Collaborative regulation of development but independent control of metabolism by two epidermis specific transcription factors in C. elegans

Jiaofang Shao  
Hong Kong Baptist University

Kan He  
Hong Kong Baptist University

Hao Wang  
University of Washington

Wing Sze Ho  
Hong Kong Baptist University

Xiaoliang Ren  
Hong Kong Baptist University

See next page for additional authors

This document is the authors' final version of the published article.  
Link to published article: http://dx.doi.org/10.1074/jbc.M113.487975

APA Citation  
Collaborative Regulation of Development but Independent Control of Metabolism by Two Epidermis-specific Transcription Factors in Caenorhabditis elegans

Received for publication, May 23, 2013, and in revised form, October 4, 2013. Published, JBC Papers in Press, October 6, 2013, DOI 10.1074/jbc.M113.487975

Jiaofang Shao,† Kan He,‡, Hao Wang,§, Wing Sze Ho, Xiaoliang Ren,‡, Xiaoming An,‡, Ming Kin Wong,‡, Bin Yan,‡, Dongying Xie,§, John Stamatoyannopoulos,¶, and Zhongying Zhao,‖

From the †Department of Biology, Hong Kong Baptist University, Hong Kong, China and ‡Department of Genome Sciences, University of Washington, Seattle, Washington 98195

Background: NHR-25 and ELT-3 are required for development but not for initial specification of epidermis in C. elegans. Results: Genome-wide in vivo targets of NHR-25 are identified. Conclusion: NHR-25 and ELT-3 collaboratively regulate development but differentially control metabolism of epidermis. Significance: The results provide insight into how tissue-specific transcription factors enforce cell fate specification initiated by its master regulator.

Cell fate specification is typically initiated by a master regulator, which is relayed by tissue-specific regulatory proteins (usually transcription factors) for further enforcement of cell identities, but how the factors are coordinated among each other to "finish up" the specification remains poorly understood. Caenorhabditis elegans epidermis specification is initiated by a master regulator, ELT-1, that activates its targets, NHR-25 and ELT-3, two epidermis-specific transcription factors that are important for development but not for initial specification of epidermis, thus providing a unique paradigm for illustrating how the tissue-specific regulatory proteins work together to enforce cell fate specification. Here we addressed the question through contrasting genome-wide in vivo binding targets between NHR-25 and ELT-3. We demonstrate that the two factors bind discrete but conserved DNA motifs, most of which remain in proximity, suggesting formation of a complex between the two. In agreement with this, gene ontology analysis of putative target genes suggested differential regulation of metabolism but coordinated control of epidermal development between the two factors, which is supported by quantitative analysis of expression of their specific or common targets in the presence or absence of either protein. Functional validation of a subset of the target genes showed both activating and inhibitory roles of NHR-25 and ELT-3 in regulating their targets. We further demonstrated differential control of specification of AB and C lineage-derived epidermis. The results allow us to assemble a comprehensive gene network underlying C. elegans epidermis development that is likely to be widely used across species and provides insights into how tissue-specific transcription factors coordinate with one another to enforce cell fate specification initiated by its master regulator.

Proper development of a metazoan animal demands a cascade of regulatory events typically involving transcription factors to specify various cell fates. Dissecting such regulatory hierarchy is a challenging task in the context of a multicellular animal. As a result, few of such networks have been thoroughly defined even in model organisms (1, 2). Caenorhabditis elegans epidermis, also known as hypodermis, includes 71 major epidermal cells covering most parts of the animal body and 11 minor epidermal cells surrounding the animal head and tail (3) that play essential roles during both early developmental stages and adulthood. Initial specification of epidermis requires ELT-1, a GATA-like transcription factor expressed in the major type of epidermal cells and its ancestors from 28-cell stage (4). A loss of function mutation in elt-1 produced excess neurons and muscle cells at the expense of epidermis, whereas ectopic expression of ELT-1 in early embryonic cells was sufficient to convert the entire embryo into epidermis (5). Transcription of elt-1 is dependent on a caudal-like maternal protein, PAL-1, in C lineage (6, 7), whereas its upstream regulators in AB have not been identified. Functional genomic analysis combined with RNAi identified another two epidermis-specific transcription factors, NHR-25 and ELT-3, as its direct targets (7). However, how these factors are coordinated with one another to "finish up" the specification events initiated by ELT-1 remains unknown. C. elegans epidermal cells are derived from both AB and C lineages during embryogenesis (see Fig. 1). Those from AB lineage frequently share the immediate ancestors with neuronal or pharyngeal cells, whereas those from C lineage share ancestors with body wall muscle cells (3), indicating the complexity of epidermis specification during development, but how a single tissue type like epi-
Epidermis Regulatory Network

dermis with different lineal origins is specified remains poorly understood.

NHR-25 is a member of nuclear hormone receptor (NHR) family that has undergone a surprising expansion in the *C. elegans* genome, which encodes a total of 284 members (8) as opposed to 48 in human and 18 in *Drosophila melanogaster*. As a result, only a few *C. elegans* NRs are conserved across metazoans. NHR-25 is one of the conserved members that is orthologous to FTZ transcription factor 1 (FTZ-F1) in *D. melanogaster* and is essential for embryogenesis and molting (9, 10). Analysis of a temperature-sensitive *nhr-25* allele revealed its function in regulating cell fusion of epidermal cells and in controlling vulval cell differentiation together with Hox gene *lin-39* (11). Genome-wide RNAi screening indicated the role of NHR-25 in regulating male tail morphogenesis (12). The protein was also involved in controlling fat uptake and storage by regulating acyl-CoA synthase-3 activities (13). In addition, NHR-25 cooperates with the Wnt signaling pathway to control fate differentiation of T seam cells and somatic gonad (14). Despite dramatic expansion of the NHR family, genome-wide binding targets have not been determined for any member of the family in *C. elegans*. ChIP-chip has been attempted to determine the target genes for the NHR-25 ortholog FTZ-F1 in *D. melanogaster*, but few targets were identified (15).

