s41587-020-0434-2.

Von Stetina, S. E., J. D. Watson, R. M. Fox, K. L. Olszewski, W. C. Spencer, P. J. Roy, and D. M. Miller 3rd. 2007. “Cell-Specific Microarray Profiling Experiments Reveal a Comprehensive Picture of Gene Expression in the C. Elegans Nervous System.” *Genome Biology* 8 (7): R135.

Wang, Juan, Rachel Kaletsky, Malan Silva, April Williams, Leonard A. Haas, Rebecca J. Androwski, Jessica N. Landis, et al. 2015. “Cell-Specific Transcriptional Profiling of Ciliated Sensory Neurons Reveals Regulators of Behavior and Extracellular Vesicle Biogenesis.” *Current Biology: CB* 25 (24): 3232–38.

Widmer, Yves F., Adem Bilican, Rémy Bruggmann, and Simon G. Sprecher. 2018. “Regulators of Long-Term Memory Revealed by Mushroom Body-Specific Gene Expression Profiling in Drosophila Melanogaster.” *Genetics* 209 (4): 1167–81.

Zhang, L., J. D. Ward, Z. Cheng, and A. F. Dernburg. 2015. “The Auxin-Inducible Degradation (AID) System Enables Versatile Conditional Protein Depletion in C. Elegans.” *Development* 142 (24): 4374–84.

Zhang, Yuru, Xiaoju Zou, Yihong Ding, Haizhen Wang, Xiaoyun Wu, and Bin Liang. 2013. “Comparative Genomics and Functional Study of Lipid Metabolic Genes in Caenorhabditis Elegans.” *BMC Genomics* 14 (March): 164.

Figure 1

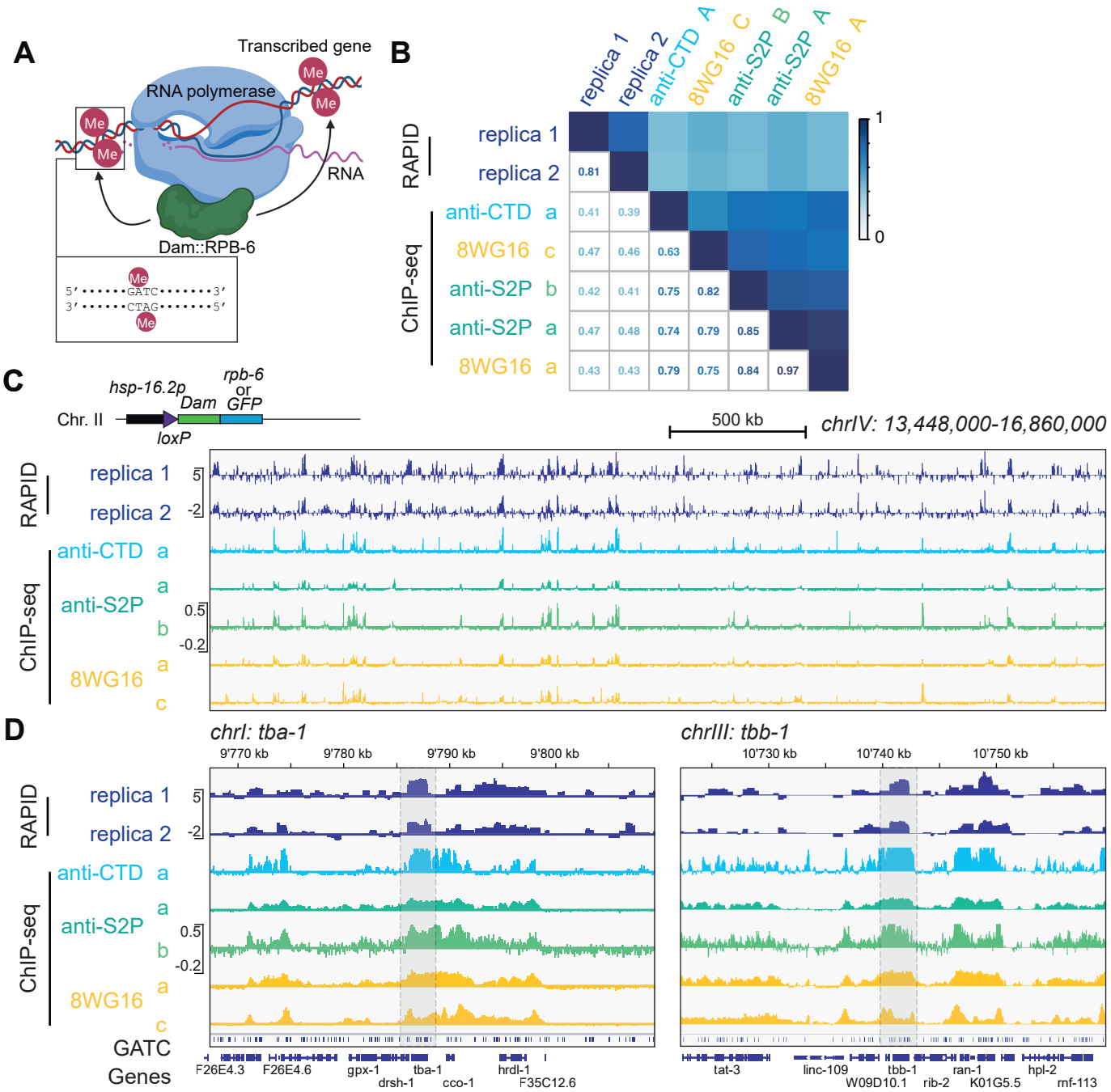

A
Transcribed gene
RNA polymerase
Me
RNA
Dam::RPB-6
5′ GATC 3′
3′ CTAG 5′
B
replica 1
replica 2
anti-CTD A
8WG16 C
anti-S2P B
anti-S2P A
8WG16 A
RAPID
ChIP-seq
0.81
0.41 0.39
0.47 0.46 0.63
0.42 0.41 0.75 0.82
0.47 0.48 0.74 0.79 0.85
0.43 0.43 0.79 0.75 0.84 0.97
1
0
C
hsp-16.2p
rpb-6 or GFP
Dam
Chr. II
loxP
500 kb
chrIV: 13,448,000-16,860,000
anti-S2P
8WG16
D
chrI: tba-1
chrIII: tbb-1
9'770 kb
9'780 kb
9'790 kb
9'800 kb
10'730 kb
10'740 kb
10'750 kb
GATC
Genes
F26E4.3
F26E4.6
gpx-1
tba-1
drsh-1
cco-1
hrdl-1
F35C12.6
tat-3
linc-109
tbb-1
W09D10.1
rib-2
ran-1
K01G5.5
hpl-2
rnf-113

