Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling

Florian A. Steiner,[1] Paul B. Talbert,[1] Sivakanthan Kasinathan,[1] Roger B. Deal,[1] and Steven Henikoff[1,2,3]

[1]Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA; [2]Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA

An understanding of developmental processes requires knowledge of transcriptional and epigenetic landscapes at the level of tissues and ultimately individual cells. However, obtaining tissue- or cell-type-specific expression and chromatin profiles for animals has been challenging. Here we describe a method for purifying nuclei from specific cell types of animal models that allows simultaneous determination of both expression and chromatin profiles. The method is based on in vivo biotin-labeling of the nuclear envelope and subsequent affinity purification of nuclei. We describe the use of the method to isolate nuclei from muscle of adult *Caenorhabditis elegans* and from mesoderm of *Drosophila melanogaster* embryos. As a case study, we determined expression and nucleosome occupancy profiles for affinity-purified nuclei from *C. elegans* muscle. We identified hundreds of genes that are specifically expressed in muscle tissues and found that these genes are depleted of nucleosomes at promoters and gene bodies in muscle relative to other tissues. This method should be universally applicable to all model systems that allow transgenesis and will make it possible to determine epigenetic and expression profiles of different tissues and cell types.

[Supplemental material is available for this article.]

Multicellular organisms are composed of multiple tissues and cell types, each of which differentiates from an undifferentiated progenitor. This differentiation involves an epigenetic reprogramming of the progenitor cell to establish the appropriate cell-type-specific expression profile. The set of genes expressed within each cell type that specifies its identity and allows it to perform its function within the organism needs to be defined during differentiation and maintained in the differentiated tissue. A combination of chromatin-based mechanisms involving transcription factor binding, nucleosome remodeling, deposition of histone variants, and post-translational histone modifications underlies these processes (Ng and Gurdon 2008; Yuan and Zhu 2011). The resulting epigenetic landscape determines how tissues develop, are maintained, and function in the context of a complete organism. How these processes are regulated in vivo is an important question that is currently difficult to address, because it is technically challenging to obtain pure populations of a certain cell type with enough yield or purity to perform expression or epigenetic profiling.

Several different approaches have been developed to investigate cell-type- or tissue-specific expression patterns in vivo. Tagged RNA-binding proteins and chemically modified RNA have been used to obtain cell type expression profiles (Roy et al. 2002; Miller et al. 2009), but these methods are not suited to study chromatin. The use of homogeneous cell cultures or ex vivo differentiated cells allows the examination of large amounts of material (Azuara et al. 2006; Fox et al. 2007), but these cells lack the proper context within the organism and cannot be used to analyze

a developmental series, since each cell line is produced independently. Techniques such as fluorescent activated cell sorting (FACS) or laser capture microdissection (LCM) make it possible to isolate specific cell or nuclei populations (Neira and Azen 2002; Von Stetina et al. 2007; Stoeckius et al. 2009; Burgemeister 2011), but these techniques involve harsh treatment of the tissue before the analysis and are challenging for certain tissues, unsuitable for rare cells, or deliver low yield. As such, these methods are generally unsuitable for genome-wide epigenomic profiling, which requires relatively large amounts of chromatin. These techniques are also expensive, relatively slow, and require specialized equipment and specific expertise.

We recently introduced a simple strategy for purifying cell-type-specific nuclei from *Arabidopsis thaliana* that circumvents these problems (Deal and Henikoff 2010). The INTACT (isolation of nuclei tagged in specific cell types) method uses affinity purification of nuclei tagged in specific cell types and delivers both RNA and chromatin for profiling. It allows the investigation of different cell types at different stages of development, requires minimal processing of the tissue, and circumvents the need for specialized equipment and training. The method was developed for the *A. thaliana* root epidermis, a relatively simple system with radial symmetry and only two cell types. However, proof of concept in morphologically complex animal systems with multiple cell types is still lacking.

In this study, we have adapted the INTACT strategy for expression and chromatin analysis to animal models. We describe a system for affinity purification of tagged nuclei from both *Caenorhabditis elegans* and *Drosophila melanogaster*. We coexpressed a nuclear envelope targeting fusion protein that is a substrate for in vivo biotinylation and the *Escherichia coli* biotin ligase BirA in muscle cells of adult *C. elegans* or mesoderm of *D. melanogaster* embryos. Biotinylated nuclei can be affinity-purified using bead-bound

streptavidin. Using this strategy, we could purify pure populations of *C. elegans* muscle nuclei and *D. melanogaster* mesoderm nuclei. We analyzed expression and chromatin profiles in *C. elegans* muscle nuclei as a case study. *C. elegans* is an ideal model system to develop this method, because it has a small number of well-characterized tissues, is relatively easily transformed, and is easily propagated. There are about 100 body-wall muscle cells per animal, containing ~10% of the somatic and 5% of the total nuclei in an adult animal. The number of muscle nuclei is sufficiently small for determining purity above background while providing sufficient abundance for expression and chromatin analyses even when starting with relatively small worm cultures. We identified hundreds of genes preferentially expressed in affinity-purified nuclei and show that they reflect expression characteristics of muscle tissue. We also found that this expression profile correlates with low nucleosome occupancy at promoters and gene bodies of the preferentially expressed genes in muscle tissue. Our results show that our method is useful for investigating the interplay between chromatin structure and gene expression in specific cell types in animal models.

Figure 1. *C. elegans* transgenes and transgenic lines for affinity purification of nuclei. (*A*) Schematic representation of the *C. elegans* nuclear targeting fusion (NTF) transgene, consisting of the muscle-specific *myo-3* promoter, the whole genomic sequence of *npp-9* for targeting to the nuclear envelope, *mCherry* for fluorescence visualization, biotin ligase recognition peptide (*BLRP*) as a substrate for biotinylation, and *3xFLAG* for immunodetection, followed by the *npp-9* 3′UTR. The biotin tag is ligated in vivo to the lysine residue shown in bold within the sequence of the BLRP. (*B*) Schematic representation of the *birA::gfp* transgene, consisting of the *E. coli* biotin ligase *birA* fused to *gfp*, flanked by the *his-72* promoter and 3′ UTR. (*C,D*) Adult transgenic *C. elegans* showing expression of the NTF targeted to the nuclei of muscle cells. mCherry fluorescence (*C*) and DIC image (*D*).

Results

A nuclear tagging strategy for *C. elegans*

We developed a system for biotin tagging of nuclei in animal cells, similar to the one previously reported for *A. thaliana* (Deal and Henikoff 2010), where an outer nuclear pore fusion protein serves as a substrate for biotinylation by *E. coli* biotin ligase (BirA), which is coexpressed in the same cells to mediate specific biotinylation. The N-terminal RANGAP1 WPP domain that was used to purify *A. thaliana* nuclei is found only in plants, thus it was not a good candidate for targeting to animal nuclear envelopes (Meier et al. 2008). For *C. elegans*, we selected the nuclear pore complex protein NPP-9 to target a fusion protein to the nuclear envelope. NPP-9 is the *C. elegans* homolog of RANBP2/NUP358 and localizes to the cytoplasmic filaments of the nuclear pore (D'Angelo and Hetzer 2008). It is therefore likely to be accessible to the biotin ligase. The *C. elegans* nuclear targeting fusion (NTF) consists of the entire NPP-9 protein fused to a tagging cassette that includes mCherry for visualization, biotin ligase recognition peptide (BLRP), a preferred substrate for BirA, and 3xFLAG for immunodetection (Fig. 1A). This fusion protein was expressed under the control of the muscle cell-specific *myo-3* promoter. *myo-3* encodes MHC A, the minor isoform of myosin heavy chain that is essential for thick filament formation (Waterston 1989). It is expressed in body-wall muscle, the sheath cells covering the hermaphrodite gonad, enteric muscle, and vulval muscles of the hermaphrodite and the diagonal muscles of the male tail. To mediate biotinylation of the NTF in vivo, we used a fusion of BirA with GFP expressed under control of the *his-72* promoter (Fig. 1B), which is expressed in multiple cell types (Ooi et al. 2006).
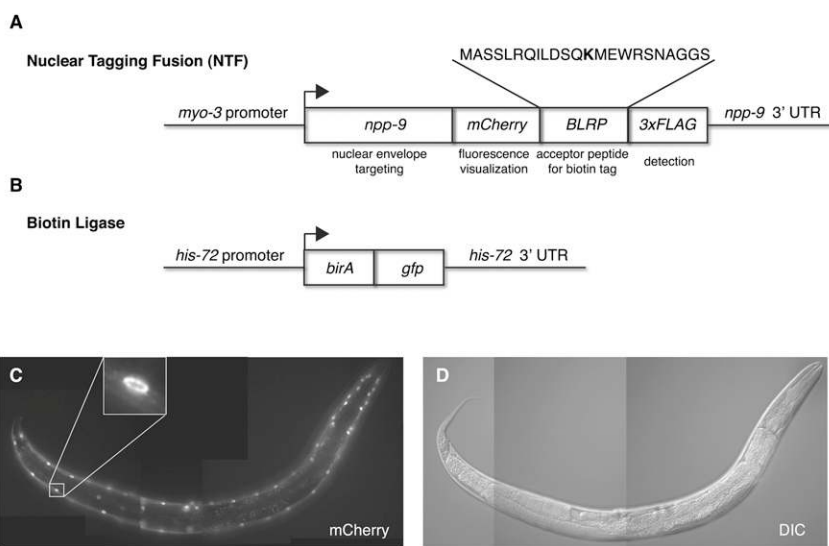
We generated strains containing these transgenes in low copy numbers using microparticle bombardment (Praitis et al. 2001). Expression of the NTF in muscle cells and its localization to the nuclei were verified by analysis of mCherry fluorescence (Fig. 1C,D). Expression is first seen in late embryos and persists through adulthood. Some cytoplasmic mCherry foci were sometimes observed, especially in aged adult animals (data not shown).

