Transcriptional regulation of gene expression in *C. elegans* [*]

Valerie Reinke[1§], Michael Krause[2§], Peter Okkema[3§]

[1]Department of Genetics, Yale University, New Haven, CT 06520

[2]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

[3]Department of Biological Sciences, University of Illinois Chicago, Chicago, IL 60607

Table of Contents

[§]To whom correspondence should be addressed. E-mail: valerie.reinke@yale.edu, michaelkr@niddk.nih.gov, or okkema@uic.edu

Abstract

Protein coding gene sequences are converted to mRNA by the highly regulated process of transcription. The precise temporal and spatial control of transcription for many genes is an essential part of development in metazoans. Thus, understanding the molecular mechanisms underlying transcriptional control is essential to understanding cell fate determination during embryogenesis, post-embryonic development, many environmental interactions, and disease-related processes. Studies of transcriptional regulation in *C. elegans* exploit its genomic simplicity and physical characteristics to define regulatory events with single-cell and minute-time-scale resolution. When combined with the genetics of the system, *C. elegans* offers a unique and powerful vantage point from which to study how chromatin-associated proteins and their modifications interact with transcription factors and their binding sites to yield precise control of gene expression through transcriptional regulation.

1. Overview

Every aspect of cellular function depends on the gene products expressed in that cell. The mechanisms regulating the expression of these gene products are diverse, and they can affect each of the steps necessary to make and maintain a steady-state level of functional RNA or protein. These mechanisms include those controlling RNA synthesis, processing, and stability, and, in the case of protein coding genes, protein translation, modification and degradation. Here we focus on the regulation of RNA transcription in *C. elegans*. Transcription is a necessary first step in gene expression, and transcriptional regulation plays a central role in organismal development and evolution (Levine and Tjian, 2003; Chen and Rajewsky 2007). Indeed, the number of specific proteins involved in transcription and its regulation increases with increasing organismal complexity (Vogel and Chothia, 2006).

Many of the general characteristics that make *C. elegans* an excellent model system (invariant lineage, simple anatomy, effectiveness of RNAi, etc.) also make it an excellent system to study transcriptional regulation. However, two specific characteristics uniquely facilitate transcriptional regulation studies in the worm. First, *C. elegans* is transparent throughout its entire life cycle making it an ideal system to use fluorescent protein reporter genes to monitor gene expression in live animals with single-cell resolution. Second, the relatively compact size of the *C. elegans* genome facilitates identification of cis-acting transcriptional regulatory elements (CREs) controlling gene expression. Examples of the success of studying transcriptional control in *C. elegans* are reviewed in WormBook and elsewhere, and they include the elucidation of the transcriptional cascade controlling specification and differentiation of the gut, nervous system, and pharynx (see The *C. elegans* intestine; Neurogenesis in the nematode *Caenorhabditis elegans*; and The *C. elegans* pharynx: a model for organogenesis). These two physical characteristics are augmented by facile forward and reverse genetics in the *C. elegans* system. Mutations affecting numerous transcription factors and even CREs have provided important insights into the transcriptional control of cell fate decisions during development, such as specifying which cells undergo apoptosis (Conradt and Xue, 2005), while characterization of dosage compensation (Meyer, 2010) and the synMuv mutants (Fay and Yochem, 2007) have broadened our general understanding gene regulation at the chromatin level. These advantages have consistently provided novel mechanistic insights into the transcription regulation of gene expression.

This chapter provides an overview of RNA Pol II transcription in *C. elegans*, focusing on what we have learned to date about gene expression in the somatic cells of the animal. Although much of the content of this review is also applicable to germline transcription, readers interested in germline gene expression are encouraged to see WormBook chapters Germline genomics and Germline chromatin, as important differences exist in the mechanisms controlling expression in these two tissue types. This chapter takes a broad-strokes approach to somatic transcription while providing references to serve as entry points for those wanting to explore particular topics in more detail. A brief introduction into the basics of gene organization and regulation sets the stage for those unfamiliar with *C. elegans* gene expression. This is followed by a discussion of transcription factor function, and the chapter ends with the known roles for chromatin in *C. elegans* gene regulation. Although some basic information may duplicate that found in previous reviews of transcriptional regulation (Transcriptional regulation; Transcription mechanisms; Gaudet and McGhee, 2010; Krause, 1995; McGhee and Krause, 1997; Van Nostrand and Kim, 2011), this review aims to augment rather than supersede these alternate overviews as each has valuable information and a unique perspective on transcriptional regulation.

2. RNA Polymerase II and associated factors

The regulation of RNA Polymerase II (Pol II)-mediated transcription in *C. elegans* can be described as typical for eucaryotes. Pol II appears to act in concert with TATA Binding Protein (TBP) and TBP-Associated Factors (TAFs) at the core promoter of protein coding genes (Dantonel et al., 2000; Kaltenbach et al., 2000; Lichtsteiner and Tjian, 1993; Walker et al., 2004; see also Transcription mechanisms). As in other eucaryotyes, the large subunit of Pol II of *C. elegans*, encoded by the *ama-1* gene (Bird and Riddle, 1989), has an extended C-terminal domain (CTD) that likely serves as a binding site for protein complexes involved in co-transcriptional mRNA processing and histone modification (Figure 1; from Phatnani and Greenleaf, 2006). Active Pol II is phosphorylated on the CTD at serine 2 and 5 (Ser2P or Ser5P) of the conserved heptad repeat (YSTPSPS) and its variants, as it is in other eucaryotes (Seydoux and Dunn, 1997; Wallenfang and Seydoux, 2002; Zhang et al., 2003). The levels of Ser2P increase while the Ser5P levels decrease with transcriptional elongation in eucaryotes (Buratowski, 2009) and there is some evidence for the same in *C. elegans* (Garrido-Lecca and Blumenthal, 2010). Interestingly, many of the antibodies used to distinguish these Pol II isoforms based on heptad repeat phosphorylation epitopes yield very similar patterns by chromatin immunoprecipitation (ChIP) in *C. elegans* (Baugh et al., 2009; Pferdehirt et al., 2011). The functions of many of the other core transcription factors are similarly conserved with one of the only major differences being the absence of the negative elongation factor NELF (reviewed in Transcription mechanisms).
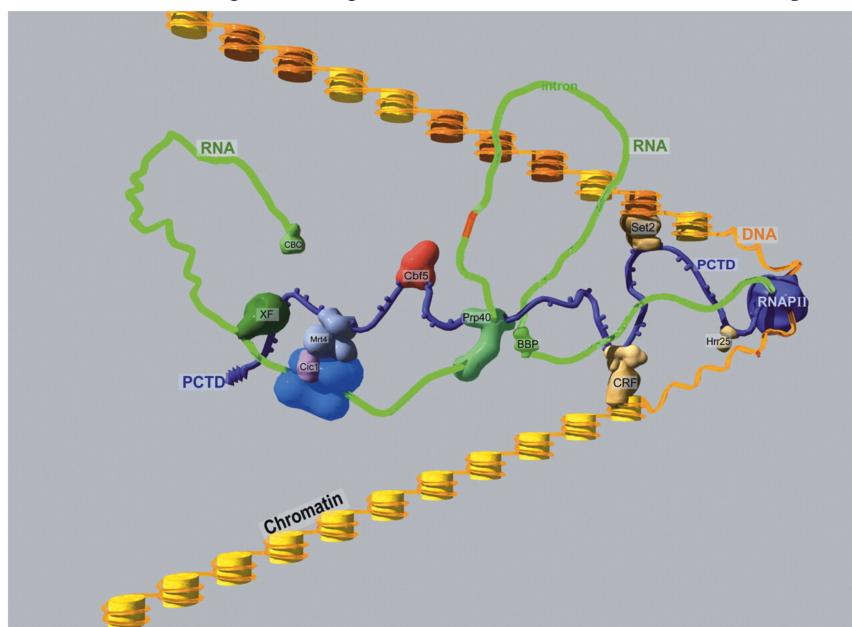
Figure 1: A hypothetical RNAPII elongation megacomplex. RNAPII (including the extended CTD with SerPO4 knobs) is purple; the globular and CTD portions are drawn approximately to scale for mammalian RNAPII. Orange DNA is wrapped around yellow nucleosomal histones; nucleosomes modified by Set2 are shaded darker. The nascent RNA transcript is green. Yeast names are used for PCAPs (e.g., Phatnani and Greenleaf, 2006), not all of which are shown. (CBC) cap-binding complex; (CRF) chromatin remodeling factor; (XF) processing/export factor. (Used with permission Phatnani and Greenleaf, 2006)

3. *C. elegans* gene organization and regulation

Transcriptional regulation results from a complex organization of cis-acting sequences that serve as binding sites for a multitude of trans-acting factors that together determine if a gene will be active or silent. In higher eucaryotes, these cis-acting sequences are typically clustered into discrete functional modules, including the core promoter, extended proximal and downstream promoter regions, positive and negative enhancers, and insulators as diagramed in Figure 2 (Levine and Tjian, 2003).
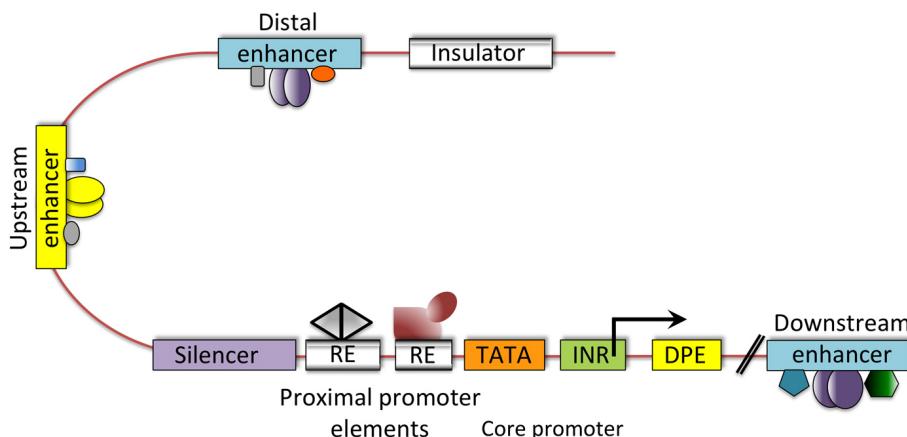
Figure 2: Metazoan regulatory modules controlling transcription. Shown is a diagram of a typical metazoan gene illustrating the complex interactions among *cis*-acting modules and *trans*-acting factors regulating gene expression. Note that both positive and negative control regions are interspersed with promoter modules, all of which can be further influenced by distal regions regulating chromatin configuration, such as insulators. (Used with permission, Levine and Tjian, 2003)

3.1. Proximal control regions

The majority of protein-coding genes in *C. elegans* are within relatively gene-dense regions of the genome. Consequently, cis-acting regulatory regions are usually close to the coding region. In fact, a good rule of thumb for *C. elegans* is that the minimal set of cis-acting sequences sufficient to regulate proper gene expression is found within 2 kb upstream of the translational start codon. Often, another gene is present on the same or opposite strand and located less than 2 kb upstream of the gene of interest. Generally in these cases, one assumes the minimal promoter is restricted to the non-coding, intergenic region. There are notable exceptions to this compact view of cis-acting sequences. For example, *egl-1* expression is controlled, in part, by elements located both upstream and more than 4 kb downstream of the coding region (Thellmann et al., 2003; Winn et al., 2011). For *lin-39*, proper reporter gene expression required inclusion of ~30 kb of genomic DNA that extends long distances upstream and downstream of the protein coding region (Wagmaister et al., 2006). Clearly *C. elegans* genes can have complex and distant control regions so the 2 kb rule of thumb should not be mistaken for dogma.

It is important to remember that the minimal promoter region is not synonymous with the natural promoter. The natural promoter may span a much larger region due to redundancy in the function of regulatory elements that ensure proper and robust regulation of the endogenous gene. One common site of additional control elements is within the introns. Most *C. elegans* introns are small (e.g., <100 bp; see Overview of gene structure) and are thus unlikely to contain elements controlling expression. However, introns larger than several hundred base pairs do often have such elements (e.g., Nam et al., 2002; Okkema et al., 1993; Kostrouchova et al., 1998). Therefore, intron size can provide a clue in searching for transcriptional control sequences.

3.2. Distal control regions

The relatively compact *C. elegans* genome may also underlie the apparent absence of long-range control mechanisms for gene regulation that are common to other metazoa. For example, CTCF in vertebrates and flies plays a key role in long-range chromatin organization and can block enhancer-stimulated gene expression, thus functioning as an insulator as shown in Figure 2 (Wallace and Felsenfeld, 2007). To date, there is no evidence for an ortholog or functional equivalent of CTCF in *C. elegans*. Thus, both the local and global organization of genes in *C. elegans* appears to be relatively simple in comparison with other metazoans, presumably simplifying our understanding of transcriptional regulation.

3.3. Transcriptional initiation and trans-splicing

Another unusual aspect of the *C. elegans* system to note when considering transcriptional regulation is the occurrence of trans-splicing, a process that replaces the initial transcript 5' untranslated region (5' UTR) with a 22 nucleotide leader sequence (SL1) for the majority of Pol II messages (Allen et al., 2011; Krause and Hirsh, 1987; Trans-splicing and operons in *C. elegans*). Therefore, mapping the 5' start site of the mature mRNA often only reveals the site of trans-splicing, not transcriptional initiation, complicating analyses that are commonplace in other

systems. The ability to trans-splice messages also provides a processing mechanism for polycistronic messages, or operons, in *C. elegans* (Spieth et al., 1993; reviewed in Trans-splicing and operons in *C. elegans*). The sequence of the spliced leader present on the mature message distinguishes the first gene of an operon (no SL or SL1 trans-spliced) versus internal messages of the operon (SL2 trans-spliced). Typically, the tightly clustered genes of an operon are co-regulated due to their polycistronic nature, although the use of different promoters upstream and within operons can result in independent transcriptional regulation of one or more of the mRNAs (Allen et al., 2011; Huang et al., 2007; Yin et al., 2010; Morton and Blumenthal, 2011; Trans-splicing and operons in *C. elegans*).