Figure 2

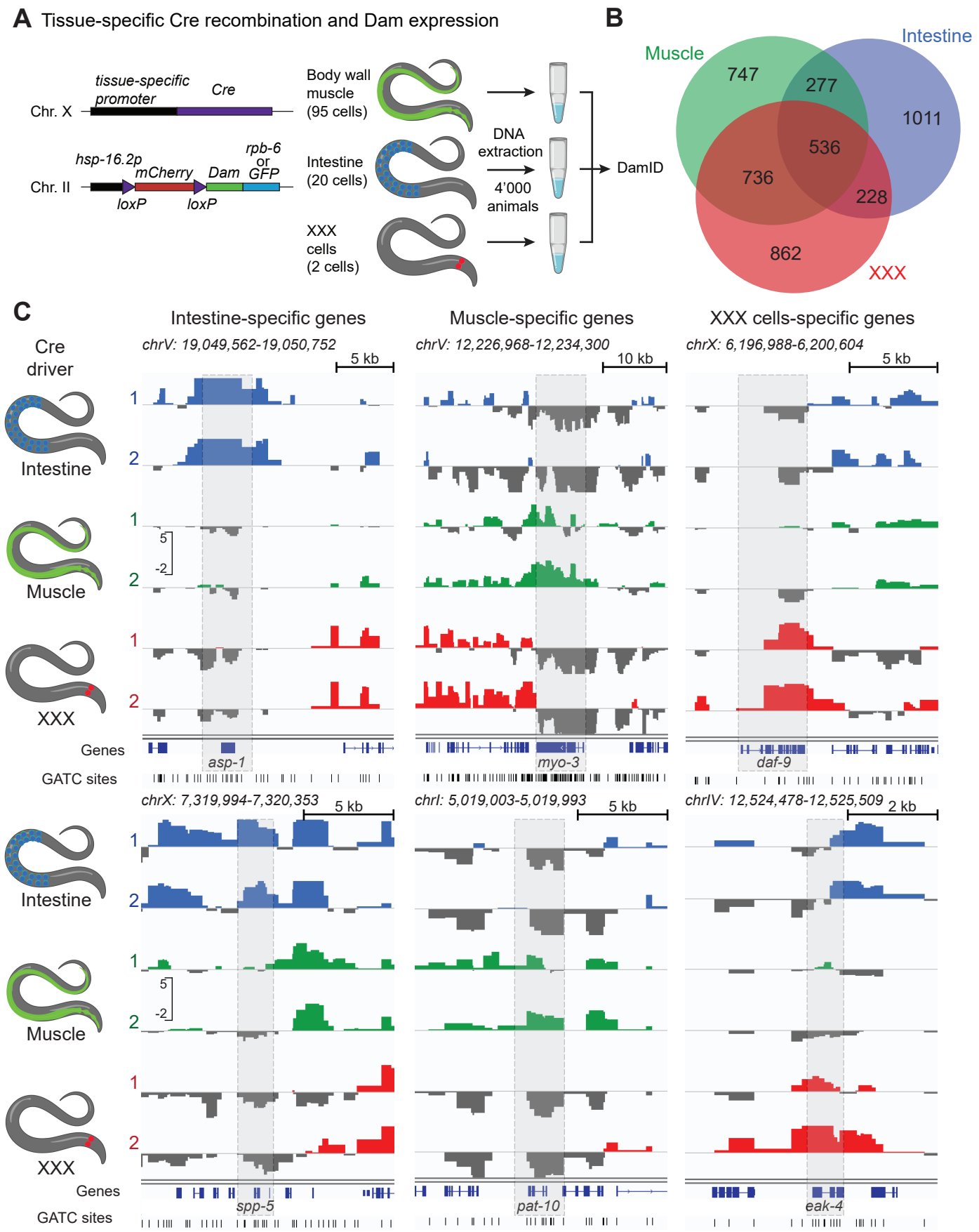

A Tissue-specific Cre recombination and Dam expression
tissue-specific promoter
Cre
Chr. X
hsp-16.2p
mCherry
Dam
rpb-6 or GFP
Chr. II
loxP
loxP
Body wall muscle (95 cells)
Intestine (20 cells)
XXX cells (2 cells)
DNA extraction
4'000 animals
DamID
B
Muscle
Intestine
XXX
747
277
1011
536
736
228
862
C
Intestine-specific genes
Muscle-specific genes
XXX cells-specific genes
Cre driver
Intestine
Muscle
XXX
chrV: 19,049,562-19,050,752
5 kb
chrV: 12,226,968-12,234,300
10 kb
chrX: 6,196,988-6,200,604
5 kb
Genes
asp-1
myo-3
daf-9
GATC sites
chrX: 7,319,994-7,320,353
5 kb
chrI: 5,019,003-5,019,993
5 kb
chrIV: 12,524,478-12,525,509
2 kb
spp-5
pat-10
eak-4