Affinity purification of nuclei from *C. elegans*

We fixed adult *C. elegans* with dimethylformamide and isolated nuclei from a strain containing both NTF and BirA and from a strain containing NTF only. In total isolated nuclei from both strains, mCherry is visible on the surface of some nuclei, confirming that the NTF remains at the nuclear envelope during the isolation procedure (Fig. 2A,B). Nuclei were incubated with streptavidin-coated magnetic beads and subsequently affinity-purified using a homemade column inserted into a Dynal Mini-MACS magnet as previously described (Deal and Henikoff 2010). Affinity purification was specific to biotinylated nuclei, as no nuclei were retained in the strain lacking expression of BirA (Fig. 2C–F). Bead-bound nuclei were positive for both mCherry and the Flag epitope (Fig. 2G–J). Starting with 1 mL of worm pellet consisting of about 500,000 adult worms, we typically obtained about 300 million input nuclei. Assuming that an adult worm contains roughly 2000 nuclei, about two-thirds of the nuclei were therefore lost during the extraction and wash procedure. Of the input nuclei, only about 4.5 million (1.5% rather than the expected 5%) were mCherry-positive, probably because muscle nuclei are less readily liberated from surrounding tissue than germline nuclei. We typically affinity-purified 1–2 million nuclei, or ~20%–40% of the mCherry-positive input nuclei. We assessed the purity of the affinity-purified
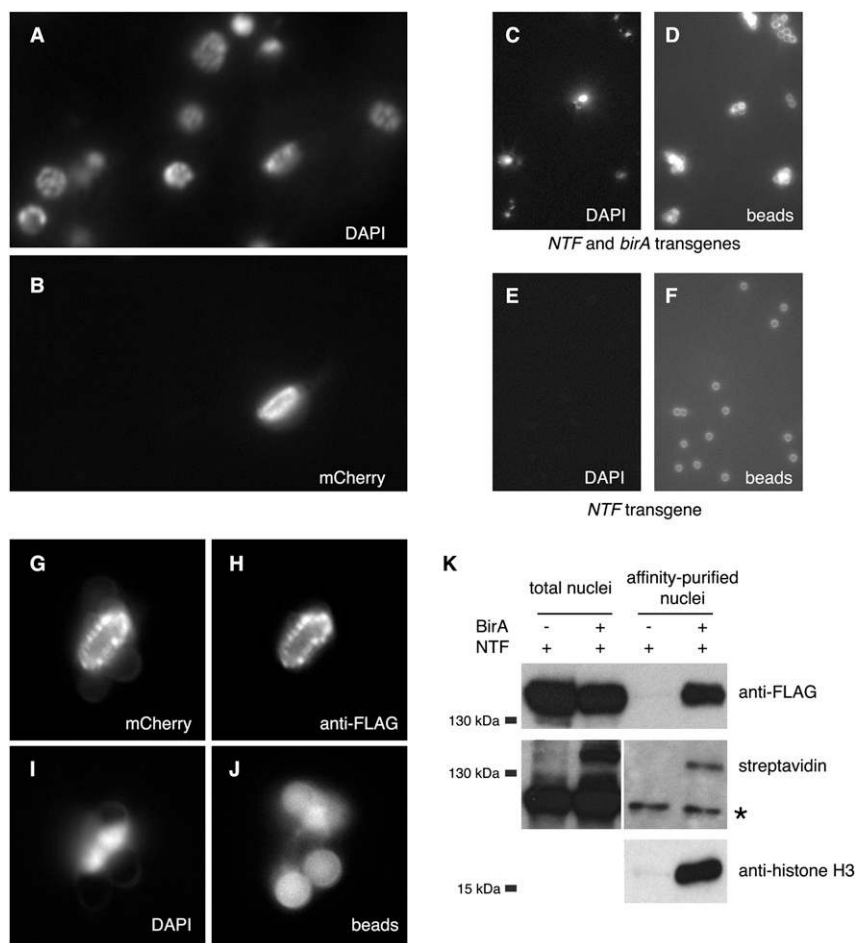
Figure 2. Affinity purification of muscle nuclei from adult *C. elegans*. (*A,B*) Total nuclei after isolation, stained with DAPI (*A*) or showing mCherry fluorescence (*B*). (*C–F*) Affinity purification of nuclei from strains containing both *NTF* and *birA* transgenes (*C,D*) or only the *NTF* transgene (*E,F*). DAPI-stained nuclei (*C,E*) and autofluorescing beads (*D,F*) are shown. (*G,H*) Example of an affinity-purified nucleus from a strain containing both *NTF* and *birA* transgenes. mCherry fluorescence (*G*), anti-Flag staining (*H*), DAPI staining (*I*), and autofluorescing beads (*J*) are shown. (*K*) Western blots detecting the NTF in total and affinity-purified nuclei, probed with an anti-Flag antibody (*top*) and streptavidin (*middle*). Streptavidin also detects an endogenous biotinylated protein of ~125 kDa that is present in all samples and serves as a loading control (asterisk). Western blot detecting histone H3 in affinity-purified nuclei (*bottom*).

soluble NTF, and that we only pulled down nuclei from strains in which the NTF was biotinylated in vivo.

A nuclear tagging strategy for *D. melanogaster*

To test the generality of our nuclear tagging system in another animal model, we transformed *D. melanogaster* with a transgene encoding a similar NTF consisting of the RANGAP protein tagged with mCherry, BLRP, and 3xFLAG (Fig. 3A). RANGAP is found in the cytoplasm and outer nuclear envelope, where it binds to NUP358, the homolog of NPP-9 (Mahajan et al. 1997; Kusano et al. 2001). The *D. melanogaster NTF* gene was expressed under the control of the *twist* (*twi*) promoter, which directs transcription in embryonic mesoderm (Thisse et al. 1988). A second *twi* promoter directs transcription of *birA* in the same cells.

Confocal imaging revealed that NTF expression was indeed mesoderm-specific. NTF was first detectable at gastrulation (stages 6–7), became conspicuous in the somites in stages 8–13 (Fig. 3B), and was still detectable at least through the end of differentiation at stage 16. mCherry staining was enriched at nuclear envelopes as expected (Fig. 3C). Biotinylation of the NTF was confirmed in fixed stage 6–7 embryos, in which streptavidin weakly detected endogenous biotinylated proteins in all cells, but robustly detected the biotinylated NTF in mesoderm in a pattern that is coincident with that detected by anti-Flag antibody (Fig. 3D–G).

NTF expression could be targeted to a large variety of cell types by placing the *NTF* and *birA* genes under the direction of a UAS sequence and making use of the large collection of cell-type-specifc GAL4 drivers in *D. melanogaster*. We generated a *UAS-NTF* line and tested expression by crossing it to *Actin 5C-GAL4*, which drives high levels of expression in larval salivary gland cells and other tissues. We found that mCherry accumulates at the nuclear envelope and in the cytoplasm of larval salivary gland cells (Fig. 3H–J), suggesting that the UAS-NTF construct is suitable for the cell-type-specific tagging of nuclei.

Affinity purification of nuclei from *D. melanogaster*

We affinity-purified mesodermal nuclei from *D. melanogaster* embryos expressing the *twi*-NTF construct using streptavidin-coated magnetic beads as described above for *C. elegans*. Clusters of streptavidin-coated beads bound to nuclei could be readily observed in samples from fresh or previously frozen 3–6.5-h-old embryos (stages 7–11). No beads with bound nuclei were observed when using nuclei from 0–2-h-old (stages 1–4) embryos, before detectable NTF expression, or nuclei from 4–6.5-h-old embryos of