3.4. Core promoter elements

Ironically, the study of common sequences among messages trans-spliced provided the first systematic information on common promoter sequences (Graber et al., 2007), revealing the presence of a consensus Kozak sequence regulating translational initiation (Kozak, 1981). A more recent study of promoter sequences has extended this analysis to define the core promoter sequence elements typically present in *C. elegans* (Grishkevich et al., 2011). The five elements commonly observed are an Sp1 like site (CNCCGCCC), T-blocks that correlate with nucleosome eviction and gene expression levels (TTTT[N/T]), TATA box (GTATA[TA][TA]AG), trans-splicing site (TTnCAG), and Kozak site that includes the translation initiation codon ([CA]AA[CA]ATG) (Grishkevich et al., 2011) (Figure 3). The sensitivity and depth of coverage of next-generation sequencing-based techniques will allow fine scale mapping of primary transcript start sites of outron-containing genes in the near future, aiding in our understanding of transcriptional initiation and requisite regulatory sequences in the near future.
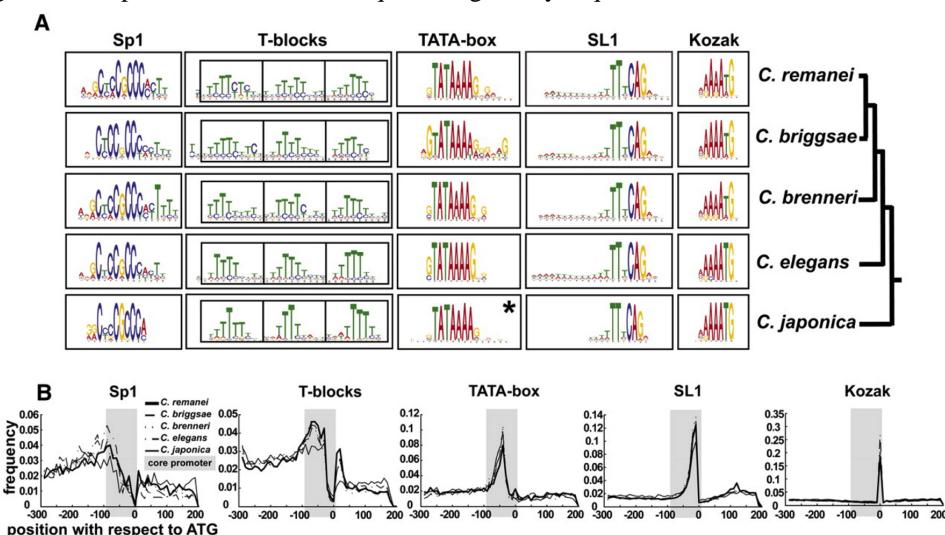

Figure 3: Core promoter motif composition among *Caenorhabditis* promoters. Motif composition of the *Caenorhabditis* core promoter. (A) Five conserved motifs in each of the five examined *Caenorhabditis* species are shown as sequence logos. (*) In contrast to all other motifs that were found in the initial search, the *Caenorhabditis japonica* TATA box motif was detected only in sequences whose orthologs contained the "TATA" motif. (B) Distribution of motifs relative to the translational start codon. The gray box in each plot corresponds to the core promoter. The area under the curve is the total frequency of occurrence within the core promoter region with the line indicating the frequency at each position as indicated by the scale to the left. The *C. japonica* SL1 motif was normalized to the length of the other species. (C) Frequency table for each sequence motif. (Figure and data used with permission, Grishkevich et al., 2011)

4. Promoter complexity

Two general types of promoter organization have been described in *C. elegans*, simple and complex. A simple promoter is defined here as one in which the cis-acting control elements necessary for proper expression are confined to a small region (a few hundreds of bp) of the genome. Housekeeping genes expressed in all tissues are good candidates for regulation by simple promoters, although few housekeeping genes in *C. elegans* have been well characterized. Among the best-characterized simple promoters are those of the *hsp-16* family of genes. This family consists of pairs of divergently transcribed genes with promoter regions sufficient for heat-regulated expression contained within the short (~350 bp) intergenic regions (Jones et al., 1986 Russnak and Candido, 1985; Stringham et al., 1992). Despite these compact promoters, distinct tissue expression patterns are induced from different *hsp-16* promoters (Stringham et al., 1992), suggesting the presence of multiple regulatory sites within these simple promoters. Another excellent example of simple promoters are in the vitellogenin (*vit*) genes, which exhibit stage-, tissue- and sex-specific expression controlled, in the case of *vit-2*, by a 247 bp promoter (MacMorris et al., 1992;

MacMorris et al., 1994). *vit-2* promoter activity depends on GATA-factor binding sites and a novel VPE2 site (TGTCAAT) conserved in *vit* gene promoters in *C. elegans* and *C. briggsae* (Spieth et al., 1985; Zucker-Aprison and Blumenthal, 1989). Certain cell cycle promoters have also been shown to be remarkably simple. Analysis of several genes expressed only in proliferative cells and encoding G1 phase regulators (e.g., cyclin D) revealed that proper regulation required as little as 67 base pairs from the promoter (Brodigan et al., 2003; Park and Krause, 1999).

In contrast to the simple promoters, complex promoters contain dispersed regulatory elements in which the overall pattern of gene expression is the result of the composite action of several dispersed elements, each influencing or contributing to the overall expression pattern. Complex promoters are often associated with regulatory genes controlling a key developmental decision. For example, the piecemeal organization of the regulatory regions for *hlh-1* and *lin-26* reflect a need for these promoters to integrate a variety of different cell lineage inputs to control proper cell fate specification in the correct cells at the appropriate time during development (see previous version of this chapter, Trancriptional regulation).

5. Transcription factors

Proper spatial and temporal regulation of gene expression depends on the binding of transcription factors to specific gene cis-regulatory sequences (Levine and Tjian, 2003). A variety of *C. elegans* transcription factors are well characterized and have important developmental roles; however it remains a challenge to accurately catalog all the transcription factors encoded in the *C. elegans* genome. Automated searches for gene ontology terms associated with transcriptional regulation can result in inclusion of false positive hits for transcription factors (Reece-Hoyes et al., 2005; Vaquerizas et al., 2009). In some cases there is ambiguity as to whether a particular domain defines the protein as a DNA-binding factor. For example, various types of zinc finger domains can bind DNA but can also serve other functions, including RNA binding and protein-protein interactions (Gamsjaeger et al., 2007; Matthews and Sunde, 2002). Likewise, for factors that modify chromatin or participate in a transcription complex, the definition of a transcription factor often lies in the eyes of the investigator. Finally, gene annotations change as new information regarding gene structure is obtained and new classes of DNA-binding transcription factors are discovered. Thus any list of transcription factors must be manually curated and periodically updated to include the latest gene annotations.

5.1. Transcription factor resources

Several groups have produced excellent catalogs of transcription factors in *C. elegans* using predictions based on gene ontology terms associated with transcription and DNA-binding domain assignments, followed by manual curation, to produce lists of transcription factor genes containing between 934 and 988 genes (Barrasa et al., 2007; Haerty et al., 2008; Reece-Hoyes et al., 2005; Wilson et al., 2008; Reece-Hoyes et al., 2011). While these lists are largely overlapping, they are not identical, so it may be useful to scan each of these lists for your favorite genes (http://edgedb.umassmed.edu/; http://www.macwormlab.net/ntfdb/index.php; http://www.transcriptionfactor.org/index.cgi?Home).

The most recent catalog is wTF2.2 (Reece-Hoyes et al., 2011), and we have updated and annotated this list with current gene names (link to Table 1. Catalog of transcription factors.). Fifty-three transcription factor genes were removed from wTF2.2 either due to reannotation of the genome, or because there was only weak evidence that motifs in these genes encoded DNA-binding domains [e.g., ZF - A20 (IPR002653), ZF - CCCH (IPR000571), ZF - DHHC (IPR001594), and ZF - MIZ(IPR004181)]. We have also added forty transcription factor genes that had been identified as DNA binding proteins in large scale yeast one-hybrid screens (Deplancke et al., 2006; Reece-Hoyes et al., 2011), or based on information from other transcription factor catalogs (Haerty et al., 2008; Wilson et al., 2008). Our catalog contains 924 transcription factor genes, which is ~4.6% of all protein coding genes (WS220). This number is slightly less than the frequency of transcription factor genes in the human genome (~6% of the protein coding genes) (Vaquerizas et al., 2009). Commercially available feeding RNAi clones targeting many of these genes are available (Table 1), facilitating functional analysis of *C. elegans* transcription factors.

5.2. Transcription factor families

C. elegans contains representatives of most major transcription factor families found in other animals (Tables 1 and 2), and 344 (37%) of the *C. elegans* transcription factor genes have been matched with clear human orthologs by reciprocal BLAST analysis or using orthology prediction programs (Table 1) (Reece-Hoyes et al., 2005; Shaye and Greenwald, 2011). Interestingly, some families of DNA-binding domains seem more highly conserved during

animal evolution than others (Figure 4). For example, 19 out of 19 (100%) MYB-like factors and 30 out of 41 bHLH factors (73%) have human orthologs. In comparison, only 6 out of 22 T-box genes (28%) and 0 out of 9 MADF-family genes (0%) have been conserved. As has been noted previously (Reece-Hoyes et al., 2005; Sluder et al., 1999), the number of likely nuclear hormone receptors (NHRs) has expanded greatly in *C. elegans* (272 members in Table 2) relative to humans (46 members; IPR001628; (Vaquerizas et al., 2009)).
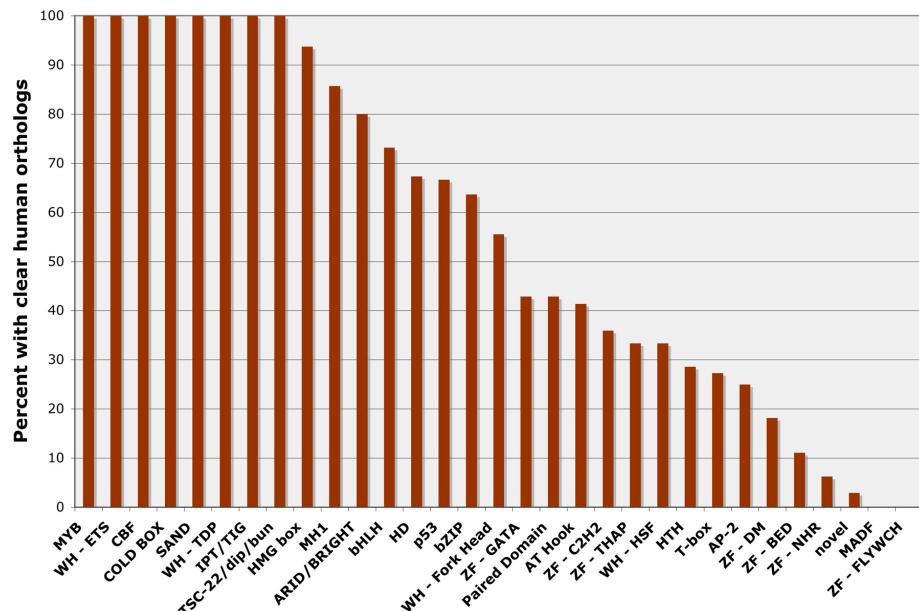
Figure 4: C. elegans transcription factor genes with clear orthologs in the human genome. The percentage of *C. elegans* genes from each DNA-binding domain family with clear orthologs in the human genome (Reece-Hoyes et al., 2005; Shaye and Greenwald, 2011). Only DNA-binding domain families with 3 or more *C. elegans* members are included.

Table 2. DNA-binding families

DNA-binding Domain	Total members
ZF - NHR	272
ZF - C2H2	217
HD	101
bHLH	41
novel	34
bZIP	33
AT Hook	29
T-box	22
MYB	19
WH - Fork Head	18
HMG box	16
ZF - GATA	14
MADF	11
ZF - DM	11
WH - ETS	10
CBF	9
ZF - BED	9
HTH	7

MH1	7
Paired Domain	7
ZF - THAP	6
ARID/BRIGHT	5
COLD BOX	5
AP-2	4
SAND	4
WH - TDP	4
ZF - FLYWCH	4
IPT/TIG	3
p53	3
TSC-22/dip/bun	3
WH - HSF	3
Brinker	2
GC-rich DNA-binding domain	2
MADS box	2
PUR	2
WH	2
ZF - NF-X1 - 10 domains	2
CP2	1
p66 family	1
RNT	1
RPEL	1
STAT	1
TEA/ATTS	1
TRAP230 family	1
WH - DAC	1
WH - RFX	1
YL1 TF	1
ZF - C2CH	1
ZF - C2HC	1

5.3. Interspecies comparisons

Genome sequencing of additional *Caenorhabditis* species allows a comparison of the *C. elegans* transcription factor gene family to those in other species (Haerty et al., 2008). *C. briggsae* and *C. remanei* contain similar numbers of transcription factor genes to *C. elegans*, and approximately 72% of the *C. elegans* transcription factor genes have detectable orthologs in both *C. briggsae* and *C. remanei*. This proportion of orthology is higher than that found overall for protein coding genes, suggesting transcription factor genes are under strong selective pressure. Transcription factor genes are not uniformly distributed on the chromosomes in *C. elegans* or *C. briggsae*, and many genes are located in clusters that are enriched for transcription factor genes compared to non-transcription factor genes. Furthermore members of gene families such as NHR, HOX, and T-box are frequently clustered in tandem arrays (Haerty et al., 2008).

5.4. Transcription factor targets

A major goal in studying transcription is to make the link between transcription factors and their target genes. These links have traditionally been made by identifying binding sites in experimentally verified targets of transcription factors by detailed promoter analyses. While this approach is still valuable, more recent techniques such as PCR based binding site selection, microarray analyses, yeast one-hybrid screens, and chromatin immunoprecipitation (ChIP) assays have expanded our ability to identify transcription factor binding sites and candidate target genes on a genome-wide scale (e.g., Deplancke et al., 2006; McElwee et al., 2003; Niu et al., 2011; Zhong et al., 2010). Our knowledge of transcription factor binding site specificity will continue to increase, but we provide references to find information about DNA binding site specificity and potential target genes for some *C. elegans* transcription factors (link to Table 3. Transcription factor targets.). As described more fully below (see Section 7), data from many of these ChIP analyses is available at modENCODE and through WormBase. Likewise candidate transcription factor binding sites predicted from published data and user generated position weight matrices can be visualized in the Genome Browser at WormBase and at modENCODE by accessing the Sequence Motif track.
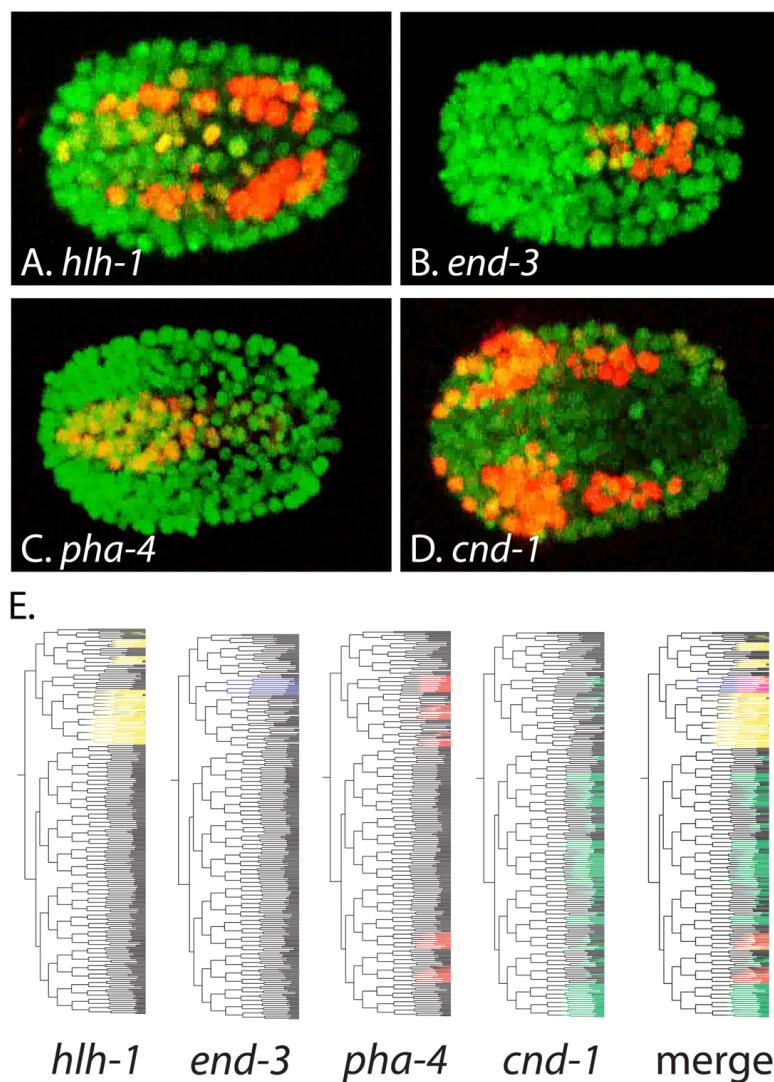
Figure 5: Automated analyses of transcription factor gene expression. (A-D) Representative frames showing expression of mCherry reporters for the indicated transcription factors (red) ovelayed on ubiquitous histone::gfp reporter expression marking all nuclei. Anterior is left. (E) Embryonic lineage trees showing expression of the indicated transcription factor::mCherry transgenes (colored), or a merged image showing expression of all four transgenes. This data was acquired and visualized using StarryNite and AceTree (Murray et al., 2008). Panels A-D were acquired from movies at http://epic.gs.washington.edu/. Panel E was adapted with permission from Murray et al., 2008.