Figure 3

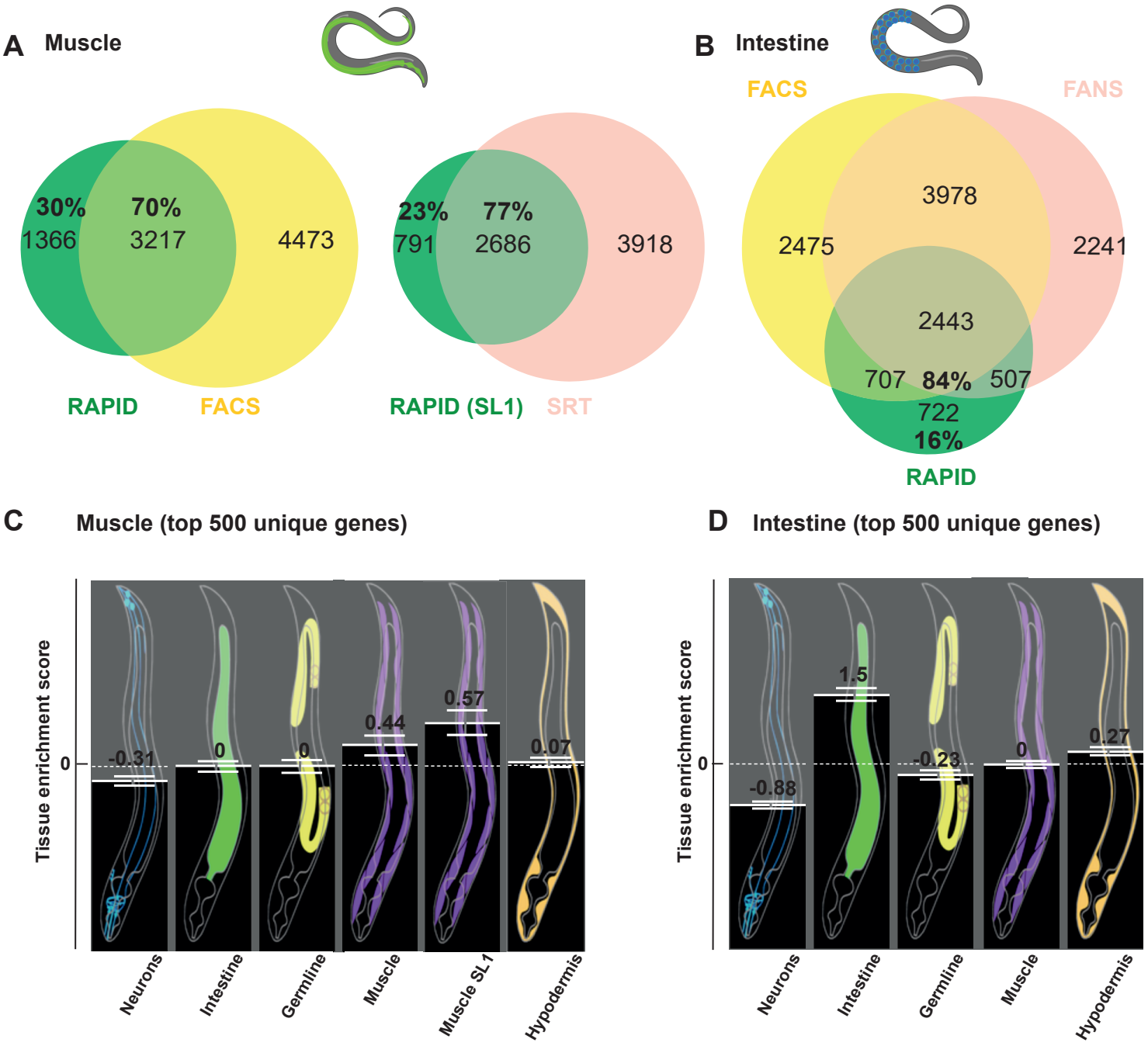

A
Muscle
30%
1366
70%
3217
4473
RAPID
FACS
23%
791
77%
2686
3918
RAPID (SL1)
SRT
B
Intestine
FACS
FANS
3978
2475
2241
2443
707
84%
507
722
16%
RAPID
C
Muscle (top 500 unique genes)
Tissue enrichment score
0
-0.31
0
0
0.44
0.57
0.07
Neurons
Intestine
Germline
Muscle
Muscle SL1
Hypodermis
D
Intestine (top 500 unique genes)
Tissue enrichment score
0
-0.88
1.5
-0.23
0
0.27
Neurons
Intestine
Germline
Muscle
Hypodermis

Figure 4

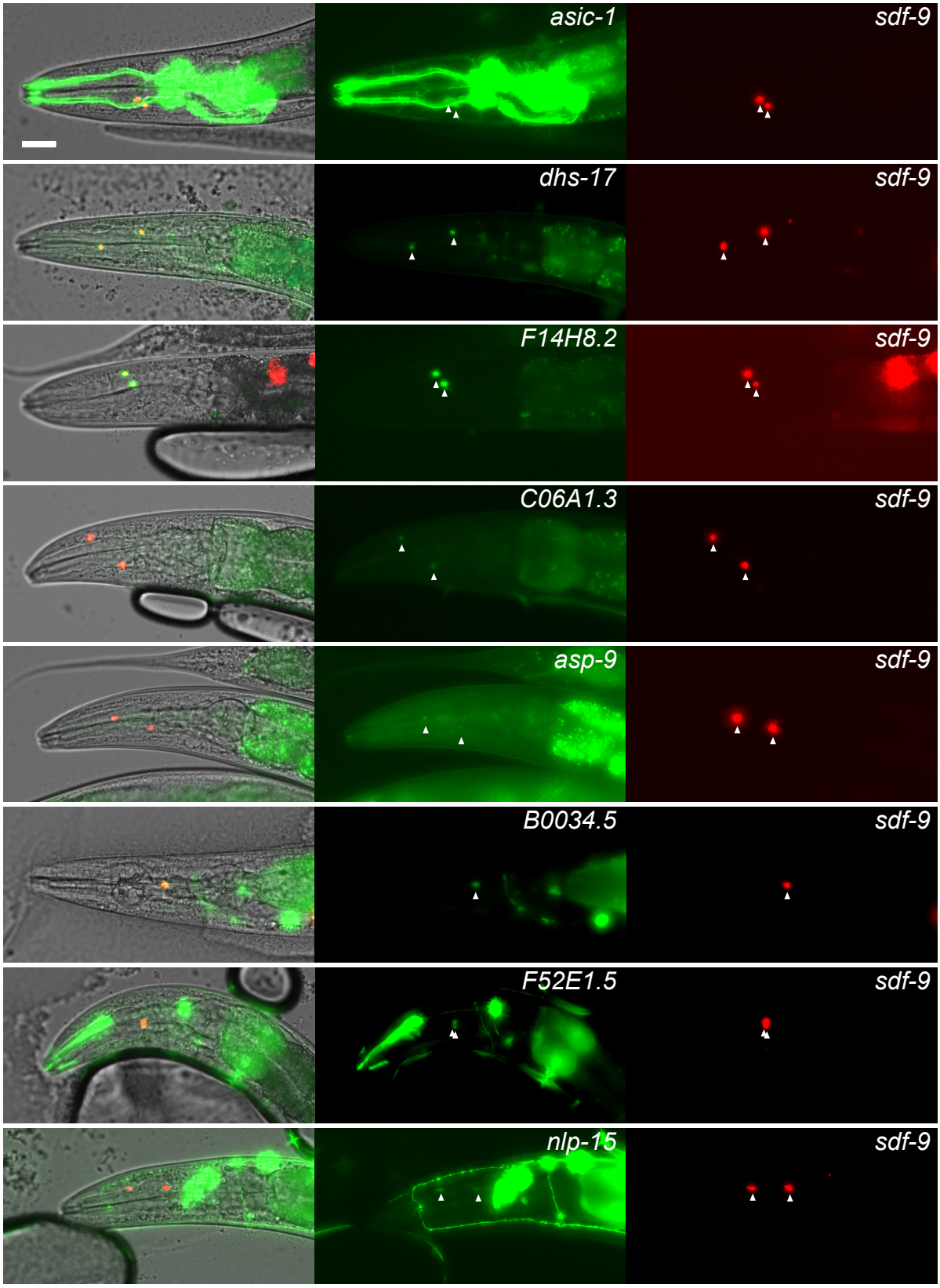

asic-1
sdf-9
dhs-17
sdf-9
F14H8.2
sdf-9
C06A1.3
sdf-9
asp-9
sdf-9
B0034.5
sdf-9
F52E1.5
sdf-9
nlp-15
sdf-9

Figure S1

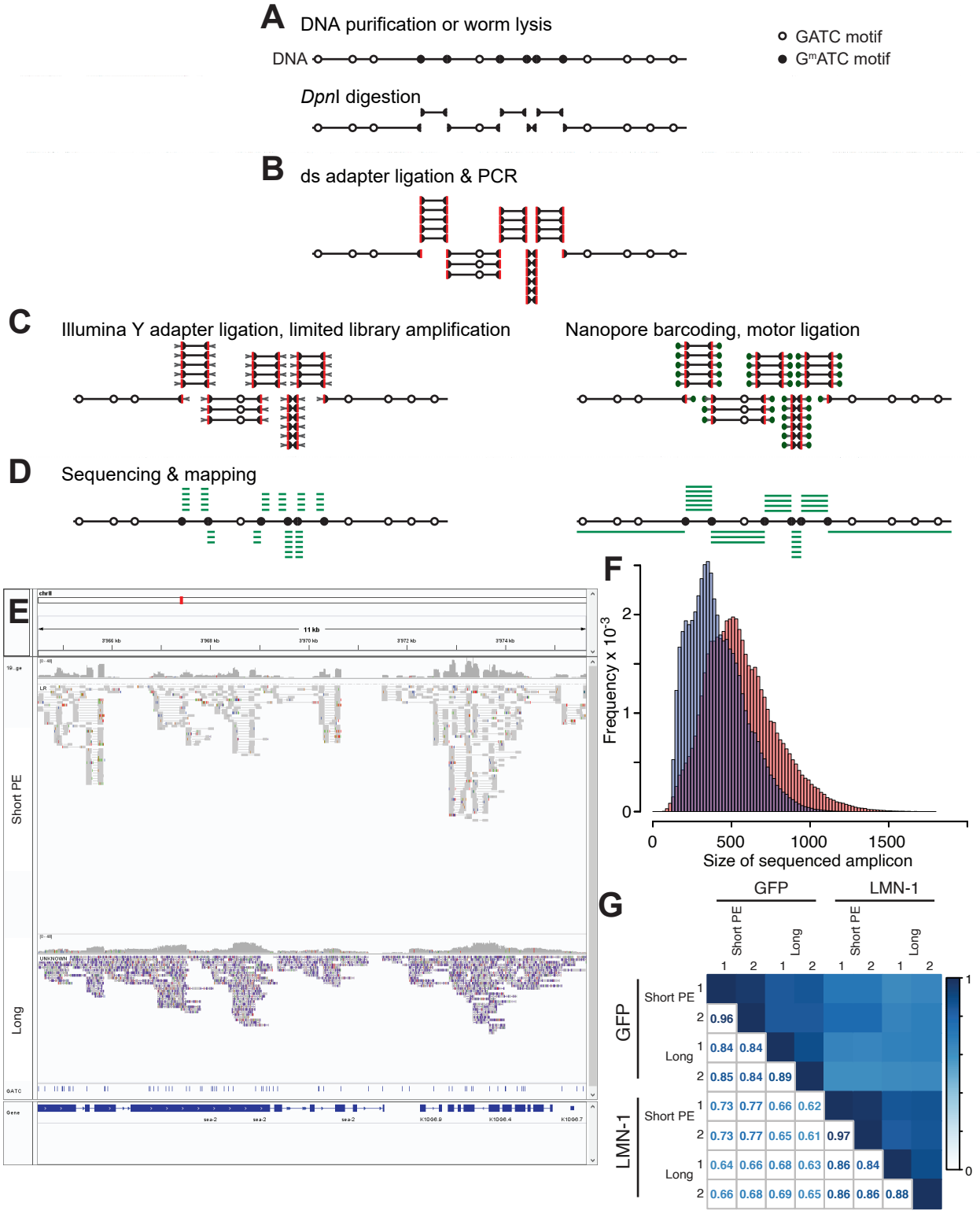
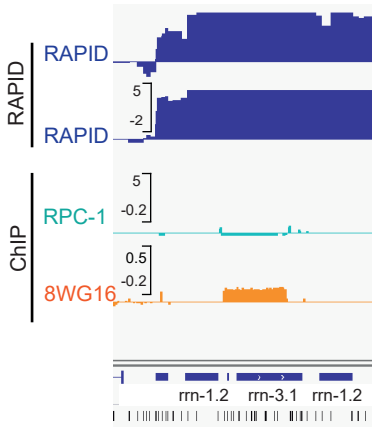
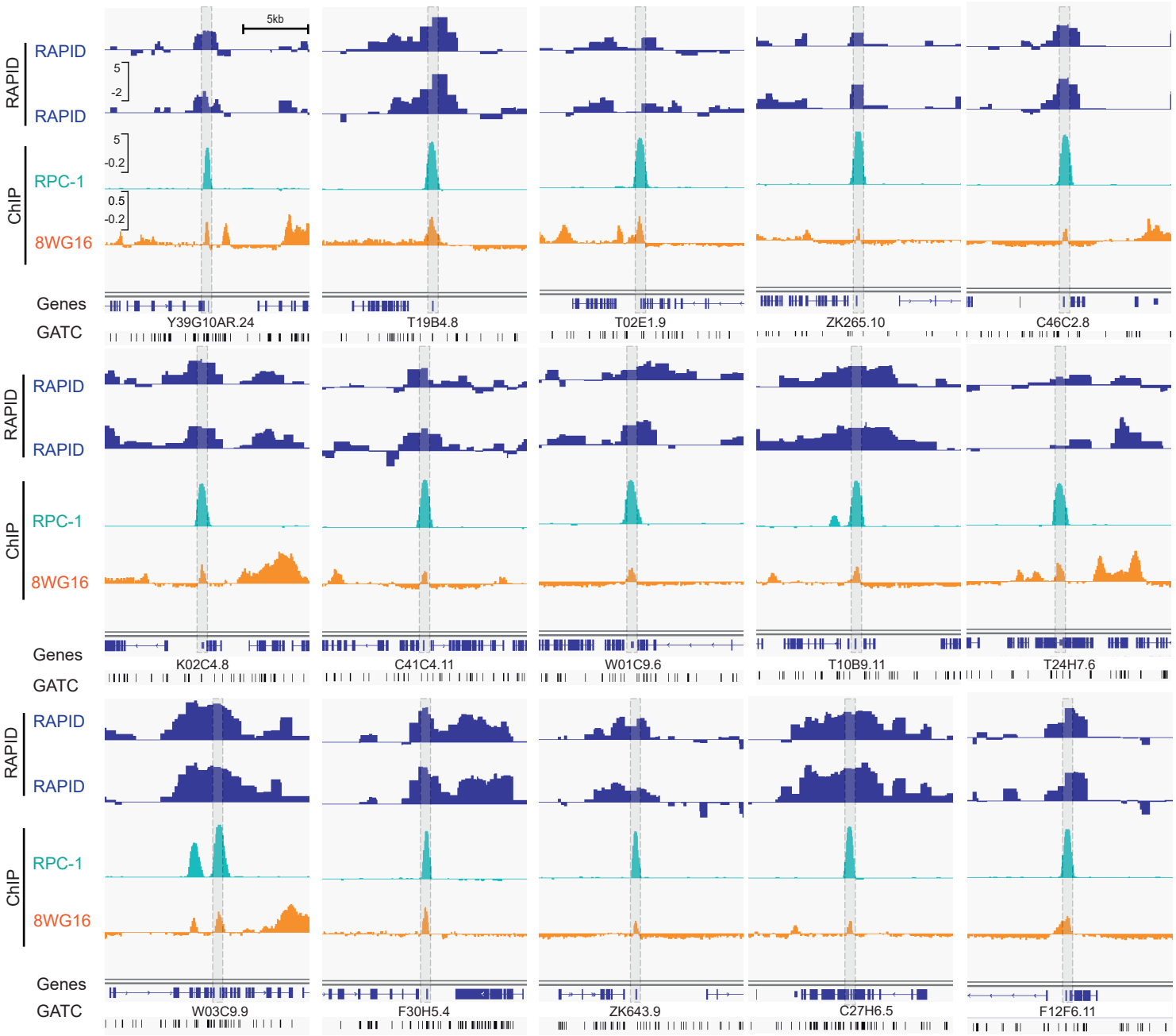

A
DNA purification or worm lysis
DNA
GATC motif
GmATC motif
DpnI digestion
B
ds adapter ligation & PCR
C
Illumina Y adapter ligation, limited library amplification
Nanopore barcoding, motor ligation
D
Sequencing & mapping
E
chrII
11 kb
Short PE
Long
GATC
Gene
F
Frequency x 10-3
Size of sequenced amplicon
G
GFP
LMN-1
Short PE
Long
0.96
0.84 0.84
0.85 0.84 0.89
0.73 0.77 0.66 0.62
0.73 0.77 0.65 0.61 0.97
0.64 0.66 0.68 0.63 0.86 0.84
0.66 0.68 0.69 0.65 0.86 0.86 0.88

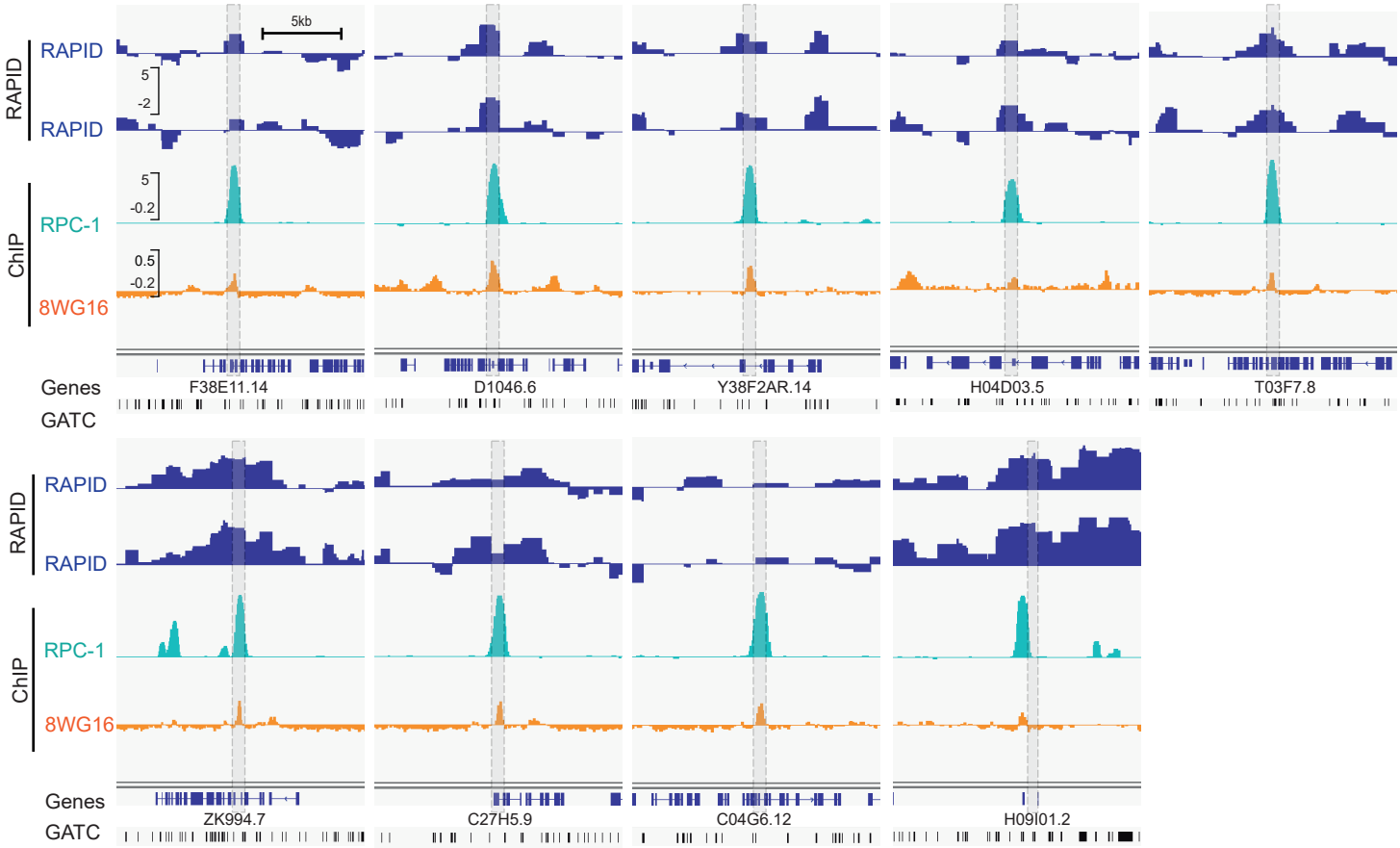
Figure S2

A RAPID signal at rDNA repeats

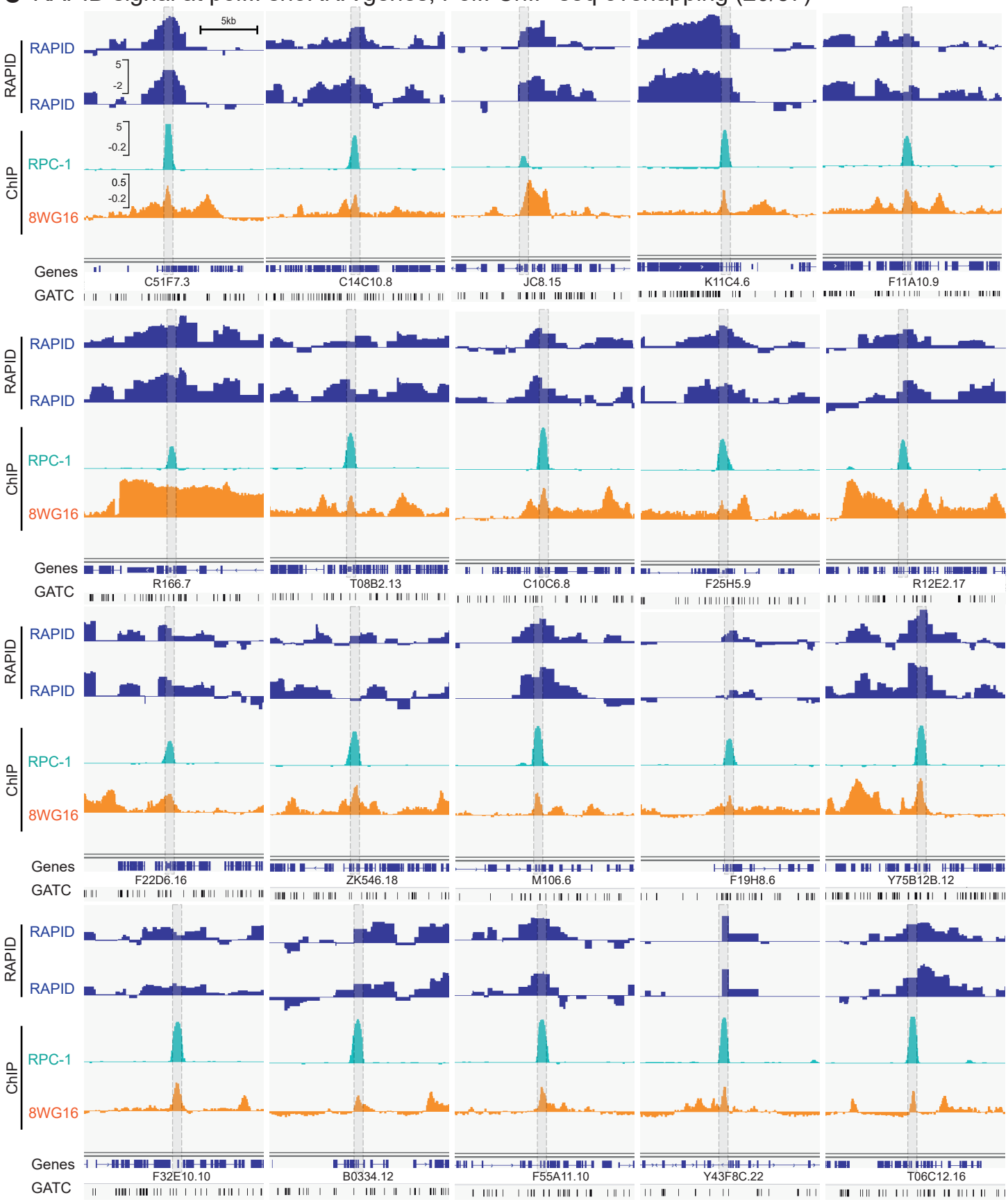


B RAPID signal at polIII snoRNA genes; no PolII overlapping (24/57)



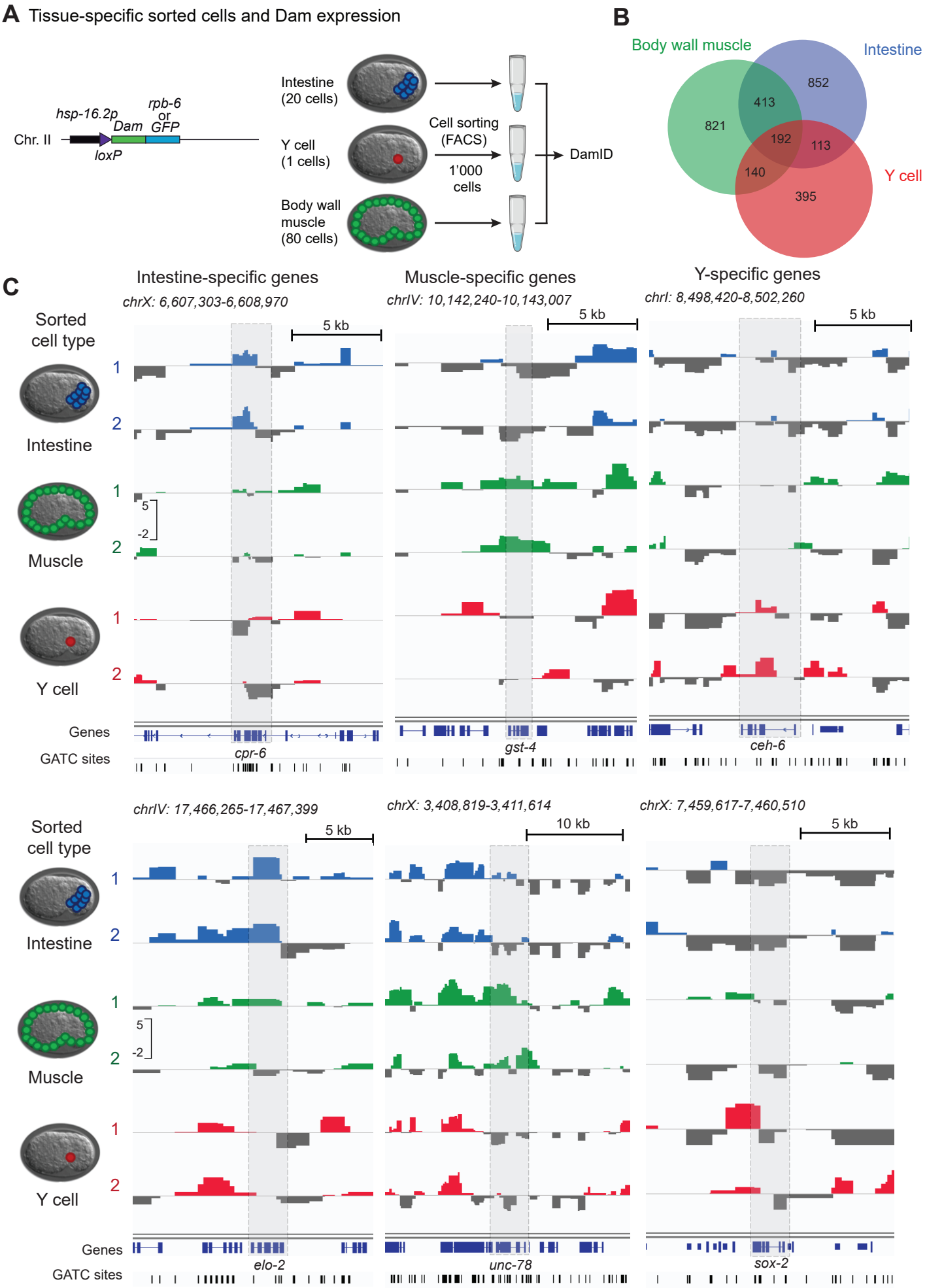

5kb
RAPID
RAPID
RAPID
5
-2
ChIP
RPC-1
5
-0.2
8WG16
0.5
-0.2
Genes
GATC
F38E11.14
D1046.6
Y38F2AR.14
H04D03.5
T03F7.8
RAPID
RAPID
RAPID
ChIP
RPC-1
8WG16
Genes
GATC
ZK994.7
C27H5.9
C04G6.12
H09I01.2

C RAPID signal at polIII snoRNA genes; PolII ChIP-seq overlapping (20/57)



D No RAPID signal (13/57)

Figure S3


A Tissue-specific sorted cells and Dam expression
Chr. II
hsp-16.2p
Dam
rpb-6
or
GFP
loxP
Intestine
(20 cells)
Y cell
(1 cells)
Body wall
muscle
(80 cells)
Cell sorting
(FACS)
1'000
cells
DamID
B
Body wall muscle
Intestine
Y cell
821
413
852
192
113
140
395
C
Intestine-specific genes
chrX: 6,607,303-6,608,970
Muscle-specific genes
chrIV: 10,142,240-10,143,007
Y-specific genes
chrI: 8,498,420-8,502,260
5 kb
Sorted
cell type
Intestine
Muscle
Y cell
Genes
GATC sites
cpr-6
gst-4
ceh-6
chrIV: 17,466,265-17,467,399
chrX: 3,408,819-3,411,614
chrX: 7,459,617-7,460,510
10 kb
elo-2
unc-78
sox-2

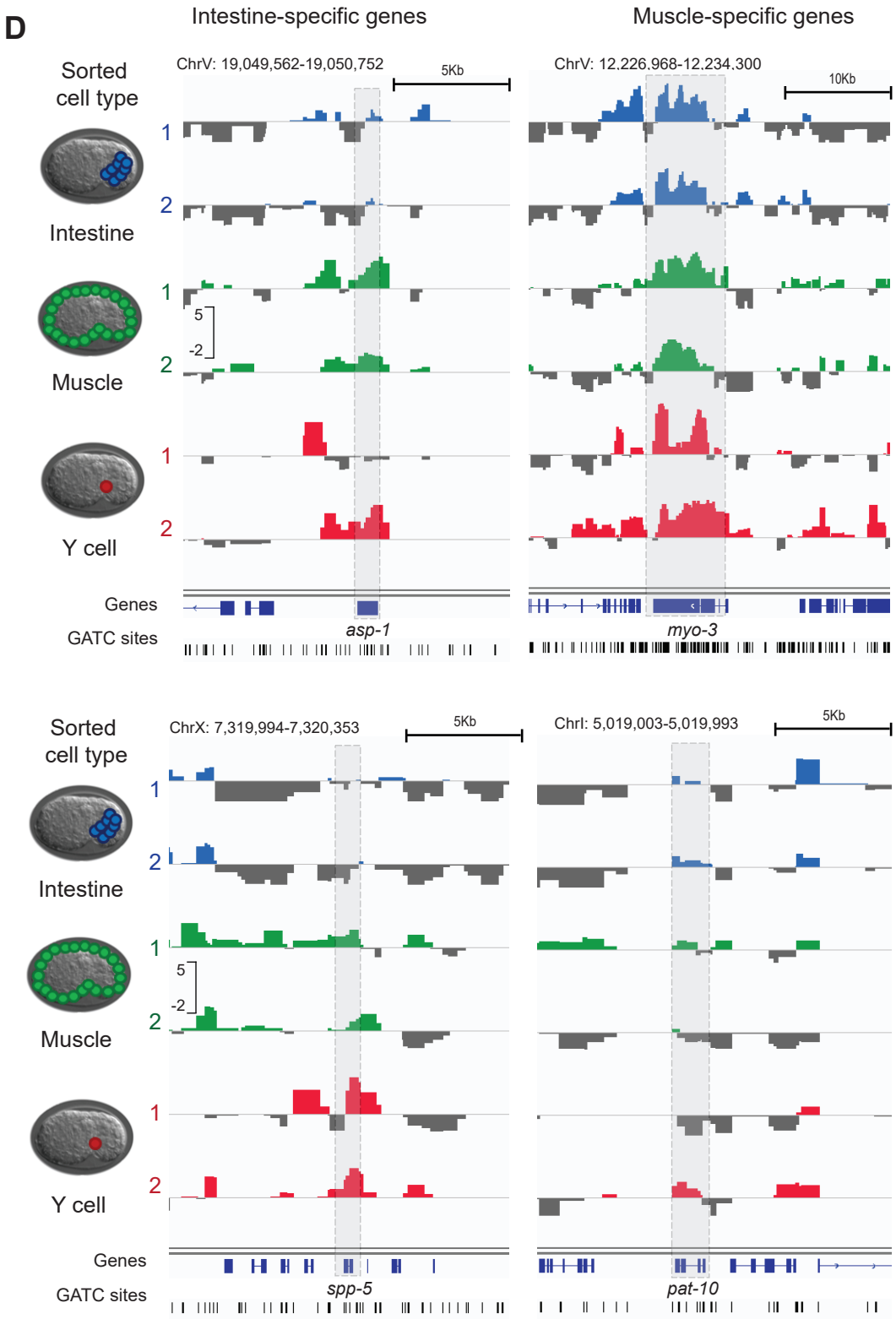

D
Intestine-specific genes
Muscle-specific genes
Sorted cell type
ChrV: 19,049,562-19,050,752
5Kb
ChrV: 12,226,968-12,234,300
10Kb
Intestine
Muscle
Y cell
1
2
5
-2
Genes
GATC sites
asp-1
myo-3
ChrX: 7,319,994-7,320,353
ChrI: 5,019,003-5,019,993
spp-5
pat-10

Figure S4

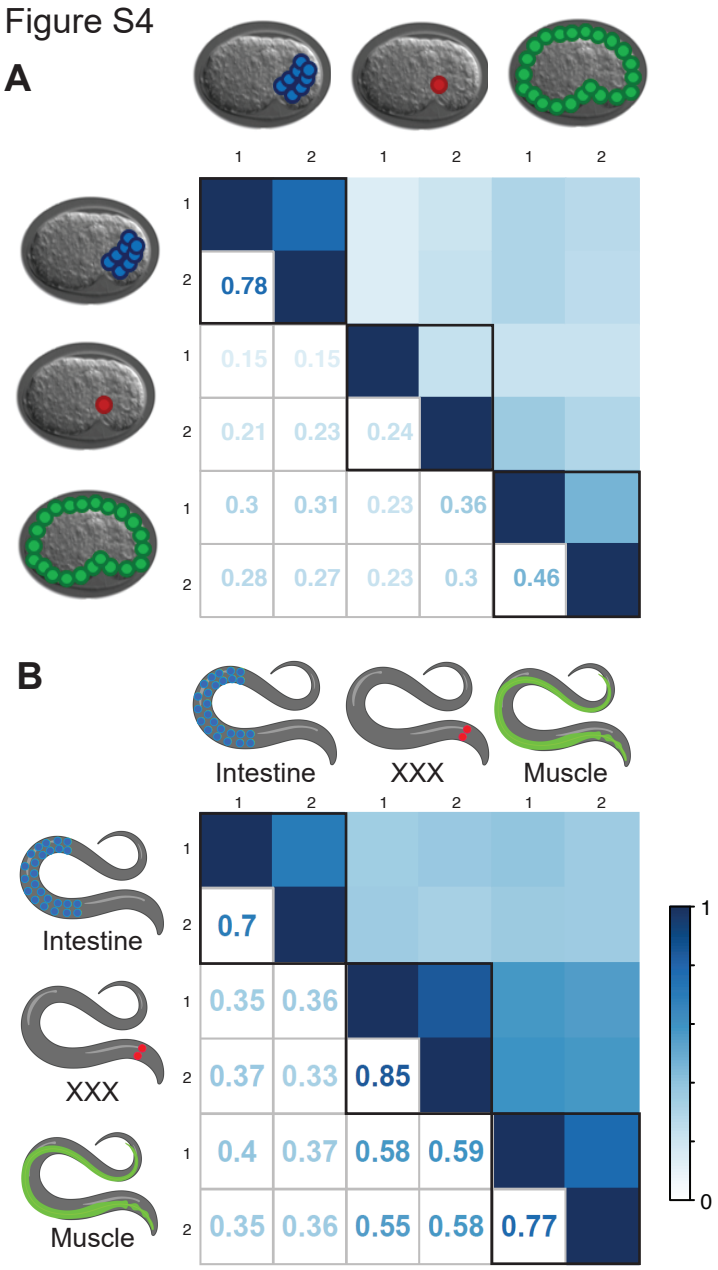
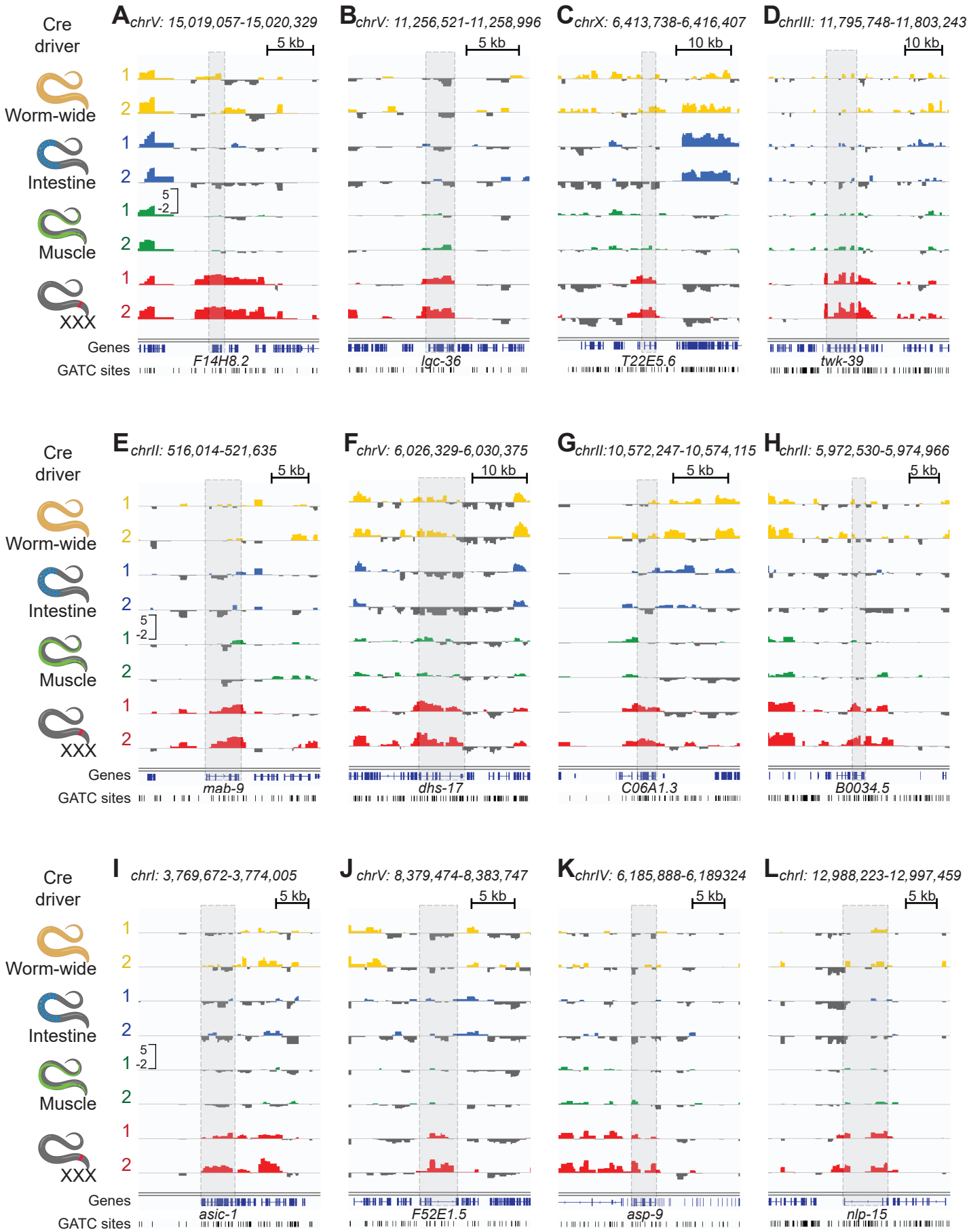

A
1 2 1 2 1 2
1 2 1 2 1 2
0.78
0.15 0.15
0.21 0.23 0.24
0.3 0.31 0.23 0.36
0.28 0.27 0.23 0.3 0.46
B
Intestine XXX Muscle
1 2 1 2 1 2
1 2 1 2 1 2
0.7
0.35 0.36
0.37 0.33 0.85
0.4 0.37 0.58 0.59
0.35 0.36 0.55 0.58 0.77
Intestine
XXX
Muscle
1
0

Figure S5


Cre driver
Worm-wide
Intestine
Muscle
XXX
Genes
GATC sites
A chrV: 15,019,057-15,020,329
5 kb
F14H8.2
B chrV: 11,256,521-11,258,996
5 kb
lqc-36
C chrX: 6,413,738-6,416,407
10 kb
T22E5.6
D chrIII: 11,795,748-11,803,243
10 kb
twk-39
E chrII: 516,014-521,635
5 kb
mab-9
F chrV: 6,026,329-6,030,375
10 kb
dhs-17
G chrII:10,572,247-10,574,115
5 kb
C06A1.3
H chrII: 5,972,530-5,974,966
5 kb
B0034.5
I chrI: 3,769,672-3,774,005
5 kb
asic-1
J chrV: 8,379,474-8,383,747
5 kb
F52E1.5
K chrIV: 6,185,888-6,189324
5 kb
asp-9
L chrI: 12,988,223-12,997,459
5 kb
nlp-15