Transcription Factor Activity Mapping of a Tissue-Specific In Vivo Gene Regulatory Network

Graphical Abstract

Highlights
- In vivo gene regulatory network based on transcription factor activity
- Promoters directly or indirectly regulated by a median of 18 transcription factors
- Network contains cell-autonomous and non-autonomous regulation
- Nested effects modeling reveals information flow between transcription factors

Authors
Lesley T. MacNeil, Carles Pons, H. Efsun Arda, Gabrielle E. Giese, Chad L. Myers, Albertha J.M. Walhout

Correspondence
marian.walhout@umassmed.edu

In Brief
MacNeil et al. use transcription factor activity to derive an in vivo intestinal gene regulatory network and use these interactions to generate a predictive model of information flow in the C. elegans intestine.
Transcription Factor Activity Mapping of a Tissue-Specific In Vivo Gene Regulatory Network

Lesley T. MacNeil, Carles Pons, H. Efsun Arda, Gabrielle E. Giese, Chad L. Myers, and Albertha J.M. Walhout

1Program in Systems Biology
2Program in Molecular Medicine
University of Massachusetts Medical School, Worcester, MA 01605, USA
3Department of Computer Science and Engineering, University of Minnesota-Twin Cities, Minneapolis, MN 55455, USA
4Present address: Department of Biochemistry and Biomedical Sciences, Farncombe Family Digestive Health Research Institute, McMaster University, Hamilton, ON L8S 4L8 Canada
5Present address: Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 93405, USA
*Correspondence: marian.walhout@umassmed.edu
http://dx.doi.org/10.1016/j.cels.2015.08.003

SUMMARY

A wealth of physical interaction data between transcription factors (TFs) and DNA has been generated, but these interactions often do not have apparent regulatory consequences. Therefore, equating physical interaction data with gene regulatory networks (GRNs) is problematic. Here, we comprehensively assay TF activity, rather than binding, to construct a network of gene regulatory interactions in the C. elegans intestine. By manually observing the in vivo tissue-specific knockdown of 921 TFs on a panel of fluorescent transcriptional reporters, we identified a GRN of 411 interactions between 19 promoters and 177 TFs. This GRN shows only a modest overlap with physical interactions, indicating that many regulatory interactions are indirect. We applied nested effects modeling to uncover the information flow between TFs in the intestine that converges on a small set of physical TF-promoter interactions. We found numerous cell non-autonomous regulatory interactions, illustrating tissue-to-tissue communication. Our study illuminates the complexity of gene regulation in the context of a living animal.

INTRODUCTION

The correct expression of genes in space and time plays a critical role not only during development but also in maintaining homeostasis and adapting to changing environmental or dietary conditions. The regulation of gene expression is accomplished, at a first level, by transcription factors (TFs) that physically bind regulatory sequences upstream within the promoter of their target genes or at distal sites in enhancers. It is widely believed that the regulation of most metazoan genes occurs through the combined action of multiple TFs that work coordinately in complex gene regulatory networks (GRNs) to deliver proper gene expression programs.

The last decade has seen the generation of genome-scale data of physical interactions between TFs and regulatory DNA elements in a number of model organisms as well as in human cells. For instance, chromatin immunoprecipitation (ChIP) has been used extensively to obtain genome-wide interaction data for individual TFs in human cultured cells, yeast, flies, and worms (Araya et al., 2014; Bernstein et al., 2012; Harbison et al., 2004; Sandmann et al., 2006; Tabuchi et al., 2011; Whittle et al., 2009). Other methods that have been used for the large-scale mapping of physical TF-DNA interactions include high-throughput yeast one-hybrid (Y1H) assays that identify the repertoire of TFs that can interact with individual regulatory elements of interest in a single experiment (Deplancke et al., 2006; Fuxman Bass et al., 2015; Gaudinier et al., 2011; Hens et al., 2011; Reece-Hoyes et al., 2013). Several studies suggest that many physical TF-DNA interactions lack an apparent regulatory consequence on gene expression (Hu et al., 2007; Jakobsen et al., 2007; Kemmeren et al., 2014; Sandmann et al., 2006; Vokes et al., 2008). Therefore, equating TF binding with gene regulation, as has been done in numerous studies, is questionable, and it is critical that changes in gene expression resulting from loss of TF activity need to be identified and integrated with physical interaction networks (Walhout, 2011).