5.5. Transcription factor gene expression

Transcription factor gene expression is often highly regulated, and understanding the spatial and temporal expression patterns of transcription factors is a key to determining their function. While a variety of transcription factors have been characterized individually using reporter genes, antibodies and in situ hybridizations (e.g., *hlh-1, pha-4, ceh-22, tbx-37, unc-86, cnd-1, end-3*) (Finney and Ruvkun, 1990; Good et al., 2004; Hallam et al., 2000; Horner et al., 1998; Kalb et al., 1998; Krause et al., 1990; Maduro et al., 2007; Okkema and Fire, 1994), high throughput techniques such as cell type-specific microarrays and SAGE analyses are providing genome-wide gene expression data (Fox et al., 2005; McKay et al., 2003; Meissner et al., 2009; Roy et al., 2002; Spencer et al., 2011; Von Stetina et al., 2007; Zhang et al., 2002). This useful data is available for transcription factor genes (and all other genes) online through WormBase and modENCODE, while additional analyses regarding tissue specificity are available at http://www.vanderbilt.edu/wormdoc/wormmap/Welcome.html (Spencer et al., 2011). In addition, automated analyses of fluorescent protein reporter gene expression are accelerating our knowledge of transcription factor gene expression during embryogenesis and in L1 larvae (Liu et al., 2009; Murray et al., 2008) (Figure 5). Lineage-based gene expression data generated using these high throughput approaches can be accessed online through EPIC and WormDB. This data is useful to investigators interested in specific transcription factors, but it also opens the exciting possibility of using computational approaches to overlay and identify correlations between gene expression patterns to understand how networks of transcription factors control development (Figure 5).

5.6. Post-transcriptional regulation of transcription factor function

Covalent post-translational modifications also play important roles in regulating transcription factor function (Tootle and Rebay, 2005), and a variety of *C. elegans* transcription factors are regulated by modifications including phosphorylation, proteolysis, ubiquitination, and SUMOylation. For example, nuclear localization and DNA-binding activity of the FoxO-family factor DAF-16 is negatively regulated by the DAF-2/insulin like signaling pathway through phosphorylation at phylogenetically conserved sites (Cahill et al., 2001; Lin et al., 2001), whereas DAF-16 protein stability is regulated by ubiquitination (Li et al., 2007). Likewise the Gli family factor TRA-1 activity is regulated by specific proteolytic cleavage and by ubiquitin mediated targeting to the proteosomal degradation pathway (Schvarzstein and Spence, 2006; Starostina et al., 2007). Both the ETS-family factor LIN-1 and the T-box factor TBX-2 require SUMOylation (Leight et al., 2005; Roy Chowdhuri et al., 2006), and at least 20 other transcription factors have been identified as possible SUMOylation targets (Kaminsky et al., 2009). SUMOylation of LIN-1 promotes interaction with MEP-1, a component of the NuRD (see Section 6.1) transcriptional repression complex (Leight et al., 2005). Transcription factor SUMOylation has also been shown to recruit *Drosophila* MEP-1 and the NuRD complex, suggesting a conserved mechanism for SUMOylation dependent transcriptional repressors (Stielow et al., 2008).

Non-covalent interactions of transcriptional co-activators or co-repressors with DNA-bound transcription factors are also crucial for transcription factor function. The transcriptional co-activator p300/CBP is important for embryonic development (Shi and Mello, 1998), while the co-repressors CtBP and SIR-2.1 play important roles regulating life span (Chen et al., 2009; Tissenbaum and Guarente, 2001). The best-studied transcriptional co-repressor in *C. elegans* is the Groucho-family factor UNC-37, which was first characterized based on its interaction with the homeodomain factor UNC-4 (Winnier et al., 1999). Not surprisingly, UNC-37/Groucho has also been shown to function with a variety of transcription factors in *C. elegans*, including POP-1, COG-1, REF-1, RNT-1 UNC-30, and MLS-1 (Calvo et al., 2001; Chang et al., 2003; Miller and Okkema, 2011; Neves and Priess, 2005; Peden et al., 2007; Xia et al., 2007). The number of transcription factors that function as UNC-37 dependent repressors is likely to be larger, as potential Groucho-interaction motifs are found in many *C. elegans* transcription factors (Copley, 2005).

6. Chromatin status and transcription

The initiation, elongation, and termination of transcription is influenced by both local and chromosome-wide chromatin configuration, and vice versa. Many excellent reviews provide in depth treatments of this topic (Berger, 2010; Yun et al., 2011; and Bannister and Kouzarides, 2011). Here, basic information that is generally applicable across species is briefly summarized prior to a discussion of chromatin regulatory complexes in *C. elegans*.

Nucleosomes, around which DNA is wound, are composed of a histone protein octamer (two each of H2A, H2B, H3, and H4) that can be post-translationally modified in a variety of ways. Typically, residues on the amino terminal tails of individual histones are modified by phosphorylation, methylation, acetylation, and ubiquitination.

The specific amino acid residue of the histone protein that is targeted, the type of modification, and the location of the nucleosome relative to the gene body, can all have effects on transcription, or be altered by transcriptional events. These effects include altering nucleosome density and changing the level of chromatin compaction to either relax or condense a region, which has been predicted to facilitate or prohibit association of different transcriptional regulatory complexes. Additionally, histone modifications can provide specific recruitment sites for different transcriptional regulatory complexes. Moreover, the number of methyl groups modifying a particular residue can have distinct effects on gene expression. As one example, monomethylation of lysine 20 of histone H4 affects transcription, while dimethylation is associated with DNA repair, and trimethylation facilitates heterochromatin formation (Balakrishnan and Milavetz, 2010). Finally, protein variants of histones, such as H2A.z or CENPA, can be incorporated into the core histone octamer, with differential effects on chromatin configuration and function as well, through regulating noncoding RNA transcription in centromeric regions (reviewed in Stimpson and Sullivan, 2010).

Over the past several years, the role of histone modifications, histone variants, and histone modifying enzymes in regulating gene expression during *C. elegans* development has become clearer. Genome-wide studies of the distribution of individual chromatin marks provide a glimpse into the complex combinatorial "codes" that are possible and that are associated with gene expression (Figure 6). Core chromatin regulatory complexes that have been studied in other organisms, such as NuRD (Section 6.1), MLL/COMPASS (Section 6.2), and PcG (Section 6.2), are also present in recognizable form in *C. elegans*. An extensive review of germline chromatin is available as a separate chapter in WormBook (Germline chromatin), so here we focus on somatic functions of these chromatin regulatory complexes.
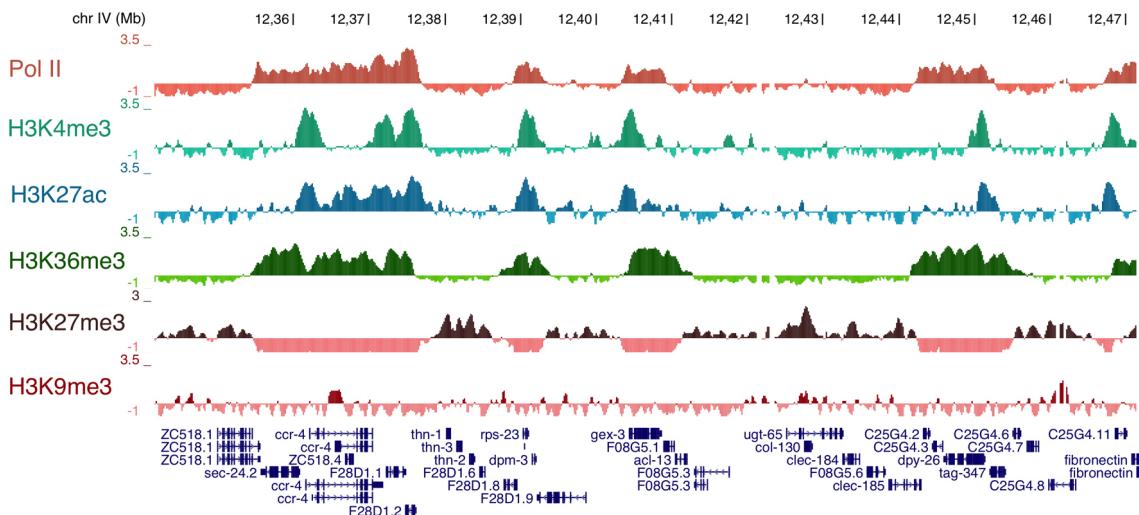

Figure 6: Examples of ChIP-chip data for RNA polymerase II (Pol II) and five different histone modifications. The X axis represents a stretch of Chr IV from nt 12,341,610 to 12,472,625. Coding genes are shown below, with arrows marking the direction of transcription 5'→3'. The Y axis represents the z-score of the log 2 ratios of IP/Input (mean centered and scaled to stdev=1). Note the opposing pattern of H3K27me, a mark associated with gene silencing, with that of activation marks such as H3K4me and H3K36me. Image courtesy of Susan Strome and Andreas Rechsteiner.

6.1. The NuRD complex

NuRD (**Nu**cleosome **R**emodeling and **D**eacetylase) is an evolutionarily conserved complex correlated with transcriptional repression that is required for proper development in mammals (reviewed in Ho and Crabtree, 2010). The NuRD complex primarily targets lysine 9 of histone H3 for deacetylation, paving the way for a histone methyltransferase (HMT) to methylate the residue. Methylated lysine 9 is then bound by HP1 (**h**eterochromatin **p**rotein 1), which induces gene silencing (Brehm et al., 1998). In *C. elegans*, NuRD consists of HDA-1 (histone deacetylase), LIN-53 (RbAp48), LIN-40 (MTA), and one of two Mi2 subunits, either LET-418 (with MEP-1) or CHD-3. The combinatorial nature of NuRD complex components allows both target gene and tissue specificity for repression. In vulval development both complexes act redundantly (von Zelewsky et al., 2000), whereas the complex containing LET-418 and MEP-1 also acts to suppress the germline fate in somatic tissues (Passannante et al., 2010; Unhavaithaya et al., 2002). The histone methyltransferases that likely methylate the deacetylated lysine 9 residue are MET-1 and/or MET-2 (Andersen and Horvitz, 2007).

6.2. The MLL/COMPASS complex

The MLL/COMPASS (**M**ixed **L**ineage **L**eukemia/**Com**plex **P**roteins **A**ssociated with **S**et1) complex is responsible for the methylation of lysine 4 of histone H3 (reviewed in Shilatifard, 2008). In contrast to lysine 9, methylation of lysine 4 is correlated with active gene expression. This complex generally contains seven subunits, all of which are present in *C. elegans*, including two MLL-like histone methyltransferases (*set-2* and *set-16*), *ash-2*, *rbbp-5*, three WDR5-like proteins (*wdr-5.1, .2,* and *.3*), *cfp-1*, and either *wdr-82* or *dpy-30* (Li and Kelly, 2011). Various components have been implicated in dosage compensation (Pferdehirt et al., 2011), vulval development (Fisher et al., 2010), neuronal development (Poole et al., 2011), and aging (Greer et al., 2010). This complex potentially functions antagonistically to NuRD/HP1 in certain aspects, such as growth and somatic gonad development (Simonet et al., 2007). Whether NuRD/HP1 and MLL/COMPASS target the same genes or distinct genes in these various developmental processes is currently unknown.

6.3. The Polycomb Group complex

The Polycomb Group (PcG) of chromatin regulators were first uncovered in the classic studies of Edward B. Lewis in *Drosophila* because of their critical roles in maintaining the repressed state of homeotic (Hox) genes regulating segmentation (reviewed in Muller and Kassis, 2006). Subsequent work by numerous groups over the years revealed that the PcG includes two types of complexes, PRC1 and PRC2. A major role for mammalian PRC2 is repression of Hox gene expression during development by promoting histone H3 lysine 27 methylation, a mark that is then bound by the silencing complex PRC1, which ubiquitylates histone H2A. The mechanism by which this modification leads to Hox gene silencing is not clear. In *C. elegans*, the PcG-related components were first identified in maternal effect sterile (MES) screens due to their role in the germline (Capowski et al., 1991), a topic explored in detail in the WormBook chapter Specification of the germ line. In the *C. elegans* soma, at least some PcG-related proteins of the PRC2 complex (MES-2 and MES-6) appear to have a role in regulating Hox gene expression, as is observed in other animals (Deng et al., 2007; Ross and Zarkower, 2003). Moreover, MES-2 is required for restricting the developmental plasticity of embryos through global changes to the chromatin state (Yuzyuk et al., 2009). Until recently, it was less clear whether *C. elegans* contained a PRC1-like complex. However, certain PRC1-related components are present including MIG-32/Bmi-1 and SPAT-3/Ring1B, which were shown to be required for H2A ubiquitylation in the soma (Karakuzu et al., 2009). Moreover, *mig-32* and *spat-3* mutants have somatic defects similar to *mes-2*, suggesting that they act to regulate the same genes. Intriguingly, however, *mig-32* and *spat-3* do not share the germline defects of *mes-2*, suggesting that the mechanism of PRC2 activity in the germline is distinct from that of the soma (Karakuzu et al., 2009).

6.4. The SynMuv pathway

Genetic studies in *C. elegans* have also uncovered a series of intertwined, genetically-linked pathways involving chromatin modifications that affect developmental pathways called the SynMuv pathway (reviewed in Fay and Yochem, 2007). Consisting of at least three genetically defined groups (A, B, and C), the SynMuv groups reveal the functional redundancy underlying gene regulation (Figure 7). Under standard conditions, genes from at least two of the SynMuv groups must be disrupted by mutation before significant effects in development occur, typically and historically read out as defects in vulva formation in the hermaphrodite. All three groups contain genes that encode gene regulatory proteins, although the SynMuv A group is unique to *C. elegans* (Davison et al., 2011). SynMuv C genes encode proteins of the TipA/HAT regulatory complex that is associated with H3K4 acetylation and active gene expression (Ceol and Horvitz, 2004). The best-studied pathway, SynMuv B, includes members of the DRM complex (Harrison et al., 2006), which is related to a repressor complex called dREAM in *Drosophila* (Korenjak et al., 2004) and DREAM/Myb-MuvB/LINC in mammals (Knight et al., 2009). In *C. elegans*, DRM includes the sequence-specific binding factor E2F called EFL-1, a pocket protein called LIN-35 and other factors with less well-understood functions, but it does not contain a Myb-like protein as in other organisms (Harrison et al., 2006). DRM potentially acts with the deacetylase NuRD complex to promote transcriptional repression, and indeed microarray studies of mutants in the SynMuv B group primarily show increased abundance of target transcripts (Kirienko and Fay, 2007). DRM components are widely expressed and appear to play diverse roles in many tissues, although in most cases the function of the DRM complex is not solely essential for normal development. However, a second mutation that disrupts a tissue-specific regulatory protein along with a mutation in a DRM complex member can cause a tissue-specific phenotype. For instance, mutation of the C2H2 zinc finger gene *slr-2* in conjunction with *lin-35* mutations disrupts intestinal function, while either mutation alone does not (Kirienko et al., 2008). Additionally, a recent report indicates that one SynMuvB component, LIN-61, which binds methylated lysine 9 of H3 (Koester-Eiserfunke and Fischle, 2011), can act as part of the DRM complex in vulval development but not in

other processes, showing that this complex has tissue-specificity (Harrison et al., 2007). Finally, environmental influences such as temperature can have an effect as well. *lin-35*, and certain other components of the DRM complex, are required to suppress the germline fate in somatic tissues (Wang et al., 2005), but only at higher temperatures is the onset of the germline fate sufficiently severe to lead to a larval arrest (Petrella et al., 2011).
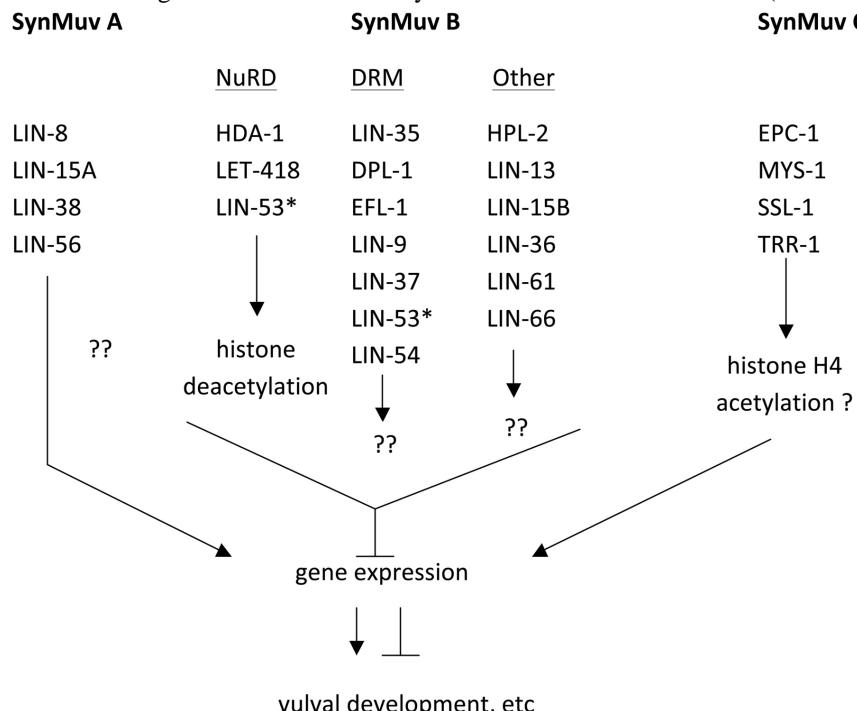

Figure 7: Outline of genes in each group of the SynMuv pathway. SynMuv A and SynMuv B are the major groups in this pathway that are redundantly required for vulval development. The relationship of these pathways with the SynMuv C pathway are less clear. Within the SynMuv B group, LIN-53, marked with an asterisk, is listed twice, as it is found in both NuRD and DRM complexes.

As in other systems, the landscape of chromatin modifications and the complexes carrying them out are generally conserved in *C. elegans*. Thus, it is likely that information from any one experimental system will inform our general understanding of chromatin influences in all systems. In the future investigating how these different chromatin regulatory complexes interact at common target loci, and how they influence key sequence-specific transcription factors and the core transcriptional machinery, will be critical for understanding gene regulatory mechanisms. The molecular and genetic advantages of *C. elegans*, combined with a relatively simple and defined cell lineage, suggest that the worm will prove particularly important for understanding the role of chromatin in developmental processes.

7. Systematic genome-scale analysis of transcription regulation

Genome-wide analyses complement detailed single-gene studies by providing a global overview that can be used to determine how broadly observations of transcriptional regulatory mechanisms at individual genes are applicable. Over the last decade microarray analysis of gene expression has been widely used to examine gene expression changes upon perturbation of various transcriptional components. However, one limitation to this type of analysis is the inability to distinguish direct from indirect effects. The more recent development of techniques such as chromatin immunoprecipitation (ChIP), which maps the binding events of a given factor throughout the genome, can overcome this limitation and provide important information about the direct activity of the factor. Chromatin fragments that immunoprecipitate with chromatin or transcriptional regulatory proteins are identified by either hybridization to a microarray (ChIP-chip) or by deep sequencing (ChIP-seq). This approach has proved to be a very powerful tool for investigating and discovering transcriptional mechanisms. In *C. elegans*, ChIP studies have been performed by individual labs focused on particular processes or factors including dosage compensation components (Ercan et al., 2007; Jans et al., 2009), the DRM complex component LIN-54 (Tabuchi et al., 2011), the histone variant HTZ-1 (Whittle et al., 2008), and transcription factors such as HLH-1 and NFI-1 (Lei et al., 2010; Whittle et al., 2009).

7.1. The modENCODE project

In addition to individual efforts, a large-scale effort by a multi-lab consortium has systematically utilized a genomics approach to exploring *C. elegans* gene expression. The modENCODE consortium (**mod**el organism **Enc**yclopedia **o**f **D**NA **E**lements), funded by the National Human Genome Research Institute (NHGRI), has in the last few years produced a wealth of genome-wide *C. elegans* datasets. These studies explore many aspects of transcriptional regulation such as transcription factor binding sites, chromatin modifications, and gene expression analysis of diverse RNAs, including small noncoding RNAs in addition to polyadenylated RNAs (Gerstein et al., 2010). A similar effort to analyze the *Drosophila melanogaster* genome is ongoing as well (Roy et al., 2010). The ultimate goal of these projects is to identify, as comprehensively as possible, all of the functional elements encoded in the DNA that are responsible for the regulation and formation of that organism.

A major effort of modENCODE is to determine the binding sites of sequence-specific transcription factors genome wide. These elements direct the temporal and spatial control of transcription, which in turn dictates an organism's development, physiology and response to the environment. Identifying these elements provides an important first step toward understanding how DNA sequence is interpreted to form a three-dimensional body plan. At the beginning of the modENCODE project, almost nothing was known about the direct targets of transcription factors in *C. elegans*. By the time the project completed its fifth year in March 2012, the genome-wide binding profiles of over 120 factors from diverse families of transcriptional regulators were collected and released to the *C. elegans* community.

7.2. modENCODE transcription factor ChIP studies

To systematically identify binding sites, transgenic lines expressing GFP-tagged transcription factors were subjected to ChIP-seq using an antibody to GFP (Zhong et al., 2010). The GFP expression patterns of these lines largely recapitulated known endogenous expression patterns, and all lines that were tested robustly rescued the mutant phenotype of that gene, indicating that the tagged transcription factors retain wild-type function. All of the detailed protocols and datasets are freely available at http://www.modencode.org/. Moreover, all the strains for which successful datasets have been produced are available in the Caenorhabditis Genetics Center (CGC). So far, 77 completed datasets are available representing 46 transcription factors, some of which have been analyzed at multiple developmental stages (Table 4).

Table 4. Expression of transcription factors.

TF genes	stage(s)	Histone modifications	stage(s)	Chromatin proteins	stage(s)
alr-1	L2	total H3	EE, L3	HPL-2	LE
aly-2	L1, L3	H3K27ac	EE, L3	LIN-15B	LE
ama-1	EE, LE, MxE, L1, starved L1, L2, L3, L4, L4/YA, YA	H3K27me1	EE, L3	pol2	EE, LE
blmp-1	L1	H3K27me3	EE, L3	ama-1	MxE
C01B12.2	L2	H3K36me1	EE, L3	HCP-3	EE, MxE
ceh-14	L2	H3K36me2	EE	CBP	MxE
ceh-26	LE	H3K36me3	EE, LE, L1, L2, L3, L4, YA	HTZ-1	MxE
ceh-30	L3	H3K4me1	EE, L3	DPY-26	MxE
ces-1	L1, L4	H3K4me2	EE, L3	DPY-27	MxE, EE, L4
daf-12	L3, L4	H3K4me3	EE, LE, L1, L2, L3, L4, YA	DPY-28	MxE
daf-16	L4	H3K79me1	MxE, L3	LEM-2	MxE
dpl-1	L1, YA	H3K79me2	EE, MxE, L3	MES-4	EE

14

dpy-27	MxE	H3K79me3	EE, MxE, L3	MIX-1	MxE
efl-1	L1, YA	H3K9me1	EE, L3	MRG-1	EE
egl-27	L1	H3K9me2	EE, L3		
egl-5	L3	H3K9me3	EE, L3		
elt-3	L1	total H4	L3		
eor-1	L3	H4acTetra	EE		
F16B12.6	L1	H4K16ac	EE		
F45C12.2	L3	H4K20me1	EE, LE, L3		
fos-1	L1, L2	H4K8ac	EE, L3		
gei-11	L2, L4				
hlh-1	MxE				
hlh-8	L3				
lin-11	L2				
lin-13	MxE				
lin-15B	L3, L4				
lin-39	L3				
mab-5	L3				
mdl-1	L1				
mef-2	L1				
mep-1	MxE				
nhr-105	L3				
nhr-111	L2				
nhr-28	L3, L4				
nhr-6	L2				
pes-1	L4				
pha-4	MxE, LE, L1, starved L1, L2, YA				
pqm-1	L3				
R02D3.7	L2				
sea-2	L3				
skn-1	L1				
unc-130	L1				
unc-62	L1, L2, L3				
W03F9.2	L4/YA				
zag-1	L1, L2, L3				
EE = early embryo, LE= late embryo, MxE= mixed stage embryos, L1-L4 = larval stages 1-4, YA = young adult					

This pipeline was first utilized on the FoxA transcription factor PHA-4 that has important roles in both organ development and environmental responses (Zhong et al., 2010). Subsequently, an analysis of the major characteristics of 22 transcription factors describes binding site features and correlations for this larger set, including a preliminary regulatory network analysis (Niu et al., 2011). Most of these factors bind thousands of sites in the genome, and the majority of these binding sites near coding genes lie within 500 bp of the predicted transcript start site. A significant insight from the properties of these 22 transcription factors was the recognition that the genome

contains hundreds of regions that are broadly permissive for non-specific binding by transcription factors (termed high-occupancy target, or "HOT", regions) (Gerstein et al., 2010). Recruitment to a HOT region does not require the sequence-specific binding property of the transcription factor, nor is binding correlated with the regulation of nearby genes. How transcription factors are recruited to HOT sites, and the possible role of chromatin or nuclear organization in this process, is unknown. For those interested in using these data to understand gene regulation by a specific transcription factor, HOT sites should be carefully distinguished from other binding sites that are either unique to or primarily bound by the factor of interest, which are more likely to result in direct regulatory events. Another limitation users should realize when interpreting the genome-wide transcription factor binding profiles is that almost all the experiments were performed in whole animals. Any binding profile for a broadly-expressed transcription factor will therefore be an amalgam of binding in multiple tissues. Currently, tissue-specific profiling techniques are being applied to circumvent this problem.

7.3. modENCODE chromatin modification ChIP studies

In addition to sequence-specific transcription factors, a second modENCODE project has focused on collecting genome-wide ChIP-chip datasets for various chromatin-associated factors as well as for histone modifications. Over 20 histone modifications and 14 factors have been analyzed to date. Unlike the transcription factor studies that analyzed transgenic GFP-tagged proteins these datasets were generated by using antibodies specific to each factor or modification to monitor their endogenous distribution; many of these antibodies are commercially available. This global analysis of chromatin states has yielded various insights, including the persistence of the germline-established chromatin state in the somatic tissues, highlighted by the existence of chromosomal domains enriched for repressive histone modifications that correlate with increased meiotic recombination rates (Gerstein et al., 2010). Additionally, the X chromosome exhibits several distinctive features relative to autosomes, including increased monomethylation of H4K20 and H3K27, and increased nucleosome density (Ercan et al., 2011; Liu et al., 2011). Importantly, when the chromatin modification status was combined with transcription factor data in predictive algorithms for gene regulatory events, the accuracy of the resulting predictions was greatly improved compared to either alone (Gerstein et al., 2010).

7.4. modENCODE transcription studies

All of these ChIP-based studies are complemented by modENCODE-based analysis of gene expression at many different developmental stages, tissues, and environmental conditions through the use of deep sequencing and tiling microarrays to monitor transcript identity and abundance. As of October 2011, more than 130 experiments have been analyzed and released to the community. These include analysis of both poly-A-selected RNAs as well as small RNAs. Most of these data have been primarily analyzed with the immediate goal of improving gene annotation, and have led to the identification of thousands of novel exons and splicing events, new small RNAs, including many microRNAs, and improved 5' and 3' UTR mapping (Hillier et al., 2009; Kato et al., 2009; Mangone et al., 2010; see also Jan et al., 2011). Improved gene models lead to improved assignment of regulatory events to the correct target genes. Moreover, as additional gene expression data is collected and analyzed, more conclusions will be drawn regarding correlations between regulatory factors and changes in gene expression levels across stages, tissues, and conditions.

Countless biological discoveries are embedded in these deep and complex modENCODE datasets. The published global observations and analyses (Gerstein et al., 2010) are just the beginning. Investigation of the data by the larger research community, with their specialized expertise in so many aspects of *C. elegans* biology, is essential to plumb the full possibilities of the data. To facilitate such endeavors by the *C. elegans* community, all modENCODE data, along with detailed descriptions of growth and collection conditions and protocols, are available at http://www.modencode.org/. The data are available for bulk download for large-scale analyses and comparisons, but the data for individual experiments or individual genes can be examined as well, for those with a particular focus on a single pathway or process. Data of interest can be selected as "tracks" for viewing on a genome browser as well. In the near future, additional changes will be made to the interface to improve selection and analysis of all interested users. Ultimately, all the data on the modENCODE website will be incorporated into WormBase. Movement of these data to cloud-based storage to increase accessibility and facilitate downloads is also a likely possibility in the near future.

8. Future Prospects

There is little doubt that the field of transcriptional regulation in *C. elegans* is in the midst of an information explosion. We are rapidly acquiring information concerning temporal and spatial patterns of gene expression using

genome wide expression assays and automated analyses of fluorescent protein reporter expression. At the same time, we are identifying binding sites for transcription factors and chromatin regulatory factors throughout the genome, and recently developed techniques for isolating specific nuclei will enhance our ability for tissue specific chromatin profiling (Deal and Henikoff, 2010). Still, we have only begun to explore other areas of transcriptional regulation. How does higher order organization within nuclei affect gene expression (Meister et al., 2010; Ikegami et al., 2010)? What impact do post-transcriptional and post-translational mechanisms have on transcription factor activity? Overall, it is an exciting time to study transcriptional regulation in *C. elegans*. Because of the relative simplicity of *C. elegans* gene promoters, we can reasonably make connections between transcription factors and their target genes. Understanding this information will help decipher how the information in the genome controls every aspect of *C. elegans* biology.