ELT-3 is a conserved GATA-type transcription factor expressed in all major types of epidermal cells during embryogenesis and early larval stages (16). Like NHR-25, expression of ELT-3 is dependent on ELT-1 activities (6, 7), and its downstream target genes have recently been identified (17). Gene knock-out coupled with an expression assay indicated that ELT-3 is sufficient but not necessary for epidermal cell specification and functions downstream of ELT-1 in a redundant pathway controlling epidermis development (5). Intriguingly, perturbation of either ELT-3 or NHR-25 did not prevent specification of epidermal cell fate, although ELT-3 was sufficient for epidermis specification in the absence of ELT-1 (5, 11, 16). We hypothesize that the two transcription factors are not responsible for the initial specification of epidermal cell fate but coordinate with each other for fine-tuning and reinforcing of epidermis cell identities that were initially specified by ELT-1. In this study, we tested the hypothesis by identifying the binding targets of NHR-25 using chromatin immunoprecipitation and sequencing (ChIP-seq) and contrasting them with the targets of ELT-3 identified by Niu et al. (17) using a similar technique. The results suggest that the two transcription factors both cooperatively regulate epidermis development and differentially control metabolisms in epidermis. The identified targets of NHR-25 provide a link between initial specification of epidermal cell fate and “finalization” of epidermal identities. Interestingly, we demonstrated that NHR-25 binding leads to both activating and inhibitory effects on its targets by reporter analysis. In addition, we explored the regulatory mechanism of genetically specific differentiation of epidermis by identifying novel genes and validating the roles of those known to function upstream of NHR-25 in both AB and C lineage-derived epidermis. The combined results allowed us to assemble a comprehensive gene regulatory hierarchy in governing epidermis development in *C. elegans*. Given the conservation of both NHR-25 and ELT-3 amino acid sequences, the network is likely to be broadly used during epidermis development across metazoan species.

MATERIALS AND METHODS

**Strain Generation and Maintenance**—Transgenic strains expressing NHR-25::GFP::3XFLAG fusion were made with a GFP-tagged fosmid (18) by biolistic bombardment as described (19). Two independent transgenic lines stably expressing NHR-25::GFP were obtained and named as RW10473 and RW10475. All of the strains were maintained on nematode growth medium plates at room temperature unless indicated otherwise.

**Complementation Test**—A functional test of the GFP-tagged transgenic animals was done by examining whether the NHR-25::GFP fusion was able to rescue the embryonic lethality of *nhr-25* loss-of-function allele ok64552. Specifically, the NHR-25::GFP allele was crossed into the strain VC469 (+/szT1[lon-2(e678)] I; *nhr-25* ok645) and GFP-expressing wild-type animals were picked for selfing. At least 32 GFP-positive F2 wild-type animals were singled on individual plates and genotyped with PCR. Those that showed GFP expression but yielded no wild-type product spanning the last exon and 3′-UTR of *nhr-25* but showed a PCR band with deletion were deemed as successful rescue of *nhr-25* mutation by the transgene.

**Analysis of Embryonic Expression**—To generate an embryonic line expression profile of NHR-25, the NHR-25::GFP transgene was crossed into a lineaging strain, RW10226 (19), to produce a strain that could be lineaged. Automated lineaging and gene expression profiling essentially were performed as described earlier (20, 21). Images of NHR-25::GFP expression in embryo and postembryonic animals were taken with a Leica SP5 confocal microscope using pinhole 1.0 and 15% laser power (488 nm).

**ChIP-seq**—Synchronized L1 transgenic animals expressing NHR-25::GFP or PHA-4::GFP were fed on OP50 *Escherichia coli* on agar plates for 6 h and cross-linked with 2% formaldehyde for 0.5 h as described with modifications (22). The fixed worms were sheared with sonication, and the cell lysate was used for ChIP with anti-GFP antibody (polyclonal goat IgG; a gift from the library of Tony Hyman).

The libraries were constructed following the Illumina standard protocol and sequenced on an Illumina Genome Analyzer IIx by the High-Throughput Genomics Unit (University of Washington) according to a standard protocol. Thirty-six-based reads were pooled and mapped to the *C. elegans* genome (WS232) with up to two mismatches. The purified DNAs from both PHA-4 ChIP and input samples were used as a template in quantitative PCR using a Roche Applied Science LightCycler 480 using the same cycling conditions as described previously (22).

---

The abbreviations used are: NHR, nuclear hormone receptor; GO, gene ontology; ChIP-seq, chromatin immunoprecipitation and sequencing; modENCODE, model organism encyclopedia of DNA elements; TSS, transcriptional start site; HOT, highly occupied target; FTZ-F1, FTZ transcription factor 1; C, Class.
**ChIP-seq Sequence Alignment and Peak Calling**—The *C. elegans* genome sequence and annotation were downloaded from WormBase (WS232), and the *C. elegans* microRNA annotation was retrieved from miRBase release 19 (23). ChIP-seq raw data for PHA-4, ELT-3, and HLH-1 were obtained from the modENCODE project (17). Highly occupied target (HOT) regions used in peak annotation were reported previously (24). Conserved motifs predicted by PhyloNet (25) were downloaded and converted into MEME format.

Raw reads from each sample were aligned against the *C. elegans* genome using bowtie2 (version 2.0.0-beta7) (26) with default parameters. Only uniquely mapped reads (termed as mappable reads) were used for the subsequent peak calling. The correlation coefficient ($r$) between replicates of ChIP enrichments was examined by comparing the whole genome coverage of mappable reads between each pair of replicates. The coverage for each base was computed and normalized using MACS2 by extending reads to a 200-nucleotide length toward the 3' direction. The whole genome was divided into 1-kb windows, and the average coverage for each window was used to calculate the Pearson correlation coefficient with R package.

Read enrichments/peaks were first calculated for each replicate with the uniquely mapped reads by MACS2 (27) with default parameters except bandwidth was set to 200. For each transcription factor, peaks were called again using the uniquely mapped reads pooled from both replicates with the same parameters as above and thus termed as merged peaks. To increase the specificity of the calculated peaks, the merged peaks were first compared against those calculated from each replicate using BEDTools (v2.16.2) (28). Peaks showing >50% overlapping regions between those called with each replicate were retained for further filtering with HOT regions (24). Only those with no overlap with HOT regions were retained for further analysis.