Large-scale efforts to generate TF activity-driven regulatory networks in complex organisms lag behind the identification of physical TF-DNA interactions. Available regulatory interaction studies in yeast (Kemmeren et al., 2014) and in multicellular organisms such as the nematode C. elegans (Gracida and Eckmann, 2013; Kouns et al., 2011; Murphy et al., 2003; Quach et al., 2013) are TF-centered because they focus on gene expression changes elicited by the perturbation of individual TFs. However, because a major goal in systems biology is to understand the regulation of each gene in a genome of interest, it is remarkable that large-scale gene-centered studies that determine the repertoire of regulatory interactions for individual genes are not yet available.

We reasoned that using promoter reporters for multiple co-expressed genes together with TF RNAi in a single tissue in living animals may provide a first step toward the delineation of a GRN that is based on TF activity rather than binding and that...
is, therefore, truly regulatory in nature. We selected C. elegans as a model system because it is a multicellular organism amenable to large-scale RNAi screening for phenotypic characterizations (Kamath et al., 2003) and for detection of tissue-specific changes in reporter gene expression in living animals (Watson et al., 2013). We focused on the C. elegans intestine, a highly dynamic organ that serves not only to digest food but also to fine-tune metabolic processes and respond to pathogenic insults. In addition, the intestine provides a technical advantage because it is robustly sensitive to RNAi and easy to see by light microscopy, enabling visual screening.

Using 19 intestinal gene promoters, we delineate an in vivo TF activity-based GRN comprising 411 interactions involving 177 TFs, indicating pervasive transcriptional regulation in the C. elegans intestine (Figures 1A and 1B). Interactions discovered by RNAi screening can be direct or indirect. We identify interactions that are likely to be direct by comparing the regulatory interactions detected by RNAi with physical interaction data for the same promoters. In agreement with prior studies in other organisms (Hu et al., 2007; Jakobsen et al., 2007; Kemmeren et al., 2014; Sandmann et al., 2006; Vokes et al., 2008), we found limited overlap between regulatory and physical TF interactions for the same promoters, suggesting that many regulatory interactions are indirect. Cellular proteins often function in the context of complex signaling pathways, and perturbation of genes that function within a pathway will result in similar regulatory effects on downstream genes. Indirect regulatory effects may therefore arise when TFs that regulate the expression of genes within a signaling pathway are perturbed. We applied nested effects modeling to the GRN to generate a hierarchical TF model that illuminates (indirect) information flow. We validated and refined this model by identifying cell-autonomous and non-autonomous effects as well as several feedforward loops (FFLs). Together, our findings indicate that regulatory networks may be wired via a complex hierarchy involving TFs and other types of regulators to maintain intestinal homeostasis.

RESULTS AND DISCUSSION

A Tissue-Specific, Gene-Centered GRN Derived from Living Animals

We first generated a near-complete RNAi library of 921 C. elegans TFs that contains 891 of the 934 predicted TFs (95%) (Reece-Hoyes et al., 2005) as well as 30 unconventional DNA binding proteins (uDBPs) (Table S1). As a starting point for in vivo GRN mapping, we selected 19 transgenic C. elegans strains in which GFP is expressed in the intestine. These strains were selected because they express sufficiently high levels of GFP to enable visual screening by light microscopy. Overall, the set of transgenes represents different aspects of intestinal biology, including responses to diet, oxidative stress, and the unfolded protein response and include promoters of both protein-coding and microRNA genes. Although all strains express GFP (or wCherry) in the gut, most also exhibit GFP expression in other tissues (Figure S1).