9. Acknowledgements

10. References

Allen, M.A., Hillier, L.W., Waterston, R.H., and Blumenthal, T. (2011). A global analysis of *C. elegans* trans-splicing. Genome Res. 21, 255-264. Abstract Article

Amin, N.M., Shi, H., and Liu, J. (2010). The FoxF/FoxC factor LET-381 directly regulates both cell fate specification and cell differentiation in *C. elegans* mesoderm development. Development *137*, 1451-1460. Article

An, J.H., and Blackwell, T.K. (2003). SKN-1 links *C. elegans* mesendodermal specification to a conserved oxidative stress response. Genes Dev. *17*, 1882-1893. Article

Andersen, E.C., and Horvitz, H.R. (2007). Two *C. elegans* histone methyltransferases repress *lin-3* EGF transcription to inhibit vulval development. Development *134*, 2991-2999. Abstract Article

Ao, W., Gaudet, J., Kent, W.J., Muttumu, S., and Mango, S.E. (2004). Environmentally induced foregut remodeling by PHA-4/FoxA and DAF-12/NHR. Science *305*, 1743-1746. Article

Arata, Y., Kouike, H., Zhang, Y., Herman, M.A., Okano, H., and Sawa, H. (2006). Wnt signaling and a Hox protein cooperatively regulate *psa-3*/Meis to determine daughter cell fate after asymmetric cell division in *C. elegans*. Dev. Cell *11*, 105-115. Article

Balakrishnan, L., and Milavetz, B. (2010). Decoding the histone H4 lysine 20 methylation mark. Crit. Rev. Biochem. Mol. Biol. *45*, 440-452. Abstract Article

Bannister A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381-395. Abstract Article

Barrasa, M.I., Vaglio, P., Cavasino, F., Jacotot, L., and Walhout, A.J. (2007). EDGEdb: a transcription factor-DNA interaction database for the analysis of *C. elegans* differential gene expression. BMC Genomics *8*, 21. Abstract Article

Baugh, L.R., Demodena, J., and Sternberg, P.W. (2009). RNA Pol II accumulates at promoters of growth genes during developmental arrest. Science *324*, 92-94. Abstract Article

Beaster-Jones, L., and Okkema, P.G. (2004). DNA binding and *in vivo* function of *C.elegans* PEB-1 require a conserved FLYWCH motif. J. Mol. Biol. *339*, 695-706. Article

Berger, M.F., Philippakis, A.A., Qureshi, A.M., He, F.S., Estep, P.W., 3rd, and Bulyk, M.L. (2006). Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. Nat. Biotechnol. *24*, 1429-1435. Article

Berger S.L. (2010). Cell signaling and transcriptional regulation via histone phsophorylation. Cold Spring Harb. Symp. Quant. Biol. *75*, 23-26. Abstract Article

Bertrand, V., and Hobert, O. (2009). Linking asymmetric cell division to the terminal differentiation program of postmitotic neurons in *C. elegans*. Dev. Cell *16*, 563-575. Article

Bird, D.M., and Riddle, D.L. (1989). Molecular cloning and sequencing of *ama-1*, the gene encoding the largest subunit of *Caenorhabditis elegans* RNA polymerase II. Mol. Cell. Biol. *9*, 4119-4130. Abstract

Blackwell, T.K., Bowerman, B., Priess, J.R., and Weintraub, H. (1994). Formation of a monomeric DNA binding domain by Skn-1 bZIP and homeodomain elements. Science *266*, 621-628. Article

Blackwell, T. K. and Walker, A. K. Transcription mechanisms (September 5, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.121.1, http://www.wormbook.org. Article

Blumenthal T., Trans-splicing and operons in *C. elegans* (November 20, 2012), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.5.2, http://www.wormbook.org. Article

Brehm, A., Miska, E.A., McCance, D.J., Reid, J.L., Bannister, A.J., and Kouzarides, T. (1998). Retinoblastoma protein recruits histone deacetylase to repress transcription. Nature *391*, 597-601. Abstract Article

Brodigan, T.M., Liu, J., Park, M., Kipreos, E.T., and Krause, M. (2003). Cyclin E expression during development in *Caenorhabditis elegans*. Dev. Biol. *254*, 102-115. Abstract Article

Broitman-Maduro, G., Maduro, M.F., and Rothman, J.H. (2005). The noncanonical binding site of the MED-1 GATA factor defines differentially regulated target genes in the *C. elegans* mesendoderm. Dev. Cell *8*, 427-433. Article

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Mol. Cell *36*, 541-546. Abstract Article

Cahill, C.M., Tzivion, G., Nasrin, N., Ogg, S., Dore, J., Ruvkun, G., and Alexander-Bridges, M. (2001). Phosphatidylinositol 3-kinase signaling inhibits DAF-16 DNA binding and function via 14-3-3-dependent and 14-3-3-independent pathways. J. Biol. Chem. *276*, 13402-13410. Abstract Article

Calvo, D., Victor, M., Gay, F., Sui, G., Luke, M.P., Dufourcq, P., Wen, G., Maduro, M., Rothman, J., and Shi, Y. (2001). A POP-1 repressor complex restricts inappropriate cell type-specific gene transcription during *Caenorhabditis elegans* embryogenesis. Embo J. *20*, 7197-7208. Abstract Article

Capowski, E.E., Martin, P., Garvin, C., and Strome, S. (1991). Identification of grandchildless loci whose products are required for normal germ-line development in the nematode *Caenorhabditis elegans*. Genetics *129*, 1061-1072. Abstract

Carroll, A.S., Gilbert, D.E., Liu, X., Cheung, J.W., Michnowicz, J.E., Wagner, G., Ellenberger, T.E., and Blackwell, T.K. (1997). SKN-1 domain folding and basic region monomer stabilization upon DNA binding. Genes Dev. *11*, 2227-2238. Article

Ceol, C.J., and Horvitz, H.R. (2004). A new class of *C. elegans* synMuv genes implicates a Tip60/NuA4-like HAT complex as a negative regulator of Ras signaling. Dev. Cell *6*, 563-576. Abstract Article

Chang, S., Johnston, R.J., Jr., and Hobert, O. (2003). A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans*. Genes Dev. *17*, 2123-2137. Abstract Article

Chen, K., and Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. Nat. Rev. Genet. *8*, 93-103. Abstract Article

Chen, S., Whetstine, J. R., Ghosh, S., Hanover, J.A., Gali, R.R., Grosu, P., and Shi, Y. (2009). The conserved NAD(H)-dependent corepressor CTBP-1 regulates *Caenorhabditis elegans* life span. Proc. Natl. Acad. Sci. U. S. A. *106*, 1496-1501. Abstract Article

Choi, K.Y., Ji, Y. J., Jee, C., Kim do, H., and Ahn, J. (2002). Characterization of CeHDA-7, a class II histone deacetylase interacting with MEF-2 in *Caenorhabditis elegans*. Biochem. Biophys. Res. Commun. *293*, 1295-1300. Article

Christensen, S., Kodoyianni, V., Bosenberg, M., Friedman, L., and Kimble, J. (1996). *lag-1*, a gene required for *lin-12* and *glp-1* signaling in *Caenorhabditis elegans*, is homologous to human CBF1 and *Drosophila* Su(H). Development *122*, 1373-1383. Abstract

Conradt, B., and Horvitz, H.R. (1999). The TRA-1A sex determination protein of *C. elegans* regulates sexually dimorphic cell deaths by repressing the *egl-1* cell death activator gene. Cell *98*, 317-327. Article

Conradt, B. and Xue D. Programmed cell death (October 06, 2005), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.32.1, http://www.wormbook.org. Abstract Article

Copley, R.R. (2005). The EH1 motif in metazoan transcription factors. BMC Genomics *6*, 169. Abstract Article

Corsi, A.K., Kostas, S.A., Fire, A., and Krause, M. (2000). *Caenorhabditis elegans* twist plays an essential role in non-striated muscle development. Development *127*, 2041-2051. Abstract

Cui, M., and Han, M. (2003). Cis regulatory requirements for vulval cell-specific expression of the *Caenorhabditis elegans* fibroblast growth factor gene *egl-17*. Dev. Biol. *257*, 104-116. Article

Dantonel, J.C., Quintin, S., Lakatos, L., Labouesse, M., and Tora, L. (2000). TBP-like factor is required for embryonic RNA polymerase II transcription in *C. elegans*. Mol. Cell *6*, 715-722. Abstract Article

Davison, E.M., Saffer, A.M., Huang, L.S., DeModena, J., Sternberg, P.W., and Horvitz, H.R. (2011). The LIN-15A and LIN-56 transcriptional regulators interact to negatively regulate EGF/Ras signaling in *Caenorhabditis elegans* vulval cell-fate determination. Genetics *187*, 803-815. Abstract Article

De Masi, F., Grove, C.A., Vedenko, A., Alibes, A., Gisselbrecht, S.S., Serrano, L., Bulyk, M.L., and Walhout, A.J. (2011). Using a structural and logics systems approach to infer bHLH-DNA binding specificity determinants. Nucleic Acids Res. *39*, 4553-4563. Article

del Castillo-Olivares, A., Kulkarni, M., and Smith, H.E. (2009). Regulation of sperm gene expression by the GATA factor ELT-1. Dev. Biol. *333*, 397-408. Article

Deal, R.B., and Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. Dev. Cell *18*, 1030-1040. Abstract Article

DeMeo, S.D., Lombel, R.M., Cronin, M., Smith, E.L., Snowflack, D.R., Reinert, K., Clever, S., and Wightman, B. (2008). Specificity of DNA-binding by the FAX-1 and NHR-67 nuclear receptors of *Caenorhabditis elegans* is partially mediated via a subclass-specific P-box residue. BMC Mol. Biol. *9*, 2. Article

Deng, H., Sun, Y., Zhang, Y., Luo, X., Hou, W., Yan, L., Chen, Y., Tian, E., Han, J., and Zhang, H. (2007). Transcription factor NFY globally represses the expression of the *C. elegans* Hox gene *Abdominal-B* homolog *egl-5*. Dev. Biol. *308*, 583-592. Abstract Article

Deplancke, B., Mukhopadhyay, A., Ao, W., Elewa, A.M., Grove, C.A., Martinez, N.J., Sequerra, R., Doucette-Stamm, L., Reece-Hoyes, J.S., Hope, I.A.*, et al.* (2006). A gene-centered *C. elegans* protein-DNA interaction network. Cell *125*, 1193-1205. Abstract Article

Dichoso, D., Brodigan, T., Chwoe, K.Y., Lee, J.S., Llacer, R., Park, M., Corsi, A.K., Kostas, S.A., Fire, A., Ahnn, J., and Krause, M. (2000). The MADS-Box factor CeMEF2 is not essential for *Caenorhabditis elegans* myogenesis and development. Dev. Biol. *223*, 431-440. Article

Duggan, A., Ma, C., and Chalfie, M. (1998). Regulation of touch receptor differentiation by the *Caenorhabditis elegans mec-3* and *unc-86* genes. Development *125*, 4107-4119. Abstract

Eastman, C., Horvitz, H.R., and Jin, Y. (1999). Coordinated transcriptional regulation of the *unc-25* glutamic acid decarboxylase and the *unc-47* GABA vesicular transporter by the *Caenorhabditis elegans* UNC-30 homeodomain protein. J. Neurosci. *19*, 6225-6234. Abstract

Egan, C.R., Chung, M.A., Allen, F.L., Heschl, M.F., Van Buskirk, C.L., and McGhee, J.D. (1995). A gut-to-pharynx/tail switch in embryonic expression of the *Caenorhabditis elegans ges-1* gene centers on two GATA sequences. Dev. Biol. *170*, 397-419. Article

Enoki, Y., and Sakurai, H. (2011). Diversity in DNA recognition by heat shock transcription factors (HSFs) from model organisms. FEBS Lett. *585*, 1293-1298. Article

Ercan, S., Giresi, P.G., Whittle, C.M., Zhang, X., Green, R.D., and Lieb, J.D. (2007). X chromosome repression by localization of the *C. elegans* dosage compensation machinery to sites of transcription initiation. Nat. Genet. *39*, 403-408. Abstract Article

Ercan, S., Lubling, Y., Segal, E., and Lieb, J.D. (2011). High nucleosome occupancy is encoded at X-linked gene promoters in *C. elegans*. Genome Res. *21*, 237-244. Abstract Article

Etchberger, J.F., Flowers, E.B., Poole, R.J., Bashllari, E., and Hobert, O. (2009). Cis-regulatory mechanisms of left/right asymmetric neuron-subtype specification in *C. elegans*. Development *136*, 147-160. Article

Fay, D.S., and Yochem, J. (2007). The SynMuv genes of *Caenorhabditis elegans* in vulval development and beyond. Dev. Biol. *306*, 1-9. Abstract Article

Finney, M., and Ruvkun, G. (1990). The *unc-86* gene product couples cell lineage and cell identity in *C. elegans*. Cell *63*, 895-905. Abstract Article

Fisher, K., Southall, S.M., Wilson, J.R., and Poulin, G.B. (2010). Methylation and demethylation activities of a *C. elegans* MLL-like complex attenuate RAS signalling. Dev. Biol. *341*, 142-153. Abstract Article

Flames, N., and Hobert, O. (2009). Gene regulatory logic of dopamine neuron differentiation. Nature *458*, 885-889. Article

Fox, R.M., Von Stetina, S.E., Barlow, S.J., Shaffer, C., Olszewski, K.L., Moore, J.H., Dupuy, D., Vidal, M., and Miller, D.M., 3rd (2005). A gene expression fingerprint of *C. elegans* embryonic motor neurons. BMC Genomics *6*, 42. Abstract Article

Fukushige, T., Goszczynski, B., Yan, J., and McGhee, J.D. (2005). Transcriptional control and patterning of the *pho-1* gene, an essential acid phosphatase expressed in the *C. elegans* intestine. Dev. Biol. *279*, 446-461. Article

Furuyama, T., Nakazawa, T., Nakano, I., and Mori, N. (2000). Identification of the differential distribution patterns of mRNAs and consensus binding sequences for mouse DAF-16 homologues. Biochem. J. *349*, 629-634. Article

Gamsjaeger, R., Liew, C.K., Loughlin, F.E., Crossley, M., and Mackay, J.P. (2007). Sticky fingers: zinc-fingers as protein-recognition motifs. Trends Biochem. Sci *32*, 63-70. Abstract Article

Garrido-Lecca, A., and Blumenthal, T. (2010). RNA polymerase II C-terminal domain phosphorylation patterns in *Caenorhabditis elegans* operons, polycistronic gene clusters with only one promoter. Mol. Cell. Biol. *30*, 3887-3893. Abstract Article

Gaudet, J., and Mango, S.E. (2002). Regulation of organogenesis by the *Caenorhabditis elegans* FoxA protein PHA-4. Science *295*, 821-825. Article

Gaudet, J., and McGhee, J.D. (2010). Recent advances in understanding the molecular mechanisms regulating *C. elegans* transcription. Dev. Dyn. *239*, 1388-1404. Abstract

Gerstein, M.B., Lu, Z.J., Van Nostrand, E.L., Cheng, C., Arshinoff, B.I., Liu, T., Yip, K.Y., Robilotto, R., Rechtsteiner, A., Ikegami, K.*, et al.* (2010). Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. Science *330*, 17751787. Abstract Article

Ghai, V., and Gaudet, J. (2008). The CSL transcription factor LAG-1 directly represses *hlh-6* expression in *C. elegans.* Dev. Biol. *322*, 334-344. Article

Good, K., Ciosk, R., Nance, J., Neves, A., Hill, R.J., and Priess, J.R. (2004). The T-box transcription factors TBX-37 and TBX-38 link GLP-1/Notch signaling to mesoderm induction in *C. elegans* embryos. Development *131*, 1967-1978. Abstract Article

Graber, J.H., Salisbury, J., Hutchins, L.N., and Blumenthal, T.(2007). *C. elegans* sequences that control trans-splicing and operon pre-mRNA processing. RNA *13*, 14091426. Abstract Article

Greer, E.L., Maures, T.J., Hauswirth, A.G., Green, E.M., Leeman, D.S., Maro, G.S., Han, S., Banko, M.R., Gozani, O., and Brunet, A. (2010). Members of the H3K4 trimethylation complex regulate lifespan in a germline-dependent manner in *C. elegans.* Nature *466*, 383-387. Abstract Article

Grishkevich, V., Hashimshony, T., and Yanai, I. (2011). Core promoter T-blocks correlate with gene expression levels in *C. elegans*. Genome Res. *21*, 707-717. Abstract Article

Grove, C.A., De Masi, F., Barrasa, M.I., Newburger, D.E., Alkema, M.J., Bulyk, M.L., and Walhout, A.J. (2009). A multiparameter network reveals extensive divergence between *C. elegans* bHLH transcription factors. Cell *138*, 314-327. Article

Haerty, W., Artieri, C., Khezri, N., Singh, R.S., and Gupta, B.P. (2008). Comparative analysis of function and interaction of transcription factors in nematodes: extensive conservation of orthology coupled to rapid sequence evolution. BMC Genomics *9*, 399. Abstract Article

Hahn-Windgassen, A., and Van Gilst, M.R. (2009). The *Caenorhabditis elegans* HNF4α Homolog, NHR-31, mediates excretory tube growth and function through coordinate regulation of the vacuolar ATPase. PLoS Genet. *5*, e1000553. Article

Hallam, S., Singer, E., Waring, D., and Jin, Y. (2000). The *C. elegans* NeuroD homolog *cnd-1* functions in multiple aspects of motor neuron fate specification. Development *127*, 4239-4252. Abstract

Harfe, B.D., and Fire, A. (1998). Muscle and nerve-specific regulation of a novel NK-2 class homeodomain factor in *Caenorhabditis elegans*. Development *125*, 421-429. Abstract

Harfe, B.D., Vaz Gomes, A., Kenyon, C., Liu, J., Krause, M., and Fire, A. (1998). Analysis of a *Caenorhabditis elegans* Twist homolog identifies conserved and divergent aspects of mesodermal patterning. Genes Dev. *12*, 2623-2635. Article

Hargitai, B., Kutnyanszky, V., Blauwkamp, T.A., Stetak, A., Csankovszki, G., Takacs-Vellai, K., and Vellai, T. (2009). *xol-1,* the master sex-switch gene in *C. elegans*, is a transcriptional target of the terminal sex-determining factor TRA-1. Development *136*, 3881-3887. Article

Harrison, M.M., Ceol, C.J., Lu, X., and Horvitz, H.R. (2006). Some *C. elegans* class B synthetic multivulva proteins encode a conserved LIN-35 Rb-containing complex distinct from a NuRD-like complex. Proc. Natl. Acad. Sci. U. S. A. *103*, 16782-16787. Abstract Article

Harrison, M.M., Lu, X., and Horvitz, H.R. (2007). LIN-61, one of two *Caenorhabditis elegans* malignant-brain-tumor-repeat-containing proteins, acts with the DRM and NuRD-like protein complexes in vulval development but not in certain other biological processes. Genetics *176*, 255-271. Abstract Article

Heard, M., Maina, C.V., Morehead, B.E., Hoener, M.C., Nguyen, T.Q., Williams, C.C., Rowan, B.G., and Gissendanner, C.R. (2010). A functional NR4A nuclear receptor DNA-binding domain is required for organ development in *Caenorhabditis elegans*. Genesis *48*, 485-491. Article

Hillier, L.W., Reinke, V., Green, P., Hirst, M., Marra, M.A., and Waterston, R.H. (2009). Massively parallel sequencing of the polyadenylated transcriptome of *C. elegans*. Genome Res. *19*, 657-666. Abstract Article

Hirose, T., Galvin, B.D., and Horvitz, H.R. (2010). Six and Eya promote apoptosis through direct transcriptional activation of the proapoptotic BH3-only gene *egl-1* in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. U. S. A. *107*, 15479-15484. Article

Ho, L., and Crabtree, G.R. (2010). Chromatin remodelling during development. Nature *463*, 474-484. Abstract Article

Hobert O. Neurogenesis in the nematode *Caenorhabditis elegans* (October 4, 2010), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.12.2, http://www.wormbook.org. Article

Horner, M.A., Quintin, S., Domeier, M.E., Kimble, J., Labouesse, M., and Mango, S.E. (1998). *pha-4*, an HNF-3 homolog, specifies pharyngeal organ identity in *Caenorhabditis elegans*. Genes Dev. *12*, 1947-1952. Abstract Article

Hristova, M., Birse, D., Hong, Y., and Ambros, V. (2005). The *Caenorhabditis elegans* heterochronic regulator LIN-14 is a novel transcription factor that controls the developmental timing of transcription from the insulin/insulin-like growth factor gene *ins-33* by direct DNA binding. Mol. Cell Biol. *25*, 11059-11072. Article

Huang, P., Pleasance, E.D., Maydan, J.S., Hunt-Newbury, R., O'Neil, N.J., Mah, A., Baillie, D.L., Marra, M.A., Moerman, D.G., and Jones, S.J. (2007). Identification and analysis of internal promoters in *Caenorhabditis elegans* operons. Genome Res. *17*, 1478-1485. Abstract Article

Hwang, B.J., and Sternberg, P.W. (2004). A cell-specific enhancer that specifies *lin-3* expression in the *C. elegans* anchor cell for vulval development. Development *131*, 143-151. Article

Ikegami, K., Egelhofer, T.A., Strome, S., and Lieb, J.D. (2010). *Caenorhabditis elegans* chromosome arms are anchored to the nuclear membrane via discontinuous association with LEM-2. Genome Biol. *11*, R120. Abstract Article

Jan, C.H., Friedman, R.C., Ruby, J.G., and Bartel, D.P. (2011). Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. Nature *469*, 97-101. Abstract

Jans, J., Gladden, J.M., Ralston, E.J., Pickle, C.S., Michel, A.H., Pferdehirt, R.R., Eisen, M.B., and Meyer, B.J. (2009). A condensin-like dosage compensation complex acts at a distance to control expression throughout the genome. Genes Dev. *23*, 602-618. Abstract Article

Jiang, Y., Shi, H., Amin, N.M., Sultan, I., and Liu, J. (2008). Mesodermal expression of the *C. elegans* HMX homolog *mls-2* requires the PBC homolog CEH-20. Mech. Dev. *125*, 451-461. Article

Jin, Y., Hoskins, R., and Horvitz, H.R. (1994). Control of type-D GABAergic neuron differentiation by *C. elegans* UNC-30 homeodomain protein. Nature *372*, 780-783. Article

Johnson, A.D., Fitzsimmons, D., Hagman, J., and Chamberlin, H.M. (2001). EGL-38 Pax regulates the ovo-related gene *lin-48* during *Caenorhabditis elegans* organ development. Development *128*, 2857-2865. Abstract

Jones, D., Russnak, R.H., Kay, R.J., and Candido, E.P. (1986). Structure, expression, and evolution of a heat shock gene locus in *Caenorhabditis elegans* that is flanked by repetitive elements. J. Biol. Chem *261*, 12006-12015. Abstract

Kalb, J.M., Lau, K.K., Goszczynski, B., Fukushige, T., Moons, D., Okkema, P.G., and McGhee, J.D. (1998). *pha-4* is Ce-*fkh-1*, a fork head/HNF-3α,β,γ homolog that functions in organogenesis of the *C. elegans* pharynx. Development *125*, 2171-2180. Abstract

Kaltenbach, L., Horner, M.A., Rothman, J.H., and Mango, S.E. (2000). The TBP-like factor CeTLF is required to activate RNA polymerase II transcription during *C. elegans* embryogenesis. Mol. Cell *6*, 705-713. Abstract Article

Kaminsky, R., Denison, C., Bening-Abu-Shach, U., Chisholm, A.D., Gygi, S.P., and Broday, L. (2009). SUMO regulates the assembly and function of a cytoplasmic intermediate filament protein in *C. elegans*. Dev. Cell *17*, 724-735. Abstract Article

Karakuzu, O., Wang, D.P., and Cameron, S. (2009). MIG-32 and SPAT-3A are PRC1 homologs that control neuronal migration in *Caenorhabditis elegans*. Development *136*, 943-953. Abstract Article

Karp, X., and Greenwald, I. (2003). Post-transcriptional regulation of the E/Daughterless ortholog HLH-2, negative feedback, and birth order bias during the AC/VU decision in *C. elegans*. Genes Dev. *17*, 3100-3111. Article

Kato, M., de Lencastre, A., Pincus, Z., and Slack, F.J. (2009). Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during *Caenorhabditis elegans* development. Genome Biol. *10*, R54. Abstract Article

Kim, K., Colosimo, M.E., Yeung, H., and Sengupta, P. (2005). The UNC-3 Olf/EBF protein represses alternate neuronal programs to specify chemosensory neuron identity. Dev. Biol. *286*, 136-148. Article

Kirienko, N.V., and Fay, D.S. (2007). Transcriptome profiling of the *C. elegans* Rb ortholog reveals diverse developmental roles. Dev. Biol. *305*, 674-684. Abstract Article

Kirienko, N.V., McEnerney, J.D., and Fay, D.S. (2008). Coordinated regulation of intestinal functions in *C. elegans* by LIN-35/Rb and SLR-2. PLoS Genet. *4*, e1000059. Abstract Article

Knight, A.S., Notaridou, M., and Watson, R.J. (2009). A Lin-9 complex is recruited by B-Myb to activate transcription of G2/M genes in undifferentiated embryonal carcinoma cells. Oncogene *28*, 1737-1747. Abstract Article

Koester-Eiserfunke, N., and Fischle, W. (2011). H3K9me2/3 binding of the MBT domain protein LIN-61 is essential for *Caenorhabditis elegans* vulva development. PLoS Genet. *7*, e1002017. Abstract Article

Koh, K., Peyrot, S.M., Wood, C.G., Wagmaister, J.A., Maduro, M.F., Eisenmann, D.M., and Rothman, J.H. (2002). Cell fates and fusion in the *C. elegans* vulval primordium are regulated by the EGL-18 and ELT-6 GATA factors – apparent direct targets of the LIN-39 Hox protein. Development *129*, 5171-5180. Abstract

Korenjak, M., Taylor-Harding, B., Binne, U.K., Satterlee, J.S., Stevaux, O., Aasland, R., White-Cooper, H., Dyson, N., and Brehm, A. (2004). Native E2F/RBF complexes contain Myb-interacting proteins and repress transcription of developmentally controlled E2F target genes. Cell *119*, 181-193. Abstract Article

Kostas, S.A., and Fire, A. (2002). The T-box factor MLS-1 acts as a molecular switch during specification of nonstriated muscle in *C. elegans*. Genes Dev. *16*, 257-269. Article

Kostrouchova, M., Krause, M., Kostrouch, Z., and Rall, J.E. (1998). CHR3: a *Caenorhabditis elegans* orphan nuclear hormone receptor required for proper epidermal development and molting. Development *125*, 1617-1626. Abstract

Kozak, M. (1981). Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. *9*, 5233-5252. Abstract Article

Krause, M. (1995). Transcription and translation. Methods Cell Biol. *48*, 483-512. Abstract Article

Krause, M., Fire, A., Harrison, S.W., Priess, J., and Weintraub, H. (1990). CeMyoD accumulation defines the body wall muscle cell fate during *C. elegans* embryogenesis. Cell *63*, 907-919. Abstract Article

Krause, M., Fire, A., White-Harrison, S., Weintraub, H., and Tapscott, S. (1992). Functional conservation of nematode and vertebrate myogenic regulatory factors. J. Cell Sci. Suppl. *16*, 111-115. Abstract

Krause, M., and Hirsh, D. (1987). A trans-spliced leader sequence on actin mRNA in *C. elegans*. Cell *49*, 753-761. Abstract Article

Krause, M., Park, M., Zhang, J.M., Yuan, J., Harfe, B., Xu, S.Q., Greenwald, I., Cole, M., Paterson, B., and Fire, A. (1997). A *C. elegans* E/Daughterless bHLH protein marks neuronal but not striated muscle development. Development *124*, 2179-2189. Abstract

Kuchenthal, C.A., Chen, W., and Okkema, P.G. (2001). Multiple enhancers contribute to expression of the NK-2 homeobox gene *ceh-22* in *C. elegans* pharyngeal muscle. Genesis *31*, 156-166. Article

Lam, N., Chesney, M.A., and Kimble, J. (2006). Wnt signaling and CEH22/tinman/Nkx2.5 specify a stem cell niche in *C. elegans*. Curr. Biol. *16*, 287-295. Article

Lazakovitch, E., Kalb, J.M., Matsumoto, R., Hirono, K., Kohara, Y., and Gronostajski, R.M. (2005). *nfi-I* affects behavior and life-span in *C. elegans* but is not essential for DNA replication or survival. BMC Dev. Biol. *5*, 24. Article

Lee, S.S., Kennedy, S., Tolonen, A.C., and Ruvkun, G. (2003). DAF-16 target genes that control *C. elegans* life-span and metabolism. Science *300*, 644-647. Article

Lei, H., Fukushige, T., Niu, W., Sarov, M., Reinke, V., and Krause, M. (2010). A widespread distribution of genomic CeMyoD binding sites revealed and cross validated by ChIP-Chip and ChIP-Seq techniques. PLoS One *5*, e15898. Abstract Article

Leight, E.R., Glossip, D., and Kornfeld, K. (2005). Sumoylation of LIN-1 promotes transcriptional repression and inhibition of vulval cell fates. Development *132*, 10471056. Abstract

Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. Nature *424*, 147-151. Abstract Article

Li, T., and Kelly, W.G. (2011). A role for Set1/MLL-related components in epigenetic regulation of the *Caenorhabditis elegans* germ line. PLoS Genet. *7*, e1001349. Abstract Article

Li, W., Gao, B., Lee, S.M., Bennett, K., and Fang, D. (2007). RLE-1, an E3 ubiquitin ligase, regulates *C. elegans* aging by catalyzing DAF-16 polyubiquitination. Dev. Cell *12*, 235-246. Abstract Article

Lichtsteiner, S., and Tjian, R. (1993). Cloning and properties of the *Caenorhabditis elegans* TATA-box-binding protein. Proc. Natl. Acad. Sci. U. S. A. *90*, 9673-9677. Abstract Article

Lin, K., Hsin, H., Libina, N., and Kenyon, C. (2001). Regulation of the *Caenorhabditis elegans* longevity protein DAF-16 by insulin/IGF-1 and germline signaling. Nat. Genet. *28*, 139-145. Abstract Article

Liu, J., and Fire, A. (2000). Overlapping roles of two Hox genes and the *exd* ortholog *ceh-20* in diversification of the *C. elegans* postembryonic mesoderm. Development *127*, 5179-5190. Abstract

Liu, T., Rechtsteiner, A., Egelhofer, T.A., Vielle, A., Latorre, I., Cheung, M.S., Ercan, S., Ikegami, K., Jensen, M., Kolasinska-Zwierz, P.*, et al.* (2011). Broad chromosomal domains of histone modification patterns in *C. elegans*. Genome Res. *21*, 227-236. Abstract Article

Liu, X., Long, F., Peng, H., Aerni, S.J., Jiang, M., Sanchez-Blanco, A., Murray, J.I., Preston, E., Mericle, B., Batzoglou, S.*, et al.* (2009). Analysis of cell fate from single-cell gene expression profiles in *C. elegans*. Cell *139*, 623-633. Abstract Article

MacMorris, M., Broverman, S., Greenspoon, S., Lea, K., Madej, C., Blumenthal, T., and Spieth, J. (1992). Regulation of vitellogenin gene expression in transgenic *Caenorhabditis elegans*: short sequences required for activation of the *vit-2* promoter. Mol. Cell. Biol. *12*, 1652-1662. Abstract

MacMorris, M., Spieth, J., Madej, C., Lea, K., and Blumenthal, T. (1994). Analysis of the VPE sequences in the *Caenorhabditis elegans vit-2* promoter with extrachromosomal tandem array-containing transgenic strains. Mol. Cell. Biol. *14*, 484-491. Abstract

Maduro, M.F., Broitman-Maduro, G., Mengarelli, I., and Rothman, J.H. (2007). Maternal deployment of the embryonic SKN-1–>MED-1,2 cell specification pathway in *C. elegans*. Dev. Biol. *301*, 590-601. Abstract Article

Maduro, M.F., Meneghini, M.D., Bowerman, B., Broitman-Maduro, G., and Rothman, J.H. (2001). Restriction of mesendoderm to a single blastomere by the combined action of SKN-1 and a GSK-3β homolog is mediated by MED-1 and -2 in *C. elegans*. Mol. Cell *7*, 475-485. Article

Mah, A.K., Armstrong, K.R., Chew, D.S., Chu, J.S., Tu, D.K., Johnsen, R.C., Chen, N., Chamberlin, H.M., and Baillie, D.L. (2007). Transcriptional regulation of AQP-8, a *Caenorhabditis elegans* aquaporin exclusively expressed in the excretory system, by the POU homeobox transcription factor CEH-6. J. Biol. Chem *282*, 28074-28086. Article

Mango, S.E. The *C. elegans* pharynx: a model for organogenesis (January 22, 2007), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.129.1, http://www.wormbook.org. Article

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

Matthews, J.M., and Sunde, M. (2002). Zinc fingers–folds for many occasions. IUBMB Life *54*, 351-355. Abstract Article

McElwee, J., Bubb, K., and Thomas, J.H. (2003). Transcriptional outputs of the *Caenorhabditis elegans* forkhead protein DAF-16. Aging Cell *2*, 111-121. Abstract Article

McGhee, J.D. 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. Article

McGhee, J.D., and Krause, M.W. (1997). Transcription Factors and Transcriptional Regulation. In *C. elegans* II, D.L. Riddle, T. Blumenthal, B.J. Meyer, and J.R. Priess, eds. (Cold Spring Harbor, NY, CSHL Press), pp. 147-184.

McGhee, J.D., Sleumer, M.C., Bilenky, M., Wong, K., McKay, S.J., Goszczynski, B., Tian, H., Krich, N.D., Khattra, J., Holt, R.A.*, et al.* (2007). The ELT-2 GATA-factor and the global regulation of transcription in the *C. elegans* intestine. Dev. Biol. *302*, 627-645. Article

McKay, S.J., Johnsen, R., Khattra, J., Asano, J., Baillie, D.L., Chan, S., Dube, N., Fang, L., Goszczynski, B., Ha, E.*, et al.* (2003). Gene expression profiling of cells, tissues, and developmental stages of the nematode *C. elegans*. Cold Spring Harb. Symp. Quant. Biol. *68*, 159-169. Abstract Article

Meissner, B., Warner, A., Wong, K., Dube, N., Lorch, A., McKay, S.J., Khattra, J., Rogalski, T., Somasiri, A., Chaudhry, I.*, et al.* (2009). An integrated strategy to study muscle development and myofilament structure in *Caenorhabditis elegans*. PLoS Genet. *5*, e1000537. Abstract Article

Meister, P., Towbin, B.D., Pike, B.L., Ponti, A., and Gasser, S.M. (2010). The spatial dynamics of tissue-specific promoters during *C. elegans* development. Genes Dev. *24*, 766-782. Abstract Article

Metzstein, M.M., and Horvitz, H.R. (1999). The *C. elegans* cell death specification gene *ces-1* encodes a snail family zinc finger protein. Mol. Cell *4*, 309-319. Article

Meyer, B.J. (2010). Targeting X chromosomes for repression. Curr. Opin. Genet. Dev. *20*,179-189. Abstract Article

Miley, G.R., Fantz, D., Glossip, D., Lu, X., Saito, R.M., Palmer, R.E., Inoue, T., Van Den Heuvel, S., Sternberg, P.W., and Kornfeld, K. (2004). Identification of residues of the *Caenorhabditis elegans* LIN-1 ETS domain that are necessary for DNA binding and regulation of vulval cell fates. Genetics *167*, 1697-1709. Article

Miller, R.R., and Okkema, P.G. (2011). The *Caenorhabditis elegans* T-box factor MLS1 requires Groucho co-repressor interaction for uterine muscle specification. PLoS Genet. *7*, e1002210. Abstract Article

WormBook.org

Moilanen, L.H., Fukushige, T., and Freedman, J.H. (1999). Regulation of metallothionein gene transcription. Identification of upstream regulatory elements and transcription factors responsible for cell-specific expression of the metallothionein genes from *Caenorhabditis elegans*. J. Biol. Chem *274*, 29655-29665. Article

Morton, J.J., and Blumenthal, T. (2011). Identification of transcription start sites of trans-spliced genes: uncovering unusual operon arrangements. RNA *17*, 327-337. Abstract Article

Muller, J., and Kassis, J.A. (2006). Polycomb response elements and targeting of Polycomb group proteins in *Drosophila*. Curr. Opin. Genet. Dev. *16*, 476-484. Abstract Article

Murray, J.I., Bao, Z., Boyle, T.J., Boeck, M.E., Mericle, B.L., Nicholas, T.J., Zhao, Z., Sandel, M.J., and Waterston, R.H. (2008). Automated analysis of embryonic gene expression with cellular resolution in *C. elegans*. Nat. Methods *5*, 703-709. Abstract Article

Nam, S., Jin, Y.H., Li, Q.L., Lee, K.Y., Jeong, G.B., Ito, Y., Lee, J., and Bae, S.C. (2002). Expression pattern, regulation, and biological role of runt domain transcription factor, run, in *Caenorhabditis elegans*. Mol. Cell. Biol. *22*, 547-554. Abstract Article

Nash, B., Colavita, A., Zheng, H., Roy, P.J., and Culotti, J.G. (2000). The forkhead transcription factor UNC-130 is required for the graded spatial expression of the UNC129 TGF-β guidance factor in *C. elegans*. Genes Dev. *14*, 2486-2500. Article

Neves, A., English, K., and Priess, J.R. (2007). Notch-GATA synergy promotes endoderm-specific expression of *ref-1* in *C. elegans*. Development *134*, 4459-4468. Article

Neves, A., and Priess, J.R. (2005). The REF-1 family of bHLH transcription factors pattern *C. elegans* embryos through Notch-dependent and Notch-independent pathways. Dev. Cell *8*, 867-879. Abstract Article

Niu, W., Lu, Z.J., Zhong, M., Sarov, M., Murray, J.I., Brdlik, C.M., Janette, J., Chen, C., Alves, P., Preston, E.*, et al.* (2011). Diverse transcription factor binding features revealed by genome-wide ChIP-seq in *C. elegans*. Genome Res. *21*, 245-254. Abstract Article

O'Meara, M.M., Bigelow, H., Flibotte, S., Etchberger, J.F., Moerman, D.G., and Hobert, O. (2009). Cis-regulatory mutations in the *Caenorhabditis elegans* homeobox gene locus *cog-1* affect neuronal development. Genetics *181*, 1679-1686. Article

Oh, S.W., Mukhopadhyay, A., Dixit, B.L., Raha, T., Green, M.R., and Tissenbaum, H.A. (2006). Identification of direct DAF-16 targets controlling longevity, metabolism and diapause by chromatin immunoprecipitation. Nat. Genet. *38*, 251-257. Abstract

Okkema, P.G., and Fire, A. (1994). The *Caenorhabditis elegans* NK-2 class homeoprotein CEH-22 is involved in combinatorial activation of gene expression in pharyngeal muscle. Development *120*, 2175-2186. Abstract

Okkema, P.G., Harrison, S.W., Plunger, V., Aryana, A., and Fire, A. (1993). Sequence requirements for myosin gene expression and regulation in *Caenorhabditis elegans*. Genetics *135*, 385-404. Abstract

Okkema, P. G. and Krause, M. Transcriptional regulation (December 23, 2005), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.45.1, http://www.wormbook.org. Article

Oliveira, R.P., Porter Abate, J., Dilks, K., Landis, J., Ashraf, J., Murphy, C.T., and Blackwell, T.K. (2009). Condition-adapted stress and longevity gene regulation by *Caenorhabditis elegans* SKN-1/Nrf. Aging Cell *8*, 524-541. Article

Oommen, K.S., and Newman, A.P. (2007). Co-regulation by Notch and Fos is required for cell fate specification of intermediate precursors during *C. elegans* uterine development. Development *134*, 3999-4009. Article

Ow, M.C., Martinez, N.J., Olsen, P.H., Silverman, H.S., Barrasa, M.I., Conradt, B., Walhout, A.J., and Ambros, V. (2008). The FLYWCH transcription factors FLH-1, FLH-2, and FLH-3 repress embryonic expression of microRNA genes in *C. elegans*. Genes Dev. *22*, 2520-2534. Article

Park, M., and Krause, M.W. (1999). Regulation of postembryonic G(1) cell cycle progression in *Caenorhabditis elegans* by a cyclin D/CDK-like complex. Development *126*, 4849-4860. Abstract

Passannante, M., Marti, C.O., Pfefferli, C., Moroni, P.S., Kaeser-Pebernard, S., Puoti, A., Hunziker, P., Wicky, C., and Muller, F. (2010). Different Mi-2 complexes for various developmental functions in *Caenorhabditis elegans*. PLoS One *5*, e13681. Abstract Article

Peden, E., Kimberly, E., Gengyo-Ando, K., Mitani, S., and Xue, D. (2007). Control of sex-specific apoptosis in *C. elegans* by the BarH homeodomain protein CEH-30 and the transcriptional repressor UNC-37/Groucho. Genes Dev. *21*, 3195-3207. Abstract Article

Petrella, L.N., Wang, W., Spike, C.A., Rechtsteiner, A., Reinke, V., and Strome, S. (2011). synMuv B proteins antagonize germline fate in the intestine and ensure *C. elegans* survival. Development *138*, 1069-1079. Abstract Article

Pferdehirt, R.R., Kruesi, W.S., and Meyer, B.J. (2011). An MLL/COMPASS subunit functions in the *C. elegans* dosage compensation complex to target X chromosomes for transcriptional regulation of gene expression. Genes Dev. *25*, 499-515. Abstract Article

Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev. *20*, 2922-2936. Abstract Article

Pocock, R., Ahringer, J., Mitsch, M., Maxwell, S., and Woollard, A. (2004). A regulatory network of T-box genes and the even-skipped homologue *vab-7* controls patterning and morphogenesis in *C. elegans*. Development *131*, 2373-2385. Article

Poole, R.J., Bashllari, E., Cochella, L., Flowers, E.B., and Hobert, O. (2011). A genome-wide RNAi screen for factors involved in neuronal specification in *Caenorhabditis elegans*. PLoS Genet. *7*, e1002109. Abstract Article

Powell-Coffman, J.A., Bradfield, C.A., and Wood, W.B. (1998). *Caenorhabditis elegans* orthologs of the aryl hydrocarbon receptor and its heterodimerization partner the aryl hydrocarbon receptor nuclear translocator. Proc. Natl. Acad. Sci. U. S. A. *95*, 2844-2849. Article

Raharjo, W.H., Logan, B.C., Wen, S., Kalb, J.M., and Gaudet, J. (2010). In vitro and in vivo characterization of *Caenorhabditis elegans* PHA-4/FoxA response elements. Dev. Dyn. *239*, 2219-2232. Article

Reece-Hoyes, J.S., Deplancke, B., Barrasa, M.I., Hatzold, J., Smit, R.B., Arda, H.E., Pope, P.A., Gaudet, J., Conradt, B., and Walhout, A.J. (2009). The *C. elegans* Snail homolog CES-1 can activate gene expression *in vivo* and share targets with bHLH transcription factors. Nucleic Acids Res. *37*, 3689-3698. Article

Reece-Hoyes, J.S., Deplancke, B., Shingles, J., Grove, C.A., Hope, I.A., and Walhout, A.J. (2005). A compendium of *Caenorhabditis elegans* regulatory transcription factors: a resource for mapping transcription regulatory networks. Genome Biol. *6*, R110. Abstract Article

Reece-Hoyes, J.S., Diallo, A., Lajoie, B., Kent, A., Shrestha, S., Kadreppa, S., Pesyna, C., Dekker, J., Myers, C.L., and Walhout, A.J. (2011). Enhanced yeast one-hybrid assays for high-throughput gene-centered regulatory network mapping. Nat. Methods *8*, 1059-1064. Abstract Article

Reinke, V. Germline genomics (January 20, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.74.1, http://www.wormbook.org. Article

Ross, J.M., and Zarkower, D. (2003). Polycomb group regulation of Hox gene expression in *C. elegans*. Dev. Cell *4*, 891-901. Abstract Article

Rougvie, A.E., and Ambros, V. (1995). The heterochronic gene *lin-29* encodes a zinc finger protein that controls a terminal differentiation event in *Caenorhabditis elegans*. Development *121*, 2491-2500. Abstract

Roush, S.F., and Slack, F.J. (2009). Transcription of the *C. elegans let-7* microRNA is temporally regulated by one of its targets, *hbl-1*. Dev. Biol. *334*, 523-534. Article

Roy Chowdhuri, S., Crum, T., Woollard, A., Aslam, S., and Okkema, P.G. (2006). The T-box factor TBX-2 and the SUMO conjugating enzyme UBC-9 are required for ABa-derived pharyngeal muscle in *C. elegans*. Dev. Biol. *295*, 664-677. Abstract Article

Roy, P.J., Stuart, J.M., Lund, J., and Kim, S.K. (2002). Chromosomal clustering of muscle-expressed genes in *Caenorhabditis elegans*. Nature *418*, 975-979. Abstract

Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., Lin, M.F.*, et al.* (2010). Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. Science *330*, 1787-1797. Abstract

Russnak, R.H., and Candido, E.P. (1985). Locus encoding a family of small heat shock genes in *Caenorhabditis elegans*: two genes duplicated to form a 3.8-kilobase inverted repeat. Mol. Cell. Biol. *5*, 1268-1278. Abstract