nuclei by counting bead-bound and non-bead-bound nuclei or by counting mCherry-positive and mCherry-negative nuclei and observed a purity of 93% ± 3% and 96% ± 4%, respectively (scored >100 nuclei, average of three experiments each). The in vivo biotinylation of the NTF and the specificity of the affinity purification were also analyzed by Western blotting (Fig. 2K). Using an anti-Flag antibody, NTF was detected in total nuclei from both strains, but in affinity-purified samples only for the strain containing both NTF and BirA. Using streptavidin, we detected several endogenous biotinylated proteins in all samples. However, as expected, we detected biotinylated NTF only in the strain containing both NTF and BirA. A small fraction of the endogenous biotinylated proteins remained in the total nuclei preparation and bound to the streptavidin-coated beads, serving as a loading control in affinity-purified samples. We also probed affinity-purified samples for histone H3 and detected a band only for the strain containing both NTF and BirA. This shows that we indeed pulled down chromatin-containing nuclei and not just
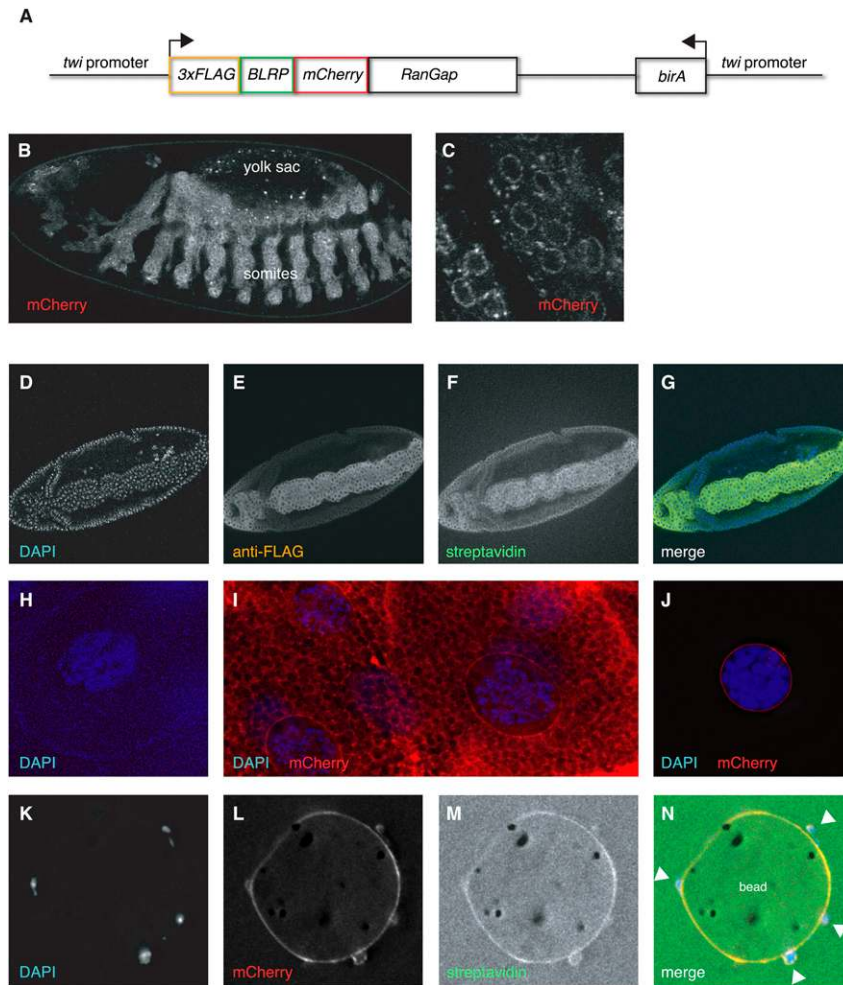
Figure 3. Nuclear tagging in *D. melanogaster*. (*A*) Schematic representation of the *D. melanogaster* nuclear targeting fusion (NTF) transgene, consisting of the mesoderm-specific *twi* promoter, the *Ran-Gap* gene tagged with *3xFLAG*, biotin ligase recognition peptide (*BLRP*), and *mCherry*. A second *twi* promoter drives transcription of *E. coli* biotin ligase gene (*birA*). (*B,C*) Confocal imaging of mCherry localization in a living stage 13 embryo. Anterior is to the *left*. (*B*) Mesodermal expression pattern is evident in somites. Some autofluorescence was also observed in particles in the yolk sac at this stage. (*C*) mCherry is present in the nuclear envelope and in the surrounding cytoplasm. (*D–G*) Detection of biotin and Flag epitopes in a fixed stage 7 embryo. (*D*) DAPI. (*E*) Anti-Flag. (*F*) Streptavidin, detecting prominent mesodermal signal from NTF and weak peripheral signal from endogenous biotinylated proteins. (*G*) Merge of *D* (blue), *E* (yellow), and *F* (green). (*H–J*) Localization of NTF to larval salivary gland nuclei in *Actin5C-GAL4; UAS-NTF* flies. (*H*) *Actin 5C/+* cells, not expressing NTF. (*I*) *Actin 5C/+; UAS-NTF/+* cells. (*J*) *Actin 5C/+; UAS-NTF/+* nucleus, broken free of cytoplasm. mCherry fluorescence is shown in red, DAPI staining in blue. (*K–N*) Affinity-purified nuclei (arrowheads) bound to an anti-Flag-coated bead. (*K*) DAPI. (*L*) mCherry. (*M*) Streptavidin. (*N*) Merge of *K* (blue), *L* (red), and *M* (green). Note that anti-Flag-coated beads are 25× larger than streptavidin-coated beads.

fluorescence and streptavidin-FITC (Fig. 3K–N), indicating that the purity of these samples is very high.

Expression analysis of affinity-purified mesoderm nuclei from *D. melanogaster*

To test whether the affinity-purified *D. melanogaster* nuclei show mesoderm-specific expression, we analyzed the transcript levels of mesoderm-specific genes in total and affinity-purified nuclei by qPCR. We and others have shown that RNA pools derived from nuclei and total cells are highly correlated (Supplemental Fig. 1; Barthelson et al. 2007; Jacob et al. 2007; Deal and Henikoff 2010), making nuclei a reasonable source for RNA for expression analysis. We determined relative transcript levels of *twi* and three other mesodermal genes—*bagpipe* (*bap*), *Myocyte enhancer factor 2* (*Mef2*), and *tinman* (*tin*)—in 3–6-h-old (stage 7–11) embryos. Expression of *bap*, *Mef2*, and *tin* genes is dependent on *twi*, peaks after 4 h, and persists beyond the 6-h end point of our collection window (Furlong et al. 2001).

A total of two purifications using anti-Flag-coated beads and three using streptavidin-coated beads were assayed. As expected for mesodermal genes, we consistently observed enrichment of *bap*, *Mef2*, and *tin* in purified nuclei, although the extent of enrichment varied between experiments. A representative data set is shown in Figure 4. Interestingly, endogenous *twi* transcript was not enriched in purified nuclei, despite the fact that the *twi* promoter drove expression of NTF and BirA. This apparent anomaly might reflect the fact that *twi* mRNA levels peak at 3 h (stages 6–7) and later decrease (Thisse et al. 1988; Thisse et al. 1991), whereas NTF protein persists through 13 h (stage 16). Therefore affinity-purified nuclei might be enriched for later-stage nuclei in which *twi* transcript levels are relatively low, whereas mRNA levels of *bap*, *Mef2*, and *tin* peak later (4 h) and thus are more likely to be enriched. The strong enrichment for these mesoderm-specific genes in affinity-purified nuclei relative to total nuclei confirms that these nuclei contain expression signatures of mesoderm nuclei and validates the method for analysis of purified, cell-type-specific *D. melanogaster* nuclei.

Expression analysis of affinity-purified muscle nuclei from *C. elegans*

Using affinity-purified nuclei from *C. elegans*, we set out to validate the method for genome-wide expression and chromatin profiling. For the expression analysis, we determined genome-wide mRNA profiles using nuclear RNA. We isolated RNA from total and

a strain that does not contain NTF, showing that purification is specific to nuclei from cells expressing NTF. Starting with 0.5–1.0 g of embryos, we obtained 170–230 million input nuclei, of which ~25% should be mesoderm nuclei (Leptin and Grunewald 1990). We affinity-purified ~5 × 10^6 to 11 × 10^6 nuclei (8%–27% theoretical yield), with 98% of nuclei bound to bead clusters.

Detection of mCherry on nuclei bound to streptavidin-coated beads is difficult due to the strong autofluorescence of these beads. We therefore affinity-purified nuclei from 3–6-h-old embryos using anti-Flag-coated magnetic beads. Ninety-five percent of these nuclei clearly contained the NTF on their nuclear envelopes (scored 200 nuclei, average of two experiments), as detected by mCherry
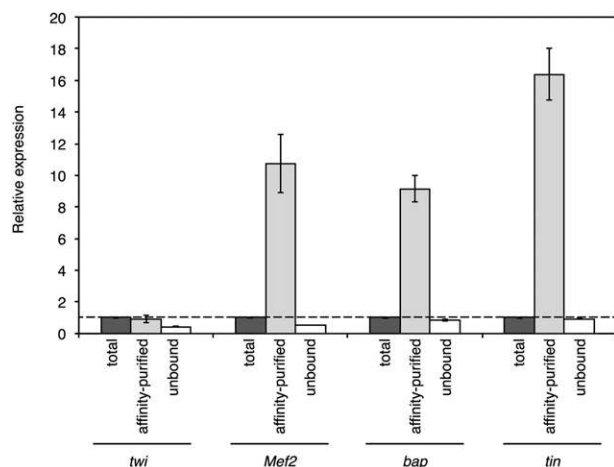
Figure 4. Relative expression of *D. melanogaster* mesodermal genes. qPCR amplification of oligo(dT)-primed *twi*, *Mef2*, *bap*, and *tin* cDNA was measured in samples from total nuclei, affinity-purified nuclei bound to anti-Flag-coated beads, and unbound nuclei. *RpL32* was used as an endogenous control. Relative expression levels were normalized to those in total nuclei. Means and standard deviations of three technical replicates are shown.

affinity-purified nuclei and generated cDNA libraries. After labeling the material with Cy dye, we hybridized it to Roche NimbleGen whole-genome tiling microarrays along with size-matched fragmented genomic DNA labeled with the complementary Cy dye. Expression scores of 20,301 genes were calculated for four biological replicates. To rule out the possibility that differences between total and nuclear RNA pools could influence the outcome of our study, we also isolated total RNA from two biological replicates and determined genome-wide expression profiles. We found that there was a strong correlation ($R = 0.96$) between total and total nuclear RNA pools (Supplemental Fig. 1), validating the use of nuclear RNA for expression analysis and confirming previously published results (Barthelson et al. 2007; Jacob et al. 2007; Deal and Henikoff 2010).

To rank genes by differential expression between total and muscle nuclei, we used the Cyber-T program (Supplemental Table 1; Baldi and Long 2001). We first compared expression of known components of the sarcomere as muscle markers (Riddle et al. 1997; Moerman and Williams 2006; Fox et al. 2007) with marker genes for the intestinal enterocyte (McGhee 2007) and for the germline (Supplemental Table 1; Hubbard and Greenstein 2000). We found that the muscle-marker genes were enriched in affinity-purified nuclei compared with total nuclei, whereas intestine- and germline-marker genes tended to be depleted (Fig. 5A). Next, we extracted WormBase annotations for the 200 most enriched and depleted genes (http://www.wormbase.org) (Supplemental Table 1). We found that 74% of the enriched genes are either exclusively or predominantly expressed in body-wall muscle or annotated as muscle-enriched (Fig. 5B). Twenty-one percent of the genes have no expression annotation, 3% are expressed or enriched in muscle as well as other tissues, and 2% are expressed or enriched in non-muscle tissues. Among the depleted genes, 77% are expressed or enriched in non-muscle tissues, for 12% the expression pattern is unknown, 10% are expressed or enriched in multiple tissues including muscle, and 1% are muscle-enriched. These annotations reflect the expected enrichment for muscle-expressed genes and the depletion for non-muscle-expressed genes in affinity-purified muscle nuclei.

In the absence of a training set for adult muscle-specific genes, we selected a cutoff of $q < 0.01$ to define a set of significantly muscle-enriched genes (Storey 2003). We assayed this set of enriched genes for overrepresentation of Gene Ontology (GO) terms (Ashburner et al. 2000). Eighty-nine GO terms were significantly overrepresented (Supplemental Table 2); the top 20 GO terms are listed in Figure 5C. Strikingly, most of these terms are associated with muscle anatomy or function, indicating that the genes with the highest relative expression are required for the specific function of this cell type within the organism. Genes associated with the structure of the sarcomere are almost completely represented in this gene set (13/17 for "M-band," 5/5 for "I-band," and 5/6 for "A-band"), confirming that this is a nearly complete set of genes with enriched expression in adult muscle tissue.

Several previous studies have identified muscle-enriched genes. Methods include FACS sorting of *myo-3*::GFP-expressing cells from dissociated embryos (Fox et al. 2007), analysis of muscle expression resulting from myogenic conversion of early blastomeres (Fukushige et al. 2006), and immunoprecipitation of tagged poly(A)-binding protein expressed by the muscle-specific *myo-3* promoter in L1 larvae (Roy et al. 2002). When comparing the results from these three studies with our set of significantly enriched genes, we found that 62% of our muscle-enriched genes were found in at least one of these studies (Fig. 5D; Supplemental Table 1), with an overlap of 45%, 45%, and 20%, respectively (Fig. 5E–G). This overlap is substantial considering that we analyzed adult worms, whereas these studies investigated stages early in development.

The expression profile of the affinity-purified muscle nuclei thus fits the characteristics of muscle tissue and validates the method for expression analysis. It also confirms the observations from the cytological and biochemical analysis that the method is highly specific, and muscle nuclei can be isolated with high purity.

To test the performance of the system, we assessed the effect of a mutation in the gene coding for the muscle-specific transcription factor UNC-120 on the muscle transcriptome in adult worms. UNC-120 is a member of the MADS-box family of transcription factors and is the single Serum Response Factor (SRF)–related protein in *C. elegans*. Its expression is limited to body-wall muscle cells and their precursors, and it is required for body-wall muscle differentiation and development (Baugh et al. 2005; Fukushige et al. 2006). *Unc-120*-null alleles are larval lethal. We therefore used the temperature-sensitive allele *unc-120(st364)* and let the worms develop to the L4 larval stage under permissive conditions before shifting them to the restrictive temperature, resulting in adult worms with severe locomotion defects. We then isolated total and muscle nuclei from this strain and determined the genome-wide expression profiles as described above. When comparing the *unc-120* mutant and wild type, we found that most genes that are down-regulated in the mutant are indeed also in the gene set that we defined as significantly muscle-enriched. However, the detection of down-regulated genes was much enhanced when analyzing affinity-purified nuclei compared with total nuclei. We found 459 genes to be significantly down-regulated in affinity-purified nuclei, compared with 77 genes in total nuclei (Fig. 6) ($q < 0.01$). This example illustrates the increased sensitivity that is gained by using our method to analyze changes in expression that affect only a limited number of cells within a whole organism.

Nucleosome occupancy profiling of affinity-purified muscle nuclei from *C. elegans*

The overall yields of nuclei obtained with our method should be sufficient for genome-wide chromatin analysis. To test this possi-
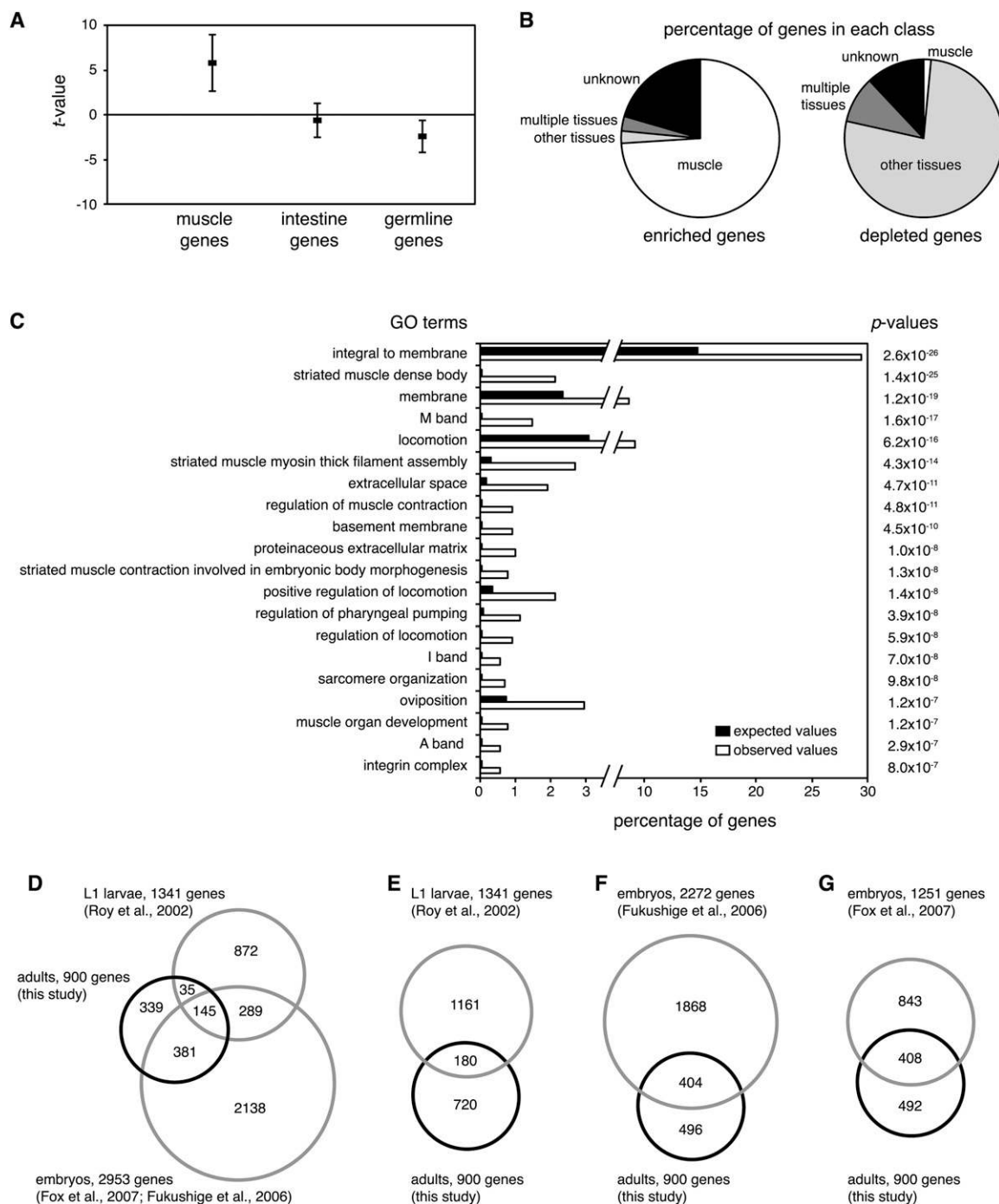
Figure 5. Expression profiling of affinity-purified *C. elegans* muscle nuclei. RNA from four independent nuclei purifications was analyzed using high-density tiling arrays. Bayesian *t*-values were determined using Cyber-T. (*A*) Enrichment or depletion of genes specific for muscle (components of the sarcomere, 50 genes), intestine (enterocyte, 76 genes), and germline (45 genes). Averages and standard deviations of *t*-values are shown. (*B*) Classification of the 200 most enriched and most depleted genes based on WormBase annotations. Genes were grouped as exclusively or predominantly expressed in muscle or muscle-enriched (white), expressed or enriched in non-muscle tissues (light gray), expressed or enriched in multiple tissues including muscle (dark gray), or without expression annotation (black). (*C*) Top 20 most significant Gene Ontology annotations of the significantly enriched genes found in this study (*q* < 0.01). Observed (white) and expected (black) percentages of genes and corrected hypergeometric *P*-values are shown. (*D*) Overlap of our set of 900 significantly enriched genes described in this study with genes previously found to be muscle-enriched in embryos and L1 larvae (Roy et al. 2002; Fukushige et al. 2006; Fox et al. 2007). (*E–G*) Overlap of our set of muscle-enriched genes with genes previously found to be muscle-enriched as in *D*, but shown separately for each study.
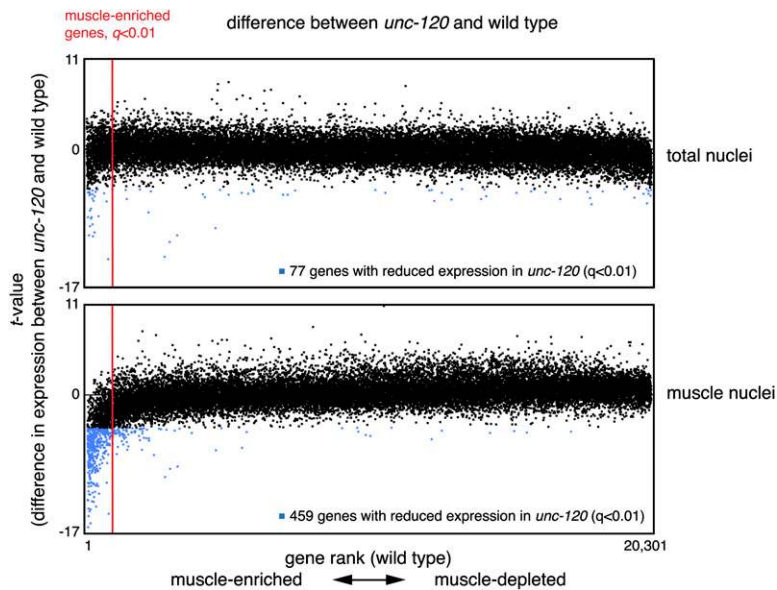
Figure 6. Nuclei purification enriches for muscle genes down-regulated in an *unc-120* mutant. Nuclei were isolated from *unc-120(st364)* and wild-type background, and RNA from total nuclei and affinity-purified muscle nuclei was analyzed using high-density tiling arrays. Averages of four biological replicates of wild type and two biological replicates of *unc-120* were used for the analysis. All 20,301 genes are ranked from muscle-enriched to muscle-depleted according to their difference in expression between total and muscle nuclei. *T*-values reflecting the differences between wild-type and *unc-120* background are plotted on the *y*-axis. The *top* panel compares total nuclei and the *bottom* panel compares muscle nuclei. Genes with significantly reduced expression in *unc-120* background are shown in blue (*q* < 0.01). The red line marks the cutoff for genes significantly enriched in muscle nuclei (*q* < 0.01).

bility, we determined nucleosome occupancy profiles of total and affinity-purified nuclei from *C. elegans* and asked whether induction of genes needed for muscle differentiation involves nucleosome loss. We treated input and affinity-purified nuclei with micrococcal nuclease (MNase), which results in protection of DNA wrapped around nucleosomes and degradation of linker DNA. As a result, regions of the genome with high nucleosome occupancy generally contain more protected fragments than regions with low nucleosome occupancy. We digested the chromatin of two biological replicates to mostly mononucleosomes, isolated the resulting protected fragments, and analyzed them by paired-end Illumina sequencing (Supplemental Fig. 2A). Alignment of genes at their 5′ and 3′ ends revealed that genic nucleosome occupancy profiles were largely overlapping in total and affinity-purified nuclei (Supplemental Fig. 2B). To determine the difference in nucleosome occupancy between total nuclei and muscle nuclei, we subtracted normalized counts of reads from total nuclei from normalized counts of reads from muscle nuclei. We created a heat map by rank-ordering genes aligned at their 5′ and 3′ ends according to decreasing difference in expression between muscle nuclei and total nuclei (Fig. 7A). For the genes that are more strongly expressed in muscle nuclei compared with total nuclei, we observed a reduction in nucleosome occupancy over promoters and gene bodies when comparing muscle nuclei with total nuclei. In other words, increases in expression are correlated with losses of nucleosome occupancy at promoters and gene bodies in muscle nuclei compared with total nuclei. This correlation holds true only for the genes that are most highly muscle-specific. The difference in nucleosome occupancy for these genes also becomes apparent when plotting the nucleosome occupancy around the 5′ and 3′ ends

of genes in total and affinity-purified nuclei (Fig. 7B, upper panel). There is no such difference in nucleosome occupancy for the non-muscle-enriched genes (Fig. 7B, lower panel).

Nuclei that are affinity-purified by the method described here can thus be used to simultaneously analyze both expression and chromatin profiles. When analyzing the interplay between gene expression and nucleosome occupancy, we identified a muscle-specific gene set that is likely to be more precise than those obtained by analyzing either expression profiles or nucleosome occupancy alone. This observation emphasizes the desirability of investigating the epigenome of cell types at multiple levels and demonstrates that our method is capable of providing such informative data sets.

Discussion

In this study, we describe a system for affinity purification of nuclei from specific cell types in animal models that can be used to define expression profiles and to analyze underlying chromatin landscapes. We used the system to isolate mesoderm nuclei from *D. melanogaster* and muscle nuclei from *C. elegans* to >90% purity. We achieved similar purity and greater overall yields than reported for *A. thaliana* epidermal cell types (Deal and Henikoff 2010). In *D. melanogaster*, we found that known mesodermal genes show greater expression in affinity-purified nuclei relative to total nuclei. In *C. elegans*, we identified a set of genes that are more highly expressed in muscle tissue and showed that a loss of nucleosome occupancy at promoters and over gene bodies accompanies this increase in expression. Our set of muscle-enriched genes mostly consists of genes previously found to be muscle-specific or muscle-enriched, and GO term analysis showed that they play a role in muscle structure or function. Moreover, genes that are down-regulated in a mutant of the muscle-specific transcription factor UNC-120 are also contained in our set of significantly muscle-enriched genes. These results confirm that the method can reliably generate accurate cell-type-specific expression and chromatin profiles.

To evaluate the performance of our method, we compared our results with muscle-enriched genes found in embryos (Fukushige et al. 2006; Fox et al. 2007) or L1 larvae (Roy et al. 2002) that were defined using different methodologies. We used the 50 sarcomere components described above as markers for known muscle genes in the comparison. Only 11 of the 50 genes were absent from our set of muscle-enriched genes, versus 12, 16, and 24 in the three previous studies, despite the fact that we called only 900 genes as muscle-enriched, versus 1251, 2272, and 1341 genes in the previous studies. Overall, 62% of the muscle-enriched genes in our data set were also found in at least one of these studies. This large overlap seems surprising, since we analyzed muscle tissues of adult worms, whereas all three previous studies analyzed stages early in development. Myoblasts arise after the end of gastrulation (Sulston et al. 1983), and the large overlap between embryonic and adult
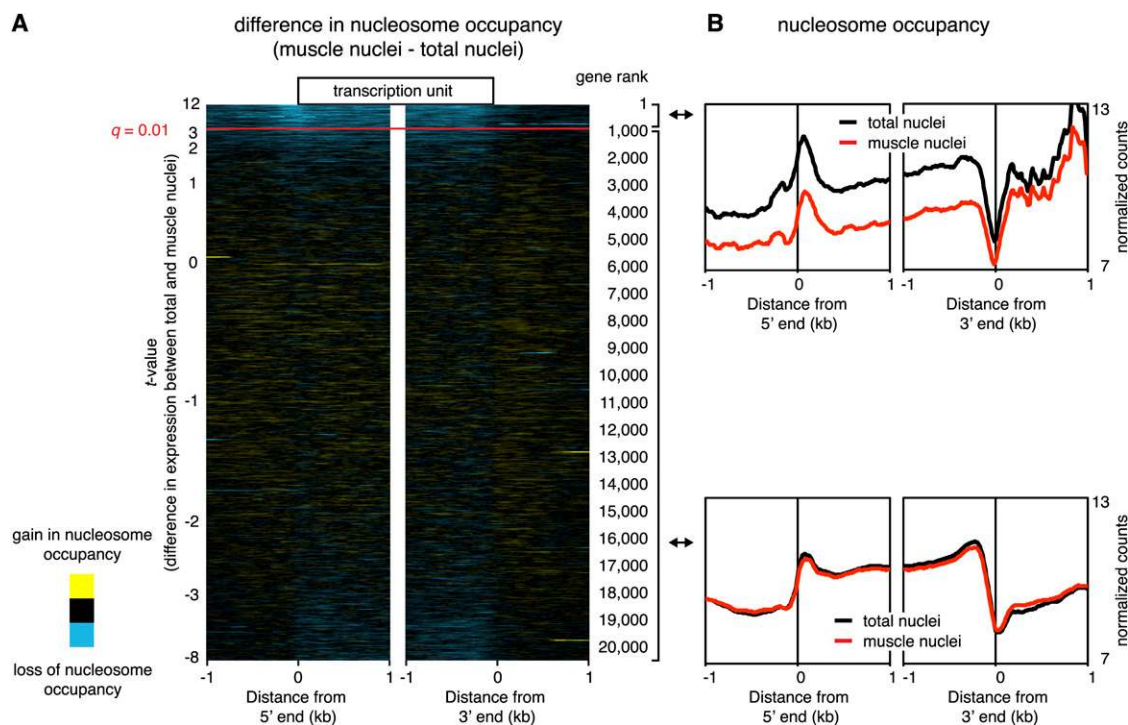
Figure 7. (*A*) Heat map showing differences in nucleosome occupancy at promoters and over gene bodies between total nuclei and muscle nuclei from *C. elegans*. Chromatin of total nuclei and affinity-purified muscle nuclei was digested to mostly mononucleosomes by MNase. Nucleosome-protected fragments were paired-end-sequenced, and the difference between normalized counts from purified nuclei and total nuclei was calculated. Genes were aligned at the 5′ end and at the 3′ end and ranked from highest to lowest difference in expression between purified nuclei and total nuclei. Yellow represents higher and blue lower nucleosome occupancy in muscle nuclei. The red line marks the *q*-value of 0.01. (*B*) Average nucleosome occupancy around the 5′ and 3′ ends for muscle-enriched genes (*q* < 0.01; *upper* panel) and non-muscle-enriched genes (*lower* panel). Profiles are shown for samples from total nuclei (black) and affinity-purified nuclei (red).

muscle expression profiles suggests that the set of genes that defines the identity of muscle cells and gives rise to proteins that perform functions specific to this cell type remains, at least in part, constant throughout development. The previous studies used FACS (Fox et al. 2007), myogenically converted total embryos (Fukushige et al. 2006), and immunoprecipitation of tagged poly(A)-binding protein (Roy et al. 2002) to determine muscle-enriched genes. It is reassuring that profiles obtained from different stages in development and using different methodologies result in expression profiles with substantial overlaps.

Regulation of gene expression is tightly linked to the regulation of the chromatin landscape. Nucleosomes serve as the basic building blocks of chromatin and therefore must be mobilized to allow for DNA transactions such as transcription. Nucleosomes are removed to allow access for transcription factors (Bernstein et al. 2004; Lee et al. 2004; Mito et al. 2005) and are evicted in cases of high-level RNA polymerase transit (McKnight and Miller 1976; Petesch and Lis 2008). We found that genes that are highly expressed in muscle cells relative to other tissues showed reduced nucleosome occupancy at promoters and over gene bodies compared with other tissues. This relationship holds true for the genes with the highest muscle-enriched expression and abruptly disappears with decreasing enrichment. This inverse relationship between nucleosome occupancy and high-level expression has been previously documented only in single-cell systems (Lee et al. 2007; Petesch and Lis 2008). Our findings confirm this relationship for a purified cell type from an animal, while also defining a set of muscle-expressed genes. This observation also validates the utility

of our method, by demonstrating that meaningful chromatin profiles can be acquired from affinity-purified nuclei.

Although it is possible to detect gene expression changes in a specific tissue by analyzing whole animals, we have demonstrated that the detection of theses changes is greatly enhanced when analyzing purified nuclei from the specific tissues. Almost six times as many genes were found to be significantly down-regulated in an *unc-120* mutant when analyzing affinity-purified muscle nuclei, compared with total nuclei. It is therefore clearly desirable to analyze specific cell types rather than whole animals for studying gene expression in a specific tissue, and our method is a useful tool in achieving this.

Our method can be easily extended to other tissues and cell types in *C. elegans* and *D. melanogaster* by exchanging the *myo-3* or *twi* promoter with different tissue- or cell-type-specific promoters. In *D. melanogaster*, this can be readily accomplished by crossing the *UAS-NTF* construct to an appropriate tissue-specific GAL4 driver. In principle, this method can be applied to all cell types within an organism. In *C. elegans*, the pattern of cell divisions within the cell lineage is essentially invariant among individuals and gives rise to 959 somatic cells in the adult hermaphrodite (Sulston and Horvitz 1977). Provided that promoters can be found that are exclusive to each of these cells, it should be possible to determine expression and chromatin profiles of each cell within a worm. Muscle cells represent ~5% of all cells in an adult *C. elegans*, and we easily obtained enough nuclei for profiling experiments starting with relatively small worm cultures. It should be possible to purify nuclei from rare cell types and individual cells within an organism by

increasing the amount of starting material. NPP-9 and RANGAP are conserved proteins in the animal kingdom, making possible the direct extension of this method to other animal models, allowing for a wide range of applications.

Methods

Worm constructs and strains

Worms were grown on peptone-rich plates seeded with Na22 bacteria as previously described (Ooi et al. 2010). Worm populations were synchronized by treating adults with sodium hypochlorite and letting embryos hatch on Na22 plates. Worms with wild-type background were grown at 20°C–23°C. *unc-120(st364)* were propagated at 15°C (permissive temperature), synchronized as described above, and shifted as L4 larvae to 23°C (restrictive temperature) for 24 h.

The *npp-9* gene and 361 bp of the *npp-9* 3′ UTR were PCR-amplified from genomic DNA and cloned into a vector containing an *unc-119* rescuing sequence. *mCherry* cDNA was PCR-amplified from pcDNA-mCherry. Biotin ligase recognition peptide (ATGGCTTCTTCTCTTCGTCAGATCCTCGACTCTCAGAAGATGGAGTGGCGTTCTAACGCTGGAGGATCT) and *3xFLAG* tag (GATTACAAGGATCACGATGGCGATTACAAGGATCACGATATCGATTACAAGGATGATGATGATAAG) sequences separated by a spacer sequence (GCGGCCGCA) were added to the *mCherry* cDNA by PCR using long oligonucleotides. The whole tagging cassette was ligated to the 3′ end of the *npp-9* gene to form a C-terminal fusion. A 2499-bp sequence upstream of the *myo-3* gene was amplified from genomic DNA and cloned upstream of the *npp-9* gene to form the plasmid pFS16. The plasmid pSO221 containing a *birA::gfp* fusion gene under the control of the *his-72* promoter was described in Ooi et al. (2010).

Transgenic lines were obtained using microparticle bombardment to transform *unc-119(ed3)* worms (Praitis et al. 2001). Strains generated for this study are JJ2285 (*zuIs262; ed3*) and JJ2286 (*zuIs263; ed3*) containing pFS16; JJ2281 (*zuIs258; ed3*) containing pSO221; JJ2289 (*zuIs258; zuIs262; ed3*); JJ2300 (*zuIs258; zuIs263; ed3*) containing both pFS16 and pSO221; and JJ2371 (*zuIs258; zuIs263; st364; ed3*).

For imaging, worms were paralyzed in M9 containing 0.1% sodium azide (Sigma-Aldrich) on a 3% agarose pad. Images were acquired on a Zeiss Axioplan microscope using the 40× lens.

Purification of affinity-tagged nuclei from adult *C. elegans*

Adult worms were washed three times in M9 buffer and collected on ice. Worms were lightly fixed in 50 mL of −20°C dimethylformadmide (Sigma-Aldrich) for 2 min and washed three times in PBS. Worms were ground to a fine powder in liquid nitrogen using a ceramic mortar and pestle. The worm powder was thawed in 5 mL of nuclei purification buffer (NPB; 10 mM Tris at pH 7.5, 40 mL NaCl, 90 mM KCl, 2 mM EDTA, 0.5 mM EGTA, 0.2 mM DTT, 0.5 mM PMSF, 0.5 mM spermidine, 0.25 mM spermine, 0.1% Triton X-100).

Worms were broken up with a glass dounce homogenizer and 30 strokes using the tight inserting pestle. Debris was spun down at 100*g* for 2 min, and the supernatant containing the nuclei was collected. More nuclei were released by resuspending the debris in NPB and light sonication at the lowest setting of a Bioruptor (Diagenode) for two cycles of 30 sec on/30 sec off. Debris was again spun down at 100*g* for 2 min, and the supernatant containing the nuclei was collected. Sonication and nuclei collection were repeated twice more. Nuclei from all collection steps were pooled and the volume brought to 45 mL with NPB. Residual debris was spun down at 100*g* for 5 min and discarded. Nuclei were collected

on a 5-mL OptiPrep density gradient medium (Sigma-Aldrich) cushion, placed at the bottom of the tube, by spinning at 1000*g* for 10 min. Nuclei were again washed twice in 45 mL of NPB and pelleted on OptiPrep cushions as described above. These washes were necessary to remove the bulk of cytoplasmic endogenous biotinylated proteins that can compete with the downstream affinity purification. Nuclei were transferred to a 15-mL tube, and the volume was adjusted to 5 mL with NPB. Aliquots of 5%–10% of these input nuclei were set aside for RNA and DNA analysis each. A 30-μL slurry of Dynabeads M-280 Streptavidin (Invitrogen) was washed with NPB and added to the nuclei. The beads were incubated with the nuclei on an end-over-end rotator for 60 min at 4°C. The mixture was subsequently diluted to 10 mL with NPB containing 0.5% Triton X-100 and drawn up into a 10-mL serological pipette. The mixture was run through a 1-mL pipette tip inserted into a MiniMacs magnet (Miltenyi Biotec) at a flow rate of ~1 mL/min. Beads and bead-bound nuclei collected at the side of the pipette tip were eluted into 1 mL of NPB. The volume was again brought up to 10 mL with NPB containing 0.5% Triton X-100, and the magnetic purification was repeated. Beads and nuclei were eluted in 1 mL of NPB and used for RNA and chromatin analysis. The pipette tips inserted into the magnet were pretreated with 1% BSA in NPB to prevent nonspecific sticking. The column purification is described in greater detail in Deal and Henikoff (2011). A small aliquot of the beads–nuclei mixture was stained with DAPI and used to assess the yield and purity of the nuclei using a Zeiss Axioplan microscope (40× lens) or a Nikon Eclipse 90i microscope (40× lens). Typical yields were 1×10^6 to 2×10^6 nuclei when starting with 1 mL of packed worm pellet (500,000 adult worms).

The Flag epitope on affinity-purified nuclei was detected using a monoclonal anti-Flag M2 antibody (Sigma-Aldrich) diluted 1:200 in NPB and a goat anti-mouse IgG-Cy5 antibody (Jackson ImmunoResearch) diluted 1:500 in NPB. Images were acquired using a Nikon Eclipse 90i microscope (100× lens).

We prepared four biological replicates of the affinity purification, testing two independent *NTF* alleles. For samples 1, 2, and 3, we used transgenic lines JJ2285 and JJ2289 carrying *NTF* allele *zuIs262*, and for experiment 4, we used transgenic lines JJ2286 and JJ2300 carrying *NTF* allele *zuIs263*, with virtually identical outcome (Supplemental Table 1). We prepared two biological replicates from the *unc-120* mutant strain JJ2371.

Western blotting

Western blotting was performed according to standard procedures. NTF was resolved on 6% and histone H3 on 18% polyacrylamide gels. The membrane was blocked and antibodies were diluted in PBS containing 0.1% Tween 20 and 5% milk for detection with anti-Flag and anti-histone H3 antibodies or with PBS containing 0.1% Tween 20 and 2% BSA for detection with streptavidin. Membranes were probed with streptavidin-HRP (GE Healthcare, 1:5000), anti-Flag M2 (Sigma-Aldrich, 1:2000), anti-histone H3 (Abcam, 1:10,000), ECL anti-mouse IgG-HRP (GE Healthcare, 1:5000), and ECL anti-rabbit IgG-HRP (GE Healthcare, 1:5000). HRP activity was detected using ECL detection reagents (Pierce).

Gene expression profiling

RNA was isolated from input and affinity-purified nuclei using the RNeasy Micro Kit (QIAGEN). RNA from total worms was isolated using the RNeasy Mini Kit (QIAGEN). DNA was removed with Turbo DNase (Ambion), and samples were cleaned up using the RNeasy Micro Kit. cDNA libraries were prepared using the Whole Transcriptome Amplification Kit (Sigma-Aldrich). Total genomic DNA used as a reference was fragmented to ~500 bp by sonication

in a Bioruptor (Diagenode) at the highest setting for six cycles of 30 sec on/30 sec off. Genomic DNA fragments and cDNA libraries were labeled with Cy3 and Cy5 dyes, respectively, as directed in the NimbleGen protocol supplied with the arrays (Roche), with the following exceptions: Strand-displacement reactions were incubated for 6 h at 37°C instead of 2 h. Thirty-four micrograms each instead of 1 μg each of Cy3- and Cy5-labeled samples were mixed, and the volume of the mixture was reduced to 12.3 μL. Samples were hybridized to custom *C. elegans* high-density NimbleGen tiling arrays C_elegans_WS170_Tiling_Iso_HX1 (Ooi et al. 2010). Hybridization and scanning were performed by the FHCRC Genomics Shared Resource.

Standard NimbleGen bi-weight-mean-adjusted $\log_2$ ratios were converted to standard deviates by subtracting the mean and dividing by the standard deviation for each probe. An expression score was calculated for each gene by averaging the standard deviate scores for all exonic probes after the 20th probe starting at the 3′ ends of the gene and moving toward the 5′ end. To find significantly enriched and depleted genes and calculate the difference in expression between total and affinity-purified nuclei, we used the program Cyber-T (Baldi and Long 2001) with the Bayesian analysis option using a window size of 101 and a confidence level of 10. We used *t*-values to rank-order genes according to enrichment or depletion.

Gene Ontology (GO) analysis was carried out using the GeneCodis 2.0 program (Carmona-Saez et al. 2007; Nogales-Cadenas et al. 2009) with a hypergeometric test and false discovery rate calculation to correct for multiple testing. Only GO terms for "Biological Process" and "Cellular Component" were selected. Sixteen genes were not recognized by the program in any format and were therefore excluded from the analysis.

Nucleosome profiling

Input nuclei and affinity-purified nuclei were pelleted by centrifugation at 1000*g* for 5 min, washed in TM2 (10 mM Tris at pH 7.5, 2 mM MgCl$_2$, 0.5 mM PMSF), and resuspended in 600 μL of TM2. Nuclei were prewarmed for 5 min at 37°C, and CaCl$_2$ to a final concentration of 2 mM, and 0.1 units (total nuclei) or 0.001 units (affinity-purified nuclei) of micrococcal nuclease (MNase; Sigma-Aldrich) was added. After 1, 2, 5, 10, and 20 min, 150-μL aliquots were taken and the reaction stopped by the addition of EDTA to a final concentration of 30 mM. Chromatin was treated with RNase and Proteinase K, and DNA was isolated with phenol:chloroform and precipitated with ethanol in the presence of glycogen. Illumina paired-end libraries were prepared for each sample. Cluster generation, followed by 25 rounds of paired-end sequencing in an Illumina Hi-Seq 2000, was performed by the FHCRC Genomics Shared Resource.

Only the input nuclei samples digested for 10 min (run 1) or 5 min (run 2) and the purified nuclei sample digested for 20 min (run 1) or 5 min (run 2) were used for subsequent analysis, because these samples contained mostly fragments protected by mononucleosomes and showed most similarity in the progress of the MNase digest (Supplemental Fig. 2A).

After processing and base-calling by Illumina software, paired-end reads were mapped to the *C. elegans* genome release WS220 using Novoalign (http://www.novocraft.com) with default parameters, except each multiple hit was mapped to one site chosen at random (Novoalign parameter -r Random). The number of inserts aligned to each 10-bp interval of the genome was counted, and the interval counts were normalized by dividing by the total number of counts for all intervals, and then scaled by multiplying by the number of bases in the genome. To select for fragments protected by nucleosomes, only inserts >140 bp were analyzed. Gene ends analysis was performed as described (Henikoff et al.

2009) using genes annotated in WormBase release WS220. Heat maps were created using Java TreeView 1.1.0 with a contrast level of 3.0.

Fly constructs and strains

Plasmid p4RG expressing the *D. melanogaster* NTF and *E. coli* BirA was constructed as follows: The *BLRP* sequence was fused to *mCherry* cDNA using long oligonucleotides as described for *C. elegans* and ligated into pRMHa3-3xFLAG-H2Av-Adh3′UTR (Bryson et al. 2010) between *3xFLAG* and *His2Av*. The 1379-bp *twi* promoter and the 2100-bp *RanGap* coding sequence were amplified from *cn bw* genomic DNA using primers 5′-AAGCTTCCGCGG CGCGCCGCATGCTGTGCTGAGGGCAGTAAATC-3′ and 5′-GGTA CCACCGGTCATTTGGTGATCTTGCTTGG-3′ for *twi*, and primers 5′-ATGCATGGCATGTCCACCTTTAACTTC-3′ and 5′-GTCGACA CAAATCGGAAAGTGGGAAA-3′ for *RanGap*, and cloned separately into pCR4-TOPO (Invitrogen). The QuikChange II site-directed mutagenesis kit (Agilent Technologies) was then used to bring the *RanGap* sequence into conformity with the corresponding sequence in FlyBase (CG9999). The *twi* promoter was ligated into pRMHa3-3xFLAG-BLRP-mCherry-H2Av-Adh3′UTR, replacing the metallothionine promoter. The *RanGap* gene was then ligated into the same plasmid, replacing *His2Av*, to form plasmid p3RG. We assembled the *E. coli birA* gene between an upstream UAS and the downstream SV40 termination sequence from pUAST (Brand and Perrimon 1993), and ligated it into CaSpeR2. The UAS was replaced with the *twi* promoter to form plasmid p2BirA. The entire *NTF* gene from *twi* promoter to *Adh* 3′ UTR was then moved from p3RG into the polylinker region of p2BirA. The resulting construct, p4RG, was transformed into *w⁻* flies by Genetic Services, Inc. The transformant line *RG2* was homozygous viable, healthy, and without developmental delays; expressed mCherry in the mesoderm; and was chosen for further characterization.

The UAS sequence for the *UAS-NTF* construct was amplified from plasmid pUAST using primers 5′-AAGCTTCCGCGGCGCGC CCATCTCTCCGGATCCAAGC-3′ and 5′-GGTTCCTTCACAAAGA TCCTCT-3′ and ligated in place of the *twi* promoters in plasmids p2BirA and p3RG to make plasmids p2UAS-BirA and p3UAS-RG. The *NTF* gene, from UAS to *Adh* 3′UTR, was then moved from p3UAS-RG into the polylinker region of p2UAS-BirA. The resulting construct, p4UAS-RG, was transformed into *w⁻* flies as above.

Purification of affinity-tagged nuclei
from *D. melanogaster* embryos

For each purification, 0.5–1 g of embryos was collected on molasses agar plates from embryo collection cages and aged at 25°C to 3–6 h after egg-laying. Embryos were dechorionated in 2.75% sodium hypochlorite and suspended in 3 mL of cold HB125 buffer (0.125 M sucrose, 15 mM Tris at pH 7.5, 15 mM NaCl, 40 mM KCl, 2 mM EDTA, 0.5 mM EGTA, 0.5 mM spermidine, 0.15 mM spermine, Roche Complete protease inhibitor cocktail) with 1.2 μL of 5 mg/mL DAPI solution, and dounce homogenized. Nuclei suspensions were filtered through Miracloth (CalBiochem), washed with 42 mL of cold HB125, and concentrated onto a 3-mL Optiprep (Sigma-Aldrich) cushion at 1000*g* for 10 min at 4°C. The supernatant and Optiprep cushion were discarded, leaving nuclei concentrated at the interface in ~1.5 mL of HB125. Isolated nuclei were incubated on an end-over-end rotator for 30 min at 4°C with either Dynabeads M-280 Streptavidin (Invitrogen) or anti-Flag M2 magnetic beads (Sigma-Aldrich) using 100 μL or 150 μL of bead slurry, respectively, per 1 g of embryos used for nuclear purification. Suspension of nuclei and beads was diluted to 10 mL with cold HB125 with added 0.1% Triton X-100, and purification of bead-bound

nuclei was carried out as for *C. elegans* with the use of HB125 containing 0.1% Triton X-100 instead of NPB.

Immunofluorescence microscopy

Expression of mCherry in live dechorionated embryos immersed in halocarbon oil 700 was monitored on an EVOS digital inverted microscope (Advanced Microscopy Group). Confocal images of live embryos were taken on a Zeiss LSM 510 META microscope. For immunofluorescence staining, embryos dechorionated in 2.75% sodium hypochlorite were fixed in a 1:1 mix of PBS, 4% formaldehyde, 0.3% Tween, and heptane for 30 min on a shaker. The aqueous phase was replaced with methanol, and embryos were shaken for 5 min at 250 rpm to burst vitelline membranes. Embryos were washed five times in methanol and rehydrated in PBS, then rinsed twice in PBST (PBS containing 0.2% Tween 20), permeabilized in PBS containing 1% Triton X-100 for 30 min, and rinsed with Starting Block (Thermo) three times before blocking for 1.5 h. Embryos were incubated overnight with anti-Flag antibody M2 (Sigma-Aldrich) diluted 1:1000 in Starting Block, then washed three times for 10 min in block, and again incubated overnight in Streptavidin Alexafluor 488 conjugate (Molecular Probes) diluted 1:500 and goat anti-mouse IgG-Cy5 (Jackson ImmunoResearch) diluted 1:100 in Starting Block. Embryos were again washed three times for 10 min with block and stained with DAPI before mounting on slides in Vectashield (Vector Labs). Slides were examined on a Zeiss LSM 510 META microscope as above.

To monitor bead binding, DAPI-stained total or purified nuclei were counted on a hemacytometer using an EVOS digital inverted microscope. Using $\sim 2 \times 10^7$ streptavidin-coated beads, we typically affinity-purified $\sim 5 \times 10^6$ to 11×10^6 nuclei from 1.7×10^8 to 2.3×10^8 total nuclei (~ 1 g of embryos). To examine bead-bound nuclei for epitope tags, DAPI-stained bead-bound nuclei isolated from embryos were placed in chamber slides on top of a magnet to immobilize the beads through subsequent manipulations. Nuclei were swelled in 0.5% sodium citrate for 10 min, fixed with 4% methanol-free formaldehyde 10–15 min, washed with PBST, and blocked with Starting Block. Expression of the NTF was detected using incubations of 1–5 h with FITC-conjugated streptavidin (Jackson ImmunoResearch) diluted 1:100 in block for nuclei bound to anti-Flag-coated beads. Stained nuclei were mounted in Vectashield and imaged on a DeltaVision RT microscope (Applied Precision).