We decided to perform a visual qualitative screen in which we assessed changes in fluorescent protein levels upon TF perturbation at the L4/young adult stage of development. We performed a qualitative screen because we aimed to delineate interactions in a single tissue in living animals. There are no methods available to quantify GFP levels in only one tissue when the fluorescent protein is also expressed in other tissues. Because the reporters used express GFP in multiple tissues, GFP quantification in whole animals would be misleading. For instance, if GFP goes down 50% in the gut but is also expressed in another tissue where it does not change, then whole-animal quantification would be meaningless. Finally, because RNAi may not result in complete knockdown, it would be challenging to interpret quantitative data.

To ensure high-quality GRN data, we performed the screens blindly, three times independently. Interactions detected at least twice were retested on larger plates with a greater number of animals. To ensure a low rate of false positives, only interactions that were confirmed during retesting were kept in the final dataset (Figure 1B).

We combined all regulatory interactions into a first-level, TF activity-driven GRN comprising 411 interactions between 19 gene promoters and 177 TFs (Figure 2A; Table S2). Most
(87%) interactions are activating (GFP down after RNAi), suggesting that transcriptional activation is more prominent in the C. elegans intestine than repression or that activation may be more readily identified by our visual screening method. The number of regulatory interactions observed per promoter ranges from 10–55, and the median is 18 TFs. Therefore, all promoters receive regulatory input through the activity of multiple TFs, either directly or indirectly.

Figure 2. A TF Activity-Based C. elegans GRN
(A) GRN depicting regulatory interactions identified by TF RNAi. Triangles indicate the 19 target promoters. Circles indicate the TFs regulating these promoters in the intestine. k indicates the out-degree or the number of promoters regulated by each TF in the intestine. Blue edges indicate activating interactions (intestinal GFP down upon TF RNAi), and orange edges indicate repressive interactions (intestinal GFP up upon TF RNAi). TFs regulating a single target gene (bottom) are not labeled.

(B) Validation of regulatory interactions. Changes in GFP expression in Pacdh-1::GFP and Pacdh-2::GFP animals following knockdown of a family of closely related NHRs was validated by crossing nhr-68, nhr-101, or nhr-114 deletion mutations into Pacdh-1::GFP or Pacdh-2::GFP strains. Based on the RNAi screen, we expected loss-of-function mutations in each of the three TFs to decrease GFP expression driven by Pacdh-1 and increase expression driven by Pacdh-2 in the intestine. Scale bar, 100 μm.
To validate the RNAi approach, we used mutants of three highly similar nuclear hormone receptor (NHR) paralogs, nhr-68, nhr-101, and nhr-114, which have identical regulatory interaction profiles in the GRN (Table S2). We crossed each mutation into reporter strains where GFP expression is driven by the acdh-1 promoter (Pacdh-1::GFP) or the acdh-2 promoter (Pacdh-2::GFP) and observed very similar effects as we did by RNAi—namely, decreased fluorescence of Pacdh-1::GFP and increased fluorescence of Pacdh-2::GFP, in the intestine (Figure 2B). These results demonstrate that the interactions detected by RNAi were not the result of off-target effects.

Assessing Cell-Autonomous Gene Regulation

Many of the 177 TFs in the GRN are expressed in the intestine, especially those that regulate two or more promoters (Figure S2). However, several TFs are not expressed in the gut but did affect promoter activity in this tissue (Table S3). For instance, LIN-26 is expressed in the hypodermis, not in the intestine, and its perturbation confers hypodermal phenotypes (Labouesse et al., 1994). The observation that intestinal expression of five target promoters is decreased following lin-26 knockdown suggests that LIN-26 regulates one or more genes in the hypodermis that, directly or indirectly, propagate(s) a signal to the intestine to which intestinal genes respond.

To distinguish cell-autonomous from cell non-autonomous effects, we examined the consequences of knocking down TFs only in the intