

## 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 *DpnI* 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 amplicon 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 PolII 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 *et al.* 2017). **B.** RAPID signal at PolIII snoRNA genes, with no PolIII 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.** Procedure schematic. Embryonic blastomeres from a strain ubiquitously expressing trace amounts of Dam constructs (either fused to GFP or RPB-6) were sorted based on the expression of fluorescent markers using fluorescence-activated cell sorting (FACS). For each

construct and each tissue, 1'000 cells were pooled before performing DamID. **B.** Venn diagram showing overlap between RAPID footprinted genes (FDR < 0.05 in both replicates). **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 adult tissues using Cre-mediated recombination for tissue-specific expression of the Dam fusions.

**Figure S5.** Extended comparative analysis of the muscular and intestinal transcript profiles obtained with RAPID and RNA-seq-based methods.

**A.** Overlap of genes footprinted using RAPID with muscle-expressed genes detected by FACS-seq (Kaletsky *et al.* 2018) and PAT-seq (Blazie *et al.* 2017). **B.** Overlap of SL1-trans-spliced genes detected by different methods in muscle: RAPID (this study), PAT-seq (Blazie *et al.* 2017), FACS ((Kaletsky *et al.* 2018) and SRT (Ma *et al.* 2016). The selection of SL1-trans-spliced protein-coding genes was made according to the annotation of modENCODE (Allen *et al.* 2011, Ma *et al.* 2016). **C.** Overlap of genes footprinted using RAPID with intestine-expressed genes detected by PAT-seq (Blazie *et al.* 2017), FACS (Kaletsky *et al.* 2018) and FANS (Haenni *et al.* 2012). **D.** Tables showing the percentage of genes commonly detected by method pairs (listed in A-C, respectively). For the muscle transcriptome, two tables are shown: one with all the genes detected by the methods in A (left) and another with the SL1-trans-spliced genes as used for the analysis in B (middle). **E.** Cross-method consistency as assessed with the fractions of genes detected by one, two or three methods. The percentage of genes detected by at least one method is indicated in the last column ( $\Sigma$ ). The analysis corresponding to muscle and muscle-SL1 is shown in the left and middle tables. The analysis corresponding to the intestine is shown in the table to the right.

**Figure S6.** Muscle transcriptome GO-analysis.

Comparison of gene sets commonly identified by two methods (Common) and gene sets identified by only one method (Only-RAPID, Only-FACS, Only-PAT-seq). **A.** Comparison between RAPID (this study) and FACS (Kaletsky *et al.* 2018). **B.** Comparison between RAPID (this study) and PAT-seq (Blazie *et al.* 2017). **C.** Comparison between FACS ((Kaletsky *et al.* 2018) and PAT-seq (Blazie *et al.* 2017). Functional enrichment analysis was performed using WormCat. The pervasively represented category "unknown" was removed from the depicted

heatmap. Green asterisks highlight expected muscle-specific categories and red asterisks highlight less expected categories for this tissue (more typical in other tissues; Holdorf *et al.* 2020).

**Figure S7.** Intestine transcriptome GO-analysis.

Comparison of gene sets commonly identified by two methods (Common) and gene sets identified by only one method (Only-RAPID, Only FACS, Only PAT-seq). **A.** Comparison between RAPID (this study) and FACS (Kaletsky *et al.* 2018). **B.** Comparison between RAPID (this study) and PAT-seq (Blazie *et al.* 2017). **C.** Comparison between FACS (Kaletsky *et al.* 2018) and PAT-seq (Blazie *et al.* 2017). Functional enrichment analysis was performed using WormCat. The pervasively represented category "unknown" was removed from the depicted heatmap. Green asterisks highlight expected intestine-specific categories and red asterisks highlight less expected categories for this tissue (more typical in other tissues; Holdorf *et al.* 2020)

**Figure S8.** Semi-quantitative comparison between RAPID and sciRNA-seq (L2 larval stage; Cao *et al.* 2017) in intestine (**A**) and muscle (**B**). All genes are represented on the x axis ranked, from left to right, based on the RAPID signal in the considered tissue (shown on left y axis). Averages of transcripts per million (tpm) for those genes in cells identified as intestine (511 cells) or muscle (10'508 cells) were calculated, using the genes falling into each bin of 690 genes on the x axis (values on right y axis). The muscular and intestinal marker genes presented in Fig. 2C are indicated in green and blue, respectively.

**Figure S9.** RAPID profile of candidate genes with XXX-specific expression.

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

**Figure S10.** Estimation of the true positive rate in XXX-specific RAPID hits. Cumulative distribution of the predicted number of true positive hits in the XXX RAPID dataset, based on the fraction of true positives in the gene subset tested with the reporter gene approach. Resampling was performed with 10'000 bootstrap iterations. Boundaries of the 95% confidence interval are depicted in red.

## **Supplementary tables**

Table S1 (excel sheet)

Strains used in this study

Table S2 (excel sheet)

Plasmids used in this study

Table S3 (excel sheet)

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

Table S4 (excel sheet)

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

Table S5 (excel sheet)

Transcribed genes analysis determined by RAPID. Genes for each tissue are listed. Detected and unique genes, and Gene ontology attributes are tabulated in different sheets.

Table S6 (excel sheet)

List of categories found in the functional enrichment analysis for genes detected with different methods in the intestine and muscle.

Table S7 (excel sheet)

Comparison between the different methods for cell-type specific transcriptome profiling.

Table S8 (excel sheet)

Sequencing depth for the RAPID libraries.