Salivary glands were dissected in HB125 containing DAPI, mounted on a slide in HB125, and examined on a DeltaVision RT microscope.

qPCR analysis

Total RNA was prepared from purified *D. melanogaster* nuclei using the QIAGEN RNeasy Mini Kit and treated with Turbo DNAse (Ambion). Oligo(dT)-primed cDNA was synthesized using M-MLV Reverse Transcriptase (Ambion) according to the manufacturer's instructions. qPCR amplification was performed using 5 ng of reverse-transcribed cDNA with SYBR Green PCR Master Mix on an Applied Biosystems StepOne Plus instrument. Relative quantities of each transcript were calculated using the ΔΔCT method (Livak and Schmittgen 2001) with *Ribosomal protein L32* (*RpL32*) as endogenous control (LaLonde et al. 2006). Primer sequences are listed in Supplemental Table 3.

Data access

The raw microarray and Illumina sequencing data have been submitted to the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE31602.

References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. 2000. Gene Ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25:** 25–29.

Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, et al. 2006. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* **8:** 532–538.

Baldi P, Long AD. 2001. A Bayesian framework for the analysis of microarray expression data: Regularized t-test and statistical inferences of gene changes. *Bioinformatics* **17:** 509–519.

Barthelson RA, Lambert GM, Vanier C, Lynch RM, Galbraith DW. 2007. Comparison of the contributions of the nuclear and cytoplasmic compartments to global gene expression in human cells. *BMC Genomics* **8:** 340. doi: 10.1186/1471-2164-8-340.

Baugh LR, Wen JC, Hill AA, Slonim DK, Brown EL, Hunter CP. 2005. Synthetic lethal analysis of *Caenorhabditis elegans* posterior embryonic patterning genes identifies conserved genetic interactions. *Genome Biol* **6:** R45. doi: 10.1186/gb-2005-6-5-r45.

Bernstein BE, Liu CL, Humphrey EL, Perlstein EO, Schreiber SL. 2004. Global nucleosome occupancy in yeast. *Genome Biol* **5:** R62. doi: 10.1186/gb-2004-5-9-r62.

Brand AH, Perrimon N. 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118:** 401–415.

Bryson TD, Weber CM, Henikoff S. 2010. Baculovirus-encoded protein expression for epigenomic profiling in *Drosophila* cells. *Fly (Austin)* **4**. doi: 10.4161/fly.4.3.12177.

Burgemeister R. 2011. Laser capture microdissection of FFPE tissue sections bridging the gap between microscopy and molecular analysis. *Methods Mol Biol* **724:** 105–115.

Carmona-Saez P, Chagoyen M, Tirado F, Carazo JM, Pascual-Montano A. 2007. GENECODIS: A web-based tool for finding significant concurrent annotations in gene lists. *Genome Biol* **8:** R3. doi: 10.1186/gb-2007-8-1-r3.

D'Angelo MA, Hetzer MW. 2008. Structure, dynamics and function of nuclear pore complexes. *Trends Cell Biol* **18:** 456–466.

Deal RB, Henikoff S. 2010. A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev Cell* **18:** 1030–1040.

Deal RB, Henikoff S. 2011. The INTACT method for cell type-specific gene expression and chromatin profiling in *Arabidopsis thaliana*. *Nat Protoc* **6:** 56–68.

Fox RM, Watson JD, Von Stetina SE, McDermott J, Brodigan TM, Fukushige T, Krause M, Miller DM III. 2007. The embryonic muscle transcriptome of *Caenorhabditis elegans*. *Genome Biol* **8:** R188. doi: 10.1186/gb-2007-8-9-r188.

Fukushige T, Brodigan TM, Schriefer LA, Waterston RH, Krause M. 2006. Defining the transcriptional redundancy of early bodywall muscle development in *C. elegans*: Evidence for a unified theory of animal muscle development. *Genes Dev* **20:** 3395–3406.

Furlong EE, Andersen EC, Null B, White KP, Scott MP. 2001. Patterns of gene expression during *Drosophila* mesoderm development. *Science* **293:** 1629–1633.

Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K. 2009. Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Res* **19:** 460–469.

Hubbard EJ, Greenstein D. 2000. The *Caenorhabditis elegans* gonad: A test tube for cell and developmental biology. *Dev Dyn* **218:** 2–22.

Jacob Y, Mongkolsiriwatana C, Veley KM, Kim SY, Michaels SD. 2007. The nuclear pore protein AtTPR is required for RNA homeostasis, flowering time, and auxin signaling. *Plant Physiol* **144:** 1383–1390.

Kusano A, Staber C, Ganetzky B. 2001. Nuclear mislocalization of enzymatically active RanGAP causes segregation distortion in *Drosophila*. *Dev Cell* **1:** 351–361.

LaLonde M, Janssens H, Yun S, Crosby J, Redina O, Olive V, Altshuller YM, Choi SY, Du G, Gergen JP, et al. 2006. A role for phospholipase D in *Drosophila* embryonic cellularization. *BMC Dev Biol* **6:** 60. doi: 10.1186/1471-213X-6-60.

Lee CK, Shibata Y, Rao B, Strahl BD, Lieb JD. 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet* **36:** 900–905.

Lee W, Tillo D, Bray N, Morse RH, Davis RW, Hughes TR, Nislow C. 2007. A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* **39:** 1235–1244.

Leptin M, Grunewald B. 1990. Cell shape changes during gastrulation in *Drosophila*. *Development* **110:** 73–84.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{(-\Delta\Delta C(T))}$ method. *Methods* **25:** 402–408.

Mahajan R, Delphin C, Guan T, Gerace L, Melchior F. 1997. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell* **88:** 97–107.

McGhee JD. 2007. The *C. elegans* intestine (March 27, 2007) *WormBook* (ed. The *C. elegans* Research Community, Wormbook). doi: 10.1895/wormbook.1.133.1. http://www.wormbook.org.

McKnight SL, Miller OL Jr. 1976. Ultrastructural patterns of RNA synthesis during early embryogenesis of *Drosophila melanogaster*. *Cell* **8:** 305–319.

Meier I, Xu XM, Brkljacic J, Zhao Q, Wang HJ. 2008. Going green: Plants' alternative way to position the Ran gradient. *J Microsc* **231:** 225–233.

Miller MR, Robinson KJ, Cleary MD, Doe CQ. 2009. TU-tagging: Cell type-specific RNA isolation from intact complex tissues. *Nat Methods* **6:** 439–441.

Mito Y, Henikoff J, Henikoff S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* **37:** 1090–1097.

Moerman DG, Williams BD. 2006. Sarcomere assembly in *C. elegans* muscle (January 16, 2006) *WormBook* (ed. The *C. elegans* Research Community, WormBook). doi: 10.1895/wormbook.1.81.1. http://www.wormbook.org.

Neira M, Azen E. 2002. Gene discovery with laser capture microscopy. *Methods Enzymol* **356:** 282–289.

Ng RK, Gurdon JB. 2008. Epigenetic inheritance of cell differentiation status. *Cell Cycle* **7:** 1173–1177.

Nogales-Cadenas R, Carmona-Saez P, Vazquez M, Vicente C, Yang X, Tirado F, Carazo JM, Pascual-Montano A. 2009. GeneCodis: Interpreting gene lists through enrichment analysis and integration of diverse biological information. *Nucleic Acids Res* **37:** W317–W322.

Ooi S, Priess J, Henikoff S. 2006. Histone H3.3 variant dynamics in the germline of *Caenorhabditis elegans*. *PLoS Genet* **2:** e97. doi: 10.1371/journal.pgen.0020097.

Ooi SL, Henikoff JG, Henikoff S. 2010. A native chromatin purification system for epigenomic profiling in *Caenorhabditis elegans*. *Nucleic Acids Res* **38:** e26. doi: 10.1093/nar/gkp1090.

Petesch SJ, Lis JT. 2008. Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. *Cell* **134:** 74–84.

Praitis V, Casey E, Collar D, Austin J. 2001. Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157:** 1217–1226.

Riddle DL, Blumenthal T, Meyer BJ, Priess JR. 1997. Introduction to *C. elegans*. In *C. elegans II*, 2nd ed. (ed. DL Riddle, et al.), Chap. 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Roy PJ, Stuart JM, Lund J, Kim SK. 2002. Chromosomal clustering of muscle-expressed genes in *Caenorhabditis elegans*. *Nature* **418:** 975–979.

Stoeckius M, Maaskola J, Colombo T, Rahn HP, Friedlander MR, Li N, Chen W, Piano F, Rajewsky N. 2009. Large-scale sorting of *C. elegans* embryos reveals the dynamics of small RNA expression. *Nat Methods* **6:** 745–751.

Storey JD. 2003. The positive false discovery rate: A Bayesian interpretation and the q-value. *Ann Stat* **31:** 2013–2035.

Sulston JE, Horvitz HR. 1977. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev Biol* **56:** 110–156.

Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* **100:** 64–119.

Thisse B, Stoetzel C, Gorostiza-Thisse C, Perrin-Schmitt F. 1988. Sequence of the twist gene and nuclear localization of its protein in endomesodermal cells of early *Drosophila* embryos. *EMBO J* **7:** 2175–2183.

Thisse C, Perrin-Schmitt F, Stoetzel C, Thisse B. 1991. Sequence-specific transactivation of the *Drosophila twist* gene by the *dorsal* gene product. *Cell* **65:** 1191–1201.

Von Stetina SE, Watson JD, Fox RM, Olszewski KL, Spencer WC, Roy PJ, Miller DM III. 2007. Cell-specific microarray profiling experiments reveal a comprehensive picture of gene expression in the *C. elegans* nervous system. *Genome Biol* **8:** R135. doi: 10.1186/gb-2007-8-7-r135.

Waterston RH. 1989. The minor myosin heavy chain, mhcA, of *Caenorhabditis elegans* is necessary for the initiation of thick filament assembly. *EMBO J* **8:** 3429–3436.

Yuan G, Zhu B. 2011. Histone variants and epigenetic inheritance. *Biochim Biophys Acta* doi: 10.1016/j.bbagrm.2011.06.007.