**Visualization**—The “bed” format files produced by MACS2 were converted to bigWig files using the UCSC tool bedGraphToBigWig (29). Peaks and annotation files as well as conservation scores were loaded to the Integrative Genomics Viewer (version 2.1.21) (30) for visualization and peak graph production. To build signal intensity tracks for comparing different samples, the extended coverage was normalized for each sample according to the total number of uniquely mapped reads. The phastCons7way DNA conservation scores for all seven nematode species were downloaded from the UCSC Genome Browser. The average phastCons scores for each peak were calculated using the UCSC tool bigWigAverageOverBed (29).

**Motif Analysis**—The 50-bp sequences flanking each side of the peak summit were extracted for the top 100 HOT-filtered merged peaks of NHR-25 and ELT-3 based on the enrichment $p$ values. The repeat sequences were removed with RepeatMasker, and the remaining sequences were used as an input for MEME-ChIP (31) for motif identification. The motifs found by MEME-ChIP from our peaks were compared with CE_PhyloNet motifs identified based on the conservation among the genomes of *C. elegans*, *Caenorhabditis briggsae*, and *Caenorhabditis remanei* using TOMTOM (32), a built-in component of the MEME suite, with default parameters.

**Target Gene Identification**—To identify candidate target genes possibly regulated by the transcription factors used in the ChIP experiments, 5-kb regions both up- and downstream of a peak summit were first searched for transcriptional start sites (TSS) and transcriptional end sites in the *C. elegans* genome (WS232) as described previously (22). If no TSS was annotated for a gene, the TSSs were arbitrarily set to the positions 150 bases upstream of the translational start sites. The candidate genes are classified into the following categories based on their positions relative to the peak summits. Those with a peak summit within the coding region are defined as Class 0 (C0). Those with a peak summit located 0–2 and 2–5 kb upstream of the TSS are defined as Classes 1 (C1) and 2 (C2), respectively; whereas those with a peak summit located 0–2 and 2–5 kb downstream of the transcriptional end site are defined as Classes 3 (C3) and 4 (C4), respectively. Those with a peak summit located >5 kb from either its TSS or TES but with no other candidate gene positioned in between are defined as Class 5 (C5). If a single gene can be assigned into several classes because of either different distances of the same peak to various alternative transcripts or more than one peak located in its proximity, it was assigned into a class with the shortest distance to TSS. In addition, peaks located upstream of the TSS take precedence over those downstream of the transcriptional end site. For instance, if gene A has two mRNAs named A1 and A2, and a peak is in the first exon of A1 but upstream of A2, gene A will be defined as a C0 target. In another case, when peak A is located 1 kb upstream and peak B is located 4 kb downstream of the same gene, the gene will be assigned as a C1 target.

**Gene Ontology (GO) Analysis of Target Genes**—To examine whether target genes of NHR-25 and ELT-3 are enriched for particular biological processes (BP), functional annotation enrichment analysis was performed using DAVID v6.7 (33, 34) with the Class 0 and 1 target genes as input. A cutoff of significance level of 0.05 was used in the analysis. A side-by-side comparison of the top 10 enriched GO_BP terms between NHR-25 and ELT-3 (ranked by $p$ values) was made by manually removing the redundant categories of GO terms.

**Association of Binding Site Locations with Epidermis-specific Expression**—To examine whether NHR-25 or ELT-3 target genes show enriched expression compared with their non-target genes in epidermal cells, embryonic expression values of all epidermal genes were extracted from a previous study (35). For each gene, the normalized log$_2$ intensities from three replicates were averaged to give rise to its expression values for both NHR-25 and ELT-3 target genes (C0 to C4). To plot epidermal expression of different classes of target genes, those that are not the targets were used as a reference, and the plotting was performed with the “vioplot” package in R. The Mann-Whitney $U$ test was performed to calculate the significance levels ($p$ value) of gene expression difference between the reference and the target genes with a cutoff $p$ value of 10$^{-3}$.

To further examine whether NHR-25 target genes are enriched in epidermis, NHR-25 target genes (C0 to C1) were compared against those known to be expressed in epidermal cells from a reporter assay (36). The statistical significance of the overlap was calculated based on the hypergeometric distribution.
Validation of NHR-25 Downstream Target Genes—For validation with embryonic and postembryonic expression, RNAi against NHR-25 was performed by microinjection. Double-stranded RNA was injected into young adults at a concentration of 100 ng/μl. Only progeny from parents producing >80% embryonic or larval lethality after the RNAi were used for the subsequent analysis. For profiling of embryonic expression, the injected worms were allowed to lay eggs for at least 8 h before retrieving embryos for automated lineaging and profiling of embryonic expression as described (21). Three replicate embryos were profiled for lineal expression.

For validation with postembryonic expression, fusion constructs between the promoter of the target gene and mCherry were generated using fusion PCR. The fusion product was co-injected either with pZZ56 (unc-119(+) or pCes361 (dpy-5(+) into unc-119 or dpy-5 mutants, respectively, to make transgenic animals as described (37). Fluorescence micrographs before and after the RNAi against NHR-25 were taken using a Leica SP5 confocal microscope with a laser power of 8% for GFP (488 nm) and 20% for mCherry (594 nm), respectively. Fluorescence intensities from three independent early larvae were measured using ImageJ. Each animal was measured three times at different expressing parts.

Identification of NHR-25 Upstream Regulatory Genes—Candidate upstream genes were depleted by RNAi with injection in the NHR-25 lineaging strain expressing both NHR-25::GFP and red lineaging markers (19). Automated lineaging and single cell gene expression profiling before and after gene expression were performed essentially as described (21). At least three replicates were produced for calculation of quantitative expression of a single cell over time for both control and the RNAi-treated embryos. Quantitative expression values were used for plotting with the R package.