Schafer, J.C., Haycraft, C.J., Thomas, J.H., Yoder, B.K., and Swoboda, P. (2003). XBX-1 encodes a dynein light intermediate chain required for retrograde intraflagellar transport and cilia assembly in *Caenorhabditis elegans*. Mol. Biol. Cell *14*, 2057-2070. Article

Schaner, C. E. and Kelly, W. G. Germline chromatin (January 24, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.73.1, http://www.wormbook.org. Article

Schvarzstein, M., and Spence, A.M. (2006). The *C. elegans* sex-determining GLI protein TRA-1A is regulated by sex-specific proteolysis. Dev. Cell *11*, 733-740. Abstract Article

Seydoux, G., and Dunn, M.A. (1997). Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of *Caenorhabditis elegans* and *Drosophila melanogaster*. Development *124*, 2191-2201. Abstract

Shaham, S., and Bargmann, C.I. (2002). Control of neuronal subtype identity by the *C. elegans* ARID protein CFI-1. Genes Dev. *16*, 972-983. Article

Shaye, D.D., and Greenwald, I. (2011). OrthoList: a compendium of *C. elegans* genes with human orthologs. PLoS One *6*, e20085. Abstract Article

Shemer, G., and Podbilewicz, B. (2002). LIN-39/Hox triggers cell division and represses EFF-1/fusogen-dependent vulval cell fusion. Genes Dev. *16*, 3136-3141. Article

Shi, Y., and Mello, C. (1998). A CBP/p300 homolog specifies multiple differentiation pathways in *Caenorhabditis elegans*. Genes Dev. *12*, 943-955. Abstract Article

Shilatifard, A. (2008). Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. Curr. Opin. Cell Biol. *20*, 341-348. Abstract

Shim, J., and Lee, J. (2008). Regulation of *rnt-1* expression mediated by the opposing effects of BRO-1 and DBL-1 in the nematode *Caenorhabditis elegans*. Biochem. Biophys. Res. Commun. *367*, 130-136. Article

Shim, Y.H. (1999). *elt-1*, a gene encoding a *Caenorhabditis elegans* GATA transcription factor, is highly expressed in the germ lines with *msp* genes as the potential targets. Mol. Cells *9*, 535-541. Abstract

Shim, Y.H., Bonner, J.J., and Blumenthal, T. (1995). Activity of a *C. elegans* GATA transcription factor, ELT-1, expressed in yeast. J. Mol. Biol. *253*, 665-676. Article

Shoichet, S.A., Malik, T.H., Rothman, J.H., and Shivdasani, R.A. (2000). Action of the *Caenorhabditis elegans* GATA factor END-1 in *Xenopus* suggests that similar mechanisms initiate endoderm development in ecdysozoa and vertebrates. Proc. Natl. Acad. Sci. U. S. A. *97*, 4076-4081. Article

Shostak, Y., Van Gilst, M.R., Antebi, A., and Yamamoto, K.R. (2004). Identification of *C. elegans* DAF-12-binding sites, response elements, and target genes. Genes Dev. *18*, 2529-2544. Article

Simonet, T., Dulermo, R., Schott, S., and Palladino, F. (2007). Antagonistic functions of SET-2/SET1 and HPL/HP1 proteins in *C. elegans* development. Dev. Biol. *312*, 367-383. Abstract Article

Sluder, A.E., Mathews, S.W., Hough, D., Yin, V.P., and Maina, C.V. (1999). The nuclear receptor superfamily has undergone extensive proliferation and diversification in nematodes. Genome Res. *9*, 103-120. Abstract

Spencer, W.C., Zeller, G., Watson, J.D., Henz, S.R., Watkins, K.L., McWhirter, R.D., Petersen, S., Sreedharan, V.T., Widmer, C., Jo, J.*, et al.* (2011). A spatial and temporal map of *C. elegans* gene expression. Genome Res. *21*, 325-341. Abstract Article

Spieth, J., Brooke, G., Kuersten, S., Lea, K., and Blumenthal, T. (1993). Operons in *C. elegans*: polycistronic mRNA precursors are processed by trans-splicing of SL2 to downstream coding regions. Cell *73*, 521-532. Abstract Article

Spieth, J., Denison, K., Kirtland, S., Cane, J., and Blumenthal, T. (1985). The *C. elegans* vitellogenin genes: short sequence repeats in the promoter regions and homology to the vertebrate genes. Nucleic Acids Res. *13*, 5283-5295. Abstract Article

Spieth, J. and Lawson, D. Overview of gene structure (January 18, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.65.1, http://www.wormbook.org. Article

Starostina, N.G., Lim, J.M., Schvarzstein, M., Wells, L., Spence, A.M., and Kipreos, E.T. (2007). A CUL-2 ubiquitin ligase containing three FEM proteins degrades TRA-1 to regulate *C. elegans* sex determination. Dev. Cell *13*, 127-139. Abstract Article

Stielow, B., Sapetschnig, A., Kruger, I., Kunert, N., Brehm, A., Boutros, M., and Suske, G. (2008). Identification of SUMO-dependent chromatin-associated transcriptional repression components by a genome-wide RNAi screen. Mol. Cell *29*, 742-754. Abstract Article

Stimpson, K.M., and Sullivan, B.A. (2010). Epigenomics of centromere assembly and function. Curr. Opin. Cell Biol. *22*, 772-280. Article

Stringham, E.G., Dixon, D.K., Jones, D., and Candido, E.P. (1992). Temporal and spatial expression patterns of the small heat shock (*hsp16*) genes in transgenic *Caenorhabditis elegans*. Mol. Biol. Cell *3*, 221-233. Abstract

Strome, S. Specification of the germ line (July 28, 2005), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.9.1, http://www.wormbook.org. Article

Tabuchi, T.M., Deplancke, B., Osato, N., Zhu, L.J., Barrasa, M.I., Harrison, M.M., Horvitz, H.R., Walhout, A.J., and Hagstrom, K.A. (2011). Chromosome-biased binding and gene regulation by the *Caenorhabditis elegans* DRM complex. PLoS Genet. *7*, e1002074. Abstract Article

Thatcher, J.D., Fernandez, A.P., Beaster-Jones, L., Haun, C., and Okkema, P.G. (2001). The *Caenorhabditis elegans peb-1* gene encodes a novel DNA-binding protein involved in morphogenesis of the pharynx, vulva, and hindgut. Dev. Biol. *229*, 480-493. Article

Thatcher, J.D., Haun, C., and Okkema, P.G. (1999). The DAF-3 Smad binds DNA and represses gene expression in the *Caenorhabditis elegans* pharynx. Development *126*, 97-107. Abstract

Thellmann, M., Hatzold, J., and Conradt, B. (2003). The Snail-like CES-1 protein of *C. elegans* can block the expression of the BH3-only cell-death activator gene *egl-1* by antagonizing the function of bHLH proteins. Development *130*, 4057-4071. Abstract Article

Tissenbaum, H.A., and Guarente, L. (2001). Increased dosage of a *sir-2* gene extends lifespan in *Caenorhabditis elegans*. Nature *410*, 227-230. Abstract Article

Tootle, T.L., and Rebay, I. (2005). Post-translational modifications influence transcription factor activity: a view from the ETS superfamily. Bioessays *27*, 285-298. Abstract Article

WormBook.org

Unhavaithaya, Y., Shin, T.H., Miliaras, N., Lee, J., Oyama, T., and Mello, C.C. (2002). MEP-1 and a homolog of the NURD complex component Mi-2 act together to maintain germline-soma distinctions in *C. elegans*. Cell *111*, 991-1002. Abstract Article

van der Linden, A.M., Nolan, K.M., and Sengupta, P. (2007). KIN-29 SIK regulates chemoreceptor gene expression via an MEF2 transcription factor and a class II HDAC. EMBO J. *26*, 358-370. Article

Van Nostrand, E.L., and Kim, S.K. (2011). Seeing elegance in gene regulatory networks of the worm. Curr. Opin. Genet. Dev. *21*, 776-786. Abstract Article

Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A., and Luscombe, N.M. (2009). A census of human transcription factors: function, expression and evolution. Nat. Rev. Genet. *10*, 252-263. Abstract

Venkatesan, K., McManus, H.R., Mello, C.C., Smith, T.F., and Hansen, U. (2003). Functional conservation between members of an ancient duplicated transcription factor family, LSF/Grainyhead. Nucleic Acids Res. *31*, 4304-4316. Article

Vilimas, T., Abraham, A., and Okkema, P.G. (2004). An early pharyngeal muscle enhancer from the *Caenorhabditis elegans ceh-22* gene is targeted by the Forkhead factor PHA-4. Dev. Biol. *266*, 388-398. Article

Vogel, C., and Chothia, C. (2006). Protein family expansions and biological complexity. PLoS Comput. Biol. *2*, e48. Abstract

Von Stetina, S.E., Watson, J.D., Fox, R.M., Olszewski, K.L., Spencer, W.C., Roy, P.J., and Miller, D.M., 3rd (2007). Cell-specific microarray profiling experiments reveal a comprehensive picture of gene expression in the *C. elegans* nervous system. Genome Biol. *8*, R135. Abstract Article

von Zelewsky, T., Palladino, F., Brunschwig, K., Tobler, H., Hajnal, A., and Muller, F. (2000). The *C. elegans* Mi-2 chromatin-remodelling proteins function in vulval cell fate determination. Development *127*, 5277-5284. Abstract

Wagmaister, J.A., Miley, G.R., Morris, C.A., Gleason, J.E., Miller, L.M., Kornfeld, K., and Eisenmann, D.M. (2006). Identification of cis-regulatory elements from the *C. elegans* Hox gene *lin-39* required for embryonic expression and for regulation by the transcription factors LIN-1, LIN-31 and LIN-39. Dev. Biol. *297*, 550-565. Abstract Article

Walker, A.K., See, R., Batchelder, C., Kophengnavong, T., Gronniger, J.T., Shi, Y., and Blackwell, T.K. (2000). A conserved transcription motif suggesting functional parallels between *Caenorhabditis elegans* SKN-1 and Cap'n'Collar-related basic leucine zipper proteins. J. Biol. Chem. *275*, 22166-22171. Abstract

Walker, A.K., Shi, Y., and Blackwell, T.K. (2004). An extensive requirement for transcription factor IID-specific TAF-1 in *Caenorhabditis elegans* embryonic transcription. J. Biol. Chem. *279*, 15339-15347. Abstract

Wallace, J.A., and Felsenfeld, G. (2007). We gather together: insulators and genome organization. Curr. Opin. Genet. Dev. *17*, 400-407. Abstract Article

Wallenfang, M.R., and Seydoux, G. (2002). *cdk-7* Is required for mRNA transcription and cell cycle progression in *Caenorhabditis elegans* embryos. Proc. Natl. Acad. Sci. U. S. A. *99*, 5527-5532. Abstract Article

Wang, D., Kennedy, S., Conte, D., Jr., Kim, J.K., Gabel, H.W., Kamath, R.S., Mello, C.C., and Ruvkun, G. (2005). Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants. Nature *436*, 593-597. Abstract Article

Wang, X., Jia, H., and Chamberlin, H.M. (2006). The bZip proteins CES-2 and ATF-2 alter the timing of transcription for a cell-specific target gene in *C. elegans*. Dev. Biol. *289*, 456-465. Article

Wang, Y., and Levy, D.E. (2006). *C. elegans* STAT: evolution of a regulatory switch. FASEB J. *20*, 1641-1652. Article

Way, J.C., and Chalfie, M. (1988). *mec-3*, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. Cell *54*, 5-16. Article

Wenick, A.S., and Hobert, O. (2004). Genomic cis-regulatory architecture and transacting regulators of a single interneuron-specific gene battery in *C. elegans*. Dev. Cell *6*, 757-770. Article

Whittle, C.M., Lazakovitch, E., Gronostajski, R.M., and Lieb, J.D. (2009). DNA-binding specificity and in vivo targets of *Caenorhabditis elegans* nuclear factor I. Proc. Natl. Acad. Sci. U. S. A. *106*, 12049-12054. Abstract Article

Whittle, C.M., McClinic, K.N., Ercan, S., Zhang, X., Green, R.D., Kelly, W.G., and Lieb, J.D. (2008). The genomic distribution and function of histone variant HTZ-1 during *C. elegans* embryogenesis. PLoS Genet *4*, e1000187. Abstract Article

Wilson, D., Charoensawan, V., Kummerfeld, S.K., and Teichmann, S.A. (2008). DBD—taxonomically broad transcription factor predictions: new content and functionality. Nucleic Acids Res. *36*, D88-92. Abstract Article

Winkelbauer, M.E., Schafer, J.C., Haycraft, C.J., Swoboda, P., and Yoder, B.K. (2005). The *C. elegans* homologs of nephrocystin-1 and nephrocystin-4 are cilia transition zone proteins involved in chemosensory perception. J. Cell Sci *118*, 5575-5587. Article

Winn, J., Carter, M., Avery, L., and Cameron, S. (2011). Hox and a newly identified E2F co-repress cell death in *Caenorhabditis elegans*. Genetics *188*, 897-905. Abstract Article

Winnier, A.R., Meir, J.Y., Ross, J.M., Tavernarakis, N., Driscoll, M., Ishihara, T., Katsura, I., and Miller, D.M., 3rd (1999). UNC-4/UNC-37-dependent repression of motor neuron-specific genes controls synaptic choice in *Caenorhabditis elegans*. Genes Dev. *13*, 2774-2786. Abstract

Xia, D., Zhang, Y., Huang, X., Sun, Y., and Zhang, H. (2007). The *C. elegans* CBFβ homolog, BRO-1, regulates the proliferation, differentiation and specification of the stem cell-like seam cell lineages. Dev. Biol. *309*, 259-272. Abstract Article

Yi, W., Ross, J.M., and Zarkower, D. (2000). *mab-3* is a direct *tra-1* target gene regulating diverse aspects of *C. elegans* male sexual development and behavior. Development *127*, 4469-4480. Abstract

Yi, W., and Zarkower, D. (1999). Similarity of DNA binding and transcriptional regulation by *Caenorhabditis elegans* MAB-3 and *Drosophila melanogaster* DSX suggests conservation of sex determining mechanisms. Development *126*, 873-881. Abstract

Yin, J., Yu, L., and Savage-Dunn, C. (2010). Alternative trans-splicing of *Caenorhabditis elegans sma-9/schnurri* generates a short transcript that provides tissue-specific function in BMP signaling. BMC Mol. Biol. *11*, 46. Abstract Article

Yun, M., Wu, J., Workman, J.L., and Li, B. (2011). Readers of histone modifications. Cell Res. *21*, 564-578. Abstract Article

Yuzyuk, T., Fakhouri, T. H., Kiefer, J., and Mango, S. E. (2009). The polycomb complex protein mes-2/E(z) promotes the transition from developmental plasticity to differentiation in *C. elegans* embryos. Dev. Cell *16*, 699-710. Abstract Article

Zarkower, D., and Hodgkin, J. (1993). Zinc fingers in sex determination: only one of the two *C. elegans* Tra-1 proteins binds DNA in vitro. Nucleic Acids Res. *21*, 3691-3698. Article

Zhong, M., Niu, W., Lu, Z.J., Sarov, M., Murray, J.L., Janette, J., Raha, D., Sheaffer, K.L., Lam, H.Y., Preston, E., *et al.* (2010). Genome-wide identification of binding sites defines distinct functions for *Caenorhabditis elegans* PHA-4/FOXA in development and environmental response. PLoS Genet. *6*, e1000848. Abstract Article