RESULTS

NHR-25 Is an Epidermis-specific Transcription Factor—To identify the genome-wide in vivo binding sites of NHR-25, we generated chromosomally integrated transgenic strains of C. elegans stably expressing the NHR-25::GFP::3XFLAG protein fusion using a GFP-tagged fosmid genomic construct (18) by biolistic bombardment and obtained two independent lines that showed comparable reporter expression (Fig. 1 and data not shown). We estimated the transgene copy number with the sequencing data of input samples using an approach similar to that described earlier by calculating the coverage ratio of target regions to the flanking 100-kb regions (18) and found that the two strains carry 2.7 and 1.8 copies of the transgenes, respectively. Consistent with previous studies (11, 19), NHR-25 was first observed in major epidermal cells of embryo with ~200 cells and was continuously expressed throughout embryogenesis and larval stages as judged by a combination of cell morphology and position. It was also observed in precursor/blastic cells of neuron and epidermis, which are called Pnp cells, in embryo and in seam cells during early larval stages (Fig. 1 and data not shown). Despite its expression in precursors of neurons and epidermal cells, NHR-25 appears not to be expressed in the neurons of the larval stage. Both transgenic alleles were able to rescue the lethality produced by loss of function in NHR-25, indicating that they are functional fusions. The two strains were used for the subsequent ChIP experiments.

Genome-wide in Vivo Binding Targets of NHR-25 Were Identified by ChIP-seq—To build a framework for an effective ChIP-seq pipeline in C. elegans, we repeated the ChIP-seq experiment for PHA-4 using an anti-GFP antibody but with modifications (see “Materials and Methods”). We were able to achieve similar or slightly higher enrichment compared with that of the previous study (Fig. 2, A and C), demonstrating the robustness of the method in defining in vivo binding targets. We applied the pipeline to map the genome-wide binding sites of NHR-25 with both starved and fed L1 worms. We were unable to obtain significant enrichment for the starved populations. Therefore, only ChIP data derived from fed L1 worms are shown. We did two ChIP replicates using the animals from two independent transgenic strains expressing NHR-25::GFP::3XFLAG and produced 17.1 and 17.9 million raw reads for each ChIP sample and 21.7 and 28.2 million raw reads for each input sample, respectively (Table 1). We were able to uniquely map ~60 and 77% of the raw reads from ChIP and input samples back to the C. elegans genome, respectively (Table 1). Enrichments between the two ChIP replicates were highly reproducible as judged by their correlation coefficient of genome-wide read coverage (r = 0.91) (Fig. 2B). We first called enrichment peaks using the uniquely mappable reads from each replicate separately. Despite high correspondence of overall read coverage between the replicates, the numbers of peaks called for each replicate differed substantially from each other; i.e. 935 peaks were called in one replicate, whereas 1643 peaks were called in the other. Such difference is not uncommon in the ChIP-seq data generated by the modENCODE project (17). To resolve the discrepancy and maximize the specificity for the called peaks, we adopted a strategy similar to that used previously to filter out the potential false positive bindings (17). Specifically, we pooled the uniquely mappable reads from the two replicates and performed peak calling with the reads using the same settings as those used for the individual replicate. As a result, a total of 1944 peaks were called (Table 1). The peaks were then compared against those called from individual replicates. Only those showing >50% overlap in length with peaks called with reads pooled from both replicates were retained as the candidate peaks (hereafter termed merged peaks) for all the downstream analysis. To further increase the specificity of the observed enrichment, we removed the peaks that showed any overlapping base pairs with the HOTs that were bound by many transcription factors (24). This dramatically decreased the number of candidate peaks to 498 (Table 2) but possibly increased the likelihood to identify the true specific binding events. However, some of the weak binding events (peaks) may be missed due to the stringent filtering criteria.

To allow parallel comparison of the binding profiles between NHR-25 and ELT-3, we downloaded raw reads of both ChIP and input samples for ELT-3 (L1 stage) and HLH-1 (embryo stage) from modENCODE and performed read alignment and peak calling using exactly the same method as we used for NHR-25. ELT-3 was selected for comparison because its tissue specificities are highly similar to those of NHR-25 and because of its known function in epidermis development, whereas
HLH-1 was chosen as negative control because of its tissue specificities in muscle that do not overlap with those of NHR-25 (38). We identified 1555 and 532 candidate peaks for the two transcription factors, respectively (Table 2). An example of an NHR-25 peak, which is located precisely within one of the two highly conserved intergenic regions ~5 kb upstream of an epidermis-specific gene, grh-1, is shown in Fig. 2D. The putative target gene encodes a homolog of LSF/GRH transcription factor required for normal cuticle synthesis and embryonic development (36, 39). Functional consequences of the binding remain to be determined.

**NHR-25 and ELT-3 DNA Demonstrate Discrete Binding Sites but Most of Them Remain in Proximity**—To gain a better understanding of the functional relevance of the DNA binding events, we first calculated the distances between the peak center and the TSS of the closest gene for all the peaks and plotted distances against 100-bp binned genomic windows. As shown in Fig. 3, the number of NHR-25 peaks located within 2 kb upstream of the TSS of a gene is significantly higher than that of its peaks located more than 2 kb upstream from the TSS of a gene or within a gene (p < 0.01) (Fig. 3A). However, the number of ELT-3 peaks decreases gradually in relation to their increasing distances from the TSS of a gene on both sides (Fig. 3B). Whether this difference represents any biologically relevant consequences or is due to a higher number of peaks for ELT-3 than that for NHR-25 remains unclear. Second, we examined whether the identified binding sites are more conserved than those of the overall genome. Notably, we observed that ~75% of the binding sites are enriched inside the regions with significantly higher conservation scores than that of the overall genome (0.34) for both NHR-25 and ELT-3. Only ~5% of the peaks were located inside a region with a conservation score comparable with that of the overall genome (average conservation of entire genomes between nematode species).
Intriguingly, roughly 20% of the peaks were located inside a region with a conservation score below that of the overall genome (Fig. 3C), suggesting *C. elegans*-specific bindings. Finally, we found that only about 15% of the binding sites were completely independent, whereas the remaining binding sites were overlapping between NHR-25 and ELT-3, suggesting the substantial coordination of the bindings between the two factors.

**FIGURE 2. ChIP-seq data analysis.** Pairwise comparisons of genome-wide read coverage between two independent ChIP experiments for PHA-4 (A) and NHR-25 (B) are shown. C, an example of PHA-4 peaks called from this study (red) and a previous study (pink) (22). D, an example of NHR-25 peaks called with replicate A or B (red), two inputs (green), or merged between the replicates (pink). HLH-1 merged peaks are shown at the bottom. Note that the conserved genomic region (bottom gray bars) upstream of *grh-1* is occupied by NHR-25 but not by HLH-1 (light blue).
Overview of ChIP-seq reads and peaks

A, replication A; B, replication B; M, merged from replications A and B.

<table>
<thead>
<tr>
<th>No. peaks</th>
<th>ChIP sample</th>
<th>Input sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. raw reads</td>
<td>Uniquely mapped</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHA-4</td>
<td>4,246</td>
<td>25,703,370</td>
</tr>
<tr>
<td>NHR-25A</td>
<td>935</td>
<td>17,073,633</td>
</tr>
<tr>
<td>NHR-25B</td>
<td>1,643</td>
<td>17,988,956</td>
</tr>
<tr>
<td>NHR-25M</td>
<td>1,944</td>
<td>35,062,589</td>
</tr>
<tr>
<td>ELT-3A</td>
<td>6,907</td>
<td>6,981,933</td>
</tr>
<tr>
<td>ELT-3B</td>
<td>2,736</td>
<td>4,377,247</td>
</tr>
<tr>
<td>ELT-3M</td>
<td>7,265</td>
<td>11,359,180</td>
</tr>
<tr>
<td>HLH-1A</td>
<td>794</td>
<td>1,320,973</td>
</tr>
<tr>
<td>HLH-1B</td>
<td>2,929</td>
<td>7,132,310</td>
</tr>
<tr>
<td>HLH-1M</td>
<td>4,190</td>
<td>8,453,283</td>
</tr>
</tbody>
</table>

*Derived from Niu et al. (17).*

To identify putative binding motifs for the two epidermis-specific transcription factors, we first removed the peaks overlapping with those of HLH-1 out of the merged peaks and used the top 100 peaks from those remaining as ranked by p values as an input for motif identification with MEME (40). Surprisingly, the top putative motifs identified for the two factors show few similarities between each other despite their similar tissue specificities (Fig. 3, D and E), suggesting discrete binding of DNA between the two. It is worth noting that the NHR-25 motif shows high similarity to an existing motif named as F38E1.9.3.matrix that was predicted by phylogenetic footprinting among three Caenorhabditis species (**p = 3.66e−07**) (25) (Fig. 3D), indicating that the motif is conserved across nematode species. The putative motif identified for ELT-3, a GATA-type transcription factor, is nearly identical to the predicted C17G1.3b.4.matrix (**p = 4.67e−06**). It is also similar to GATA factor motifs in the TRANSFAC database (Fig. 3E and data not shown). The results suggest coordinated DNA bindings between GATA-type transcription factors but distinct DNA bindings between NHR-25 and ELT-3 regardless of their shared tissue specificities. However, roughly 85% of their peaks overlap with each other, indicating the proximity of their binding sites. We speculate that the two factors may co-regulate at least some of their targets by forming a complex. Consistent with this, the NHR-25 ortholog FTZ-F1-PB in Drosophila was found to form a protein heterodimer (41).

**GO Analysis Suggested Coordinated Regulation of Epidermis Development but Differential Control of Metabolism between NHR-25 and ELT-3.** To understand the biological meaning of the binding events identified for both transcription factors, we attempted to assign candidate target genes likely to be regulated by the two proteins using the merged peaks after removal of those overlapping with HOT regions. We assigned a protein-coding or microRNA gene as a putative target of NHR-25 or ELT-3 and classified them into six groups (C0–C5) as described previously (17) (see “Materials and Methods”). We were able to assign a total of 1375 and 3396 protein-coding genes (loci) and 18 and 29 microRNAs as the candidate target genes of NHR-25 and ELT-3, respectively (see supplemental data). Over 40% of the NHR-25 targets have a peak located either within (9%) or 0–2 kb upstream of their TSSs (35%). 22% of the target genes have a peak located 2–5 kb upstream of their TSSs. A total of approximately 33% of NHR-25 target genes have a peak located 0–5 kb downstream of their transcriptional end sites, and only about 1% of the target genes are located more than 5 kb away from the closest peak (Fig. 4A).

Given the similar tissue specificities in epidermis between NHR-25 and ELT-3, we wondered whether their target genes share biological functions. We were also curious about whether their target genes overlap with those of the transcription factors specific for tissue other than epidermis such as body wall muscle. To address the question, we first counted the number of target genes shared between NHR-25 and ELT-3. We next calculated the number of target genes shared between either NHR-25 or ELT-3 and HLH-1, a muscle-specific transcription factor (42). We found that nearly half of NHR-25 target genes are also the targets of ELT-3, whereas less than 10% of the target genes of either NHR-25 or ELT-3 are also the targets of HLH-1 (Fig. 4B). This is consistent with the shared tissue specificities between NHR-25 and ELT-3 but differential specificities between either NHR-25 or ELT-3 and HLH-1. The common targets between NHR-25 and ELT-3 are more likely to be involved in epidermis-specific functions, whereas those shared between HLH-1 possibly involve genes with generic roles such as housekeeping or energy production that tend to be broadly expressed. In agreement with this, we found that NHR-25 target genes showed significant enrichment in the gene set with epidermal expression (36) when compared with the set consisting of all genes (**p < 0.001**; hypergeometric distribution) (data not shown).

Despite the striking number of shared target genes between NHR-25 and ELT-3, the number of target genes identified for ELT-3 is significantly higher than that of NHR-25, raising the possibility that the two factors may differentially regulate epidermis development and/or physiology. To test the hypothesis, we performed GO analysis for C0 and C1 of the targets genes for both NHR-25 and ELT-3 to infer their significantly associated biological processes. We chose the two classes of targets because the bindings by the two factors seem more likely to be functional than those of the remaining categories as described below (Fig. 5). In contrast to the significant overlap of the target genes between the two, there is only a single biological process that is shared in the top 10 (ranked by **p** values) biological processes between the two transcription factors (Fig. 4C). The top 10 biological processes of NHR-25 suggest that it is mainly involved in the regulation of epidermis development and morphogenesis, which supports its reported functions (10, 11, 43). Intriguingly, the top processes for ELT-3 allude to its major function in control of metabolisms rather than its expected roles in epidermal fate differentiation. For example, its top biological process is involved in oxidation reduction, a common antiaging process (Fig. 4C), which agrees with its reported roles in aging (44). However, this does not rule out the possibility that ELT-3 may still play important roles in epidermis cell fate specification together with NHR-25. To address the issue, we performed GO analysis again using the target genes shared between the two transcription factors. Interestingly, the top biological processes pulled out with the shared targets are mostly involved in embryogenesis, larval development, and growth regulation, reminiscent of the top processes pulled out with NHR-25 targets only (Fig. 4, C and D), suggesting that the
two transcription factors jointly regulate epidermis development at the embryonic and larval stages but may evolve differential roles in controlling metabolisms under normal growth conditions.

**NHR-25 Bindings Lead to Both Activation and Inhibition of Its Target Genes**—To validate whether the NHR-25 bindings are functional in vivo, we first examined whether the candidate targets for both NHR-25 and ELT-3 were more likely to be expressed than non-target control in epidermis by retrieving epidermis-specific expression data for different classes of targets from microarray analysis (35). The data demonstrated that the expression of NHR-25 Class 0 and 1 targets is significantly higher than that of the control, whereas that of Classes 2, 3, and 4 is elevated modestly (Fig. 5A). A similar pattern holds for ELT-3 targets except for its Class 2 targets (Fig. 5B). We next measured the changes in the expression of reporter fusions with promoters derived from a total of 14 NHR-25 candidate target genes before and after depletion of NHR-25 in either the larval stage (Fig. 6, A–Q) or embryonic stage (Fig. 6R). Class 1 and 2 target genes known to be spatially expressed in epidermis were chosen to ensure that at least one NHR-25 peak is located within their promoter sequences (data not shown). Surprisingly, although most of the tested target genes (7 of 14) demonstrated significantly reduced expression, two of them, nekl-3 (Fig. 6, A–D and Q) and idh-1 (Fig. 6Q), showed significantly elevated expression after depletion of NHR-25, indicating both activating and inhibitory roles of NHR-25 in regulating its target activities. We speculate that the dual roles of NHR-25 in gene regulation are achieved through forming a complex with other cofactors and/or selective post-translational modifications. One of the down-regulated target genes, *idh-1*, encodes a predicted isocitrate dehydrogenase, an enzyme that catalyzes the oxidative decarboxylation of isocitrate. Inactivation of its activities extended the life span of *C. elegans* (45) and reduced mitochondrial membrane potential, suggesting that NHR-25 likely plays roles in antiaging by down-regulating the activities of IDH-1. Expression in three of the 14 tested target genes (C28H8.11, rpa-2, and Y54G2A.11) remained relatively constant after perturbation of NHR-25 activities.

To further address whether NHR-25 and ELT-3 coordinate control of their common targets but differentially regulate their own targets, we examined the expression of their specific or common targets before and after depletion of its activities by quantitative PCR (Fig. 6S). Specifically, we chose a total of 12 target genes from the NHR-25- or ELT-3-specific target list or the list of their common targets with four for each category. Each of the four genes was derived from one of the top four GO classes (Fig. 4, C and D) with each from a single class. Expression of all four common targets is significantly changed compared with that of the no RNAi control after depletion of either NHR-25 or ELT-3, indicating coordinated regulation of epidermis development between the two. Intriguingly, a cell cycle-
related gene, cdc-6, is apparently inhibited by both NHR-25 and ELT-3, whereas the remaining three targets are activated by the two genes. Expression of three of the four NHR-25-specific targets is significantly reduced by depletion of NHR-25 as opposed to that of no RNAi control, whereas only one of the four is significantly lowered by depletion of ELT-3, supporting the specific regulation by NHR-25. However, it appears that NHR-25 also contributes to the control of ELT-3-specific targets, although no significant binding was identified in their proximity. Expression of three of the four ELT-3-specific tar-
gets is significantly altered after depletion of ELT-3, and similar changes were observed after depletion of NHR-25 but at a lower significance level (Fig. 6). In addition, expression of the remaining ELT-3 target gene, acox-1, shows little change after depletion of ELT-3 but a significant reduction after depletion of NHR-25, suggesting that NHR-25 may contribute to regulation of many genes other than its direct targets identified here in epidermis by forming a complex with various partners. However, it is likely to control epidermis development by directly regulating its targets.

Epidermal Cells Derived from AB and C Lineages Are under Both Differential and Coordinated Control by Regulators Upstream of NHR-25—Epidermal cells arise from multiple precursors in AB and C lineages of C. elegans embryo (Fig. 1 and data not shown) (3). Substantial knowledge of epidermal cell fate specification is derived from the study of the C lineage only (6, 7); therefore, we wondered whether specification of AB- or C-derived epidermis is under differential control by the regulatory proteins that function upstream of NHR-25. We chose putative regulatory genes that are known or likely to regulate cell fate differentiation either in AB or C lineage during embryogenesis except for SKN-1, which was known to be required for differentiation in the EMS lineage (46). The gene was chosen because of its maternal expression in epidermal precursors of AB lineage, but its roles in differentiation of AB-derived epidermis remain unknown. We evaluated the genetic interactions between these genes and NHR-25 by gene depletion followed by automatic expression profiling of NHR-25 as described (21). We found that in addition to the homeotic transformation of EMS into C-like lineage as described previously (46) depletion of SKN-1 led to conversion of the ABalp daughters from neuronal and pharyngeal fates (3) into epidermis as judged by NHR-25 expression (Fig. 7). The cell fate in remaining AB sublineages and C were not affected, indicating that SKN-1 regulates epidermal specification in AB but not in C lineage. Wnt and Notch signaling pathways are frequently involved in cell fate differentiation during metazoan development (47–51). We found that depletion of either POP-1, a sole member of the T cell factor/lymphoid enhancer factor family of the canonical and non-canonical Wnt signaling pathways, or...
LAG-1, a single copy of CSL (CBF1, Suppressor of Hairless), LAG-1 family member functioning as a downstream effector in the Notch signaling pathway, altered expression of NHR-25 in various ways (Figs. 7 and 8). For example, depletion of POP-1 by RNAi abolished NHR-25 expression in ABp daughters (data not shown), whereas ectopic expression was found in the daughters of ABal and ABarp (Fig. 8, A–D). However, expression of NHR-25 in C lineage seems preserved, although asymmetry in cell division timings was lost. Depletion of LAG-1 activities suppressed NHR-25 expression in both AB and C lineages (Fig. 7, C and D) but produced ectopic expression in EMS, resulting in conversion of MS into C-like and E into epidermis-like fates (Fig. 8, E and F).

NHR-25 was initially identified as an ELT-1 target by microarray assay (6) and was further confirmed by reporter analysis (7). We repeated the reporter assay by depletion of ELT-1 with RNAi followed by detailed profiling of NHR-25 expression with single cell resolution. We did observe decreased (Fig. 7, E–H), but not loss of, NHR-25 expression in both AB and C lineages after RNAi as described earlier (7), although the RNAi produced over 80% penetrance of embryonic lethality. Interestingly, changes in expression levels did not seem to be uniform across the NHR-25-expressing cells but were individualized among the cells (Fig. 7H), demonstrating an asymmetry in alterations between anterior and posterior sibling. The asymmetric patterns of expression changes suggest that there are some other components that regulate NHR-25 expression along with ELT-1.

PAL-1 is a caudal-like homeobox protein responsible for cell fate specification of C blastomere (52). Previous gene depletion combined with a reporter assay demonstrated that nhr-25 transcription in posterior embryo is dependent on PAL-1 (6). We repeated the RNAi against PAL-1 followed by lineal expression profiling in embryo and found that inactivation of PAL-1 activities abolished expression of NHR-25 in Caa daughters (Fig. 8, G and H), supporting that PAL-1 is a C lineage-specific regulator. Consistent with this, expression of NHR-25 is not affected in any AB lineages by inactivation of PAL-1. In addition, we observed ectopic expression of NHR-25 in Capp daughters despite the loss of asymmetry in cell division timings, suggesting that PAL-1 selectively regulates NHR-25 activities in C lineage possibly with some other genes that remain to be identified.

tbx-37 and tbx-38 are two paralogs of T-box-containing transcription factors. The two genes were found to be the targets of the first Notch signaling event during embryogenesis, leading to its suppressed expression in ABp but not in Aba, which is essential for AB descendants to produce pharyngeal tissues. Here we observed significantly reduced expression of NHR-25 in ABa daughters (n = 3, p < 0.01, Student’s t test) (Fig. 8, I and
but not in those of ABp and C (data not shown) after simultaneous depletion of the activities for both genes by RNAi. Remaining NHR-25 expression could be due to incomplete penetrance, which is not uncommon for a double RNAi assay. Alternatively, there are other regulatory proteins acting redundantly with the two T-box transcription factors in controlling NHR-25 activities. *tbx-8* and *tbx-9* are another two paralogs of T-box-containing transcription factors that were shown previously to activate both NHR-25 and ELT-3 (7). The observation is supported by our lineal expression data. In summary, we concluded that AB and C-derived epidermis is under both differential and coordinated control during *C. elegans* embryogenesis.

**DISCUSSION**

The comprehensive regulatory hierarchy underlying cell fate differentiation of metazoan animals is lacking for most tissue types even in model organisms. In this study, we identified genome-wide *in vivo* binding targets of NHR-25, an epidermis-specific transcription factor, and contrasted them with those of another epidermis-specific transcription factor in *C. elegans*. A combination of genomic assay with lineage analysis allowed us to assemble a comprehensive regulatory network underlying epidermis development.

**Primary Specification and Subsequent Reinforcement of Epidermal Cell Fate by Tissue-specific Transcription Factors—** *C. elegans* epidermis specification is initiated by a GATA-type transcription factor, ELT-1, and relayed by a few other epidermis-specific transcription factors, including NHR-25 and ELT-3 (7). First, the relationship among ELT-1, NHR-25, and ELT-3 has been established (Fig. 1) (7). Second, the cellular functions of both the master regulator (ELT-1) and its dependent transcription factors (NHR-25 and ELT-3) were characterized (5). Third, NHR-25 and ELT-3 were not required for initial specification of epidermis but were important for epidermis development. Therefore, it would be important to understand how the epidermis-specific transcription factors coordinate with each other to finish up epidermis specification initiated by
the master regulator ELT-1. A straightforward method is to compare their in vivo binding targets. Our analysis of the DNA binding sequences revealed discrete DNA binding motifs between NHR-25 and ELT-3, but most of them remain in proximity, strongly suggesting cooperative regulation of their targets by possibly forming a complex. In agreement with this, GO analysis of their common target genes suggests collaborative regulation of epidermis development between the two factors (Figs. 4 and 9). Strikingly, the top 10 biological processes of ELT-3 targets appear to be predominantly involved in stress responses and metabolic processes with only one process that is shared between those of NHR-25 targets regardless of their shared tissue specificities. For example, the top biological process of ELT-3 targets consists of 71 genes and appears to mainly regulate detoxification through oxidation-reduction, which supports the previous finding that ELT-3 formed a circuit along with ELT-5 and -6 to guide aging in C. elegans (44), although the claim is controversial (44, 53). We speculate that the roles of ELT-3 in metabolism and aging may be its specific functions because its roles in regulation of growth and epidermis development seem redundant with those of NHR-25 and ELT-1. Shared roles in embryonic and larval development between the two transcription factors might be an evolved strategy for developmental robustness, which helps maximize survival chances in case of disaster such as a loss-of-function mutation in NHR-25 for example. In agreement with this, both coordinated and differential control of cell identities by two muscle-specific transcription factors, HLH-1 and UNC-120, was observed in C. elegans muscle. The two proteins regulate their target genes that are either muscle- or non-muscle-specific (54). Nevertheless, given the fact that targets identified by expression-based analysis cannot distinguish direct or indirect targets and the complications associated with artificial enrichment of “muscle”-like cells, we provide a comparison of the direct binding targets between the two epidermis-specific transcription factors identified under parallel physiological conditions.

In contrast, the major roles of NHR-25 appear to reinforce the epidermal cell fate specified by ELT-1 through maintenance of epidermal integrity, secretion of cuticle, and regulation of housekeeping genes in epidermal cells. Many of the NHR-25 targets in the top biological processes are directly and indirectly involved in epidermal development or broadly involved in embryonic and larval development, which is consistent with its known function (Fig. 4). Notably, molting, one of the reported functions of NHR-25, is absent from the top 10 list but is...
Epidermis Regulatory Network

tent as 26th in the GO list as judged by the significance values. In contrast, roles in cuticle development have not been reported for ELT-3 but were pulled out in our assay. The results raise the possibility that molting defects frequently observed after inactivation of NHR-25 could be due to biased studies influenced by its known expression patterns in epidermis. Our data likely provide an unbiased and systematic view of the roles of the two factors in epidermis development. Taken together, our data provide a functional link between the initial specification and final enforcement of epidermal cell identities in C. elegans. One intriguing observation is that both NHR-25 and ELT-3 play activating roles for most of their targets but are inhibitory for a common target gene, cdc-6, a cell cycle component, as revealed by quantitative PCR analysis (Fig. 6S). This might be not uncommon for transcription factors responsible for downstream specification of a cell type because the final stage of cell fate differentiation likely demands arrest of cell division. Notably, despite the dramatic expansion of (NHRs in Caenorhabditis species, systematic mapping of in vivo binding target genes has not been attempted for any of the NHR-type transcription factors in C. elegans (55). Our genome-wide profiling of NHR-25 binding targets provides a new entry point for study of NHR function and evolution.

Differential Regulation of Epidermal Cell Fate with Different Lineal Origins—C. elegans epidermis arises mainly from a subset of daughter cells of both AB and C lineages, including ABarp, ABpla, ABpra, and C (Fig. 1 and data not shown). Combinatorial origins of epidermis make it difficult to dissect the gene regulatory network in the context of an entire embryo. Therefore, a simplified model focusing on the C lineage was used for deducing the gene regulatory network of epidermis and body wall muscle specification through a functional genomic approach (6, 7). These studies established PAL-1 as an activator upstream of ELT-1 through which it further regulates its downstream targets, NHR-25 and ELT-3 (Fig. 1). However, analysis of the epidermis gene regulatory network based on C lineage only may not reflect the global picture of gene regulation in epidermis if AB-derived epidermis is not taken into account. This is because single cell gene expression profiling in the L1 larval stage of C. elegans suggested that cells with identical fates derived from different cell lineages can be formed by different gene regulatory pathways (56). One open question is whether the specifications of AB- and C-derived epidermal cells are under differential control. The answer to this question requires differential profiling of NHR-25 binding sites in the cell populations derived from AB and C in vivo. To partially address the issue, we evaluated the regulatory roles of proteins that were either known or likely to function upstream of NHR-25 in both AB and C lineages by automated lineage/expression analysis. The results showed differential dependence of NHR-25 activities on its upstream regulators in AB- and C-derived epidermal cells (Figs. 7–9), supporting the previous observation that the same cell type can be produced through different regulatory pathways. Given the fact that the in vivo targets were profiled for only two of the ELT-1 downstream transcription factors, more such targets were found downstream of ELT-1 in the microarray-based analysis. For example, one transcription factor, LIN-26, was also pulled out in parallel with NHR-25 and ELT-3 as an ELT-1 target in the microarray analysis (7). Profiling of in vivo targets for both ELT-1 and LIN-26 in the future would definitely extend the network underlying epidermis development.

It is noteworthy that NHR-25 appears to inhibit ELT-3 activities through its microRNA expression (Fig. 9), whereas ELT-3 activities are not required for NHR-25 expression based on its perturbation followed by automatic profiling of reporter expression (data not shown), although the two factors were reported to mutually inhibit each other based on a microarray expression assay (7). The discrepancies are likely produced by different sensitivities between the two assay methods, incomplete RNAi penetrance, or both. In addition, both NHR-25 and ELT-3 were identified as their own targets in the ChIP-seq assay, suggesting that the two factors self-regulate their activities in vivo. Analysis of both upstream regulators and downstream targets of epidermis-specific transcription factors allowed us to assemble a comprehensive in vivo gene regulatory network controlling epidermis development and physiology in C. elegans (Fig. 9). Both NHR-25 and ELT-3 proteins are conserved among metazoan species. For example, the NHR-25 ortholog in Drosophila species is also essential for embryogenesis and molting (57), and the ELT ortholog in mammals, GATA-3, is required for establishment of the epidermal barrier and survival in the ex utero environment (58). Therefore, we speculate that the gene network focusing on tissue-specific transcription factors NHR-25 and ELT-3 is likely to be widely used across species during epidermis development.

Acknowledgments—We thank Robert H. Waterston for transgenic strains, Mihail Sarov for anti-GFP antibody, and Chung Wai Shing for logistic support. Some strains were provided by the C. elegans Genetic Center. nhr-25 deletion mutant was generated by the International C. elegans Gene Knock-out Consortium.

REFERENCES

8. 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
paralogous microRNA genes of Caenorhabditis elegans direct the cessation of molting via the conserved nuclear hormone receptors NHR-23 and NHR-25. Development 133, 4631–4641


Collaborative Regulation of Development but Independent Control of Metabolism by Two Epidermis-specific Transcription Factors in *Caenorhabditis elegans*
Jiaofang Shao, Kan He, Hao Wang, Wing Sze Ho, Xiaoliang Ren, Xiaomeng An, Ming Kin Wong, Bin Yan, Dongying Xie, John Stamatoyannopoulos and Zhongying Zhao

*J. Biol. Chem.* 2013, 288:33411-33426.
doi: 10.1074/jbc.M113.487975 originally published online October 6, 2013

Access the most updated version of this article at doi: 10.1074/jbc.M113.487975

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2013/10/06/M113.487975.DC1

This article cites 57 references, 29 of which can be accessed free at http://www.jbc.org/content/288/46/33411.full.html#ref-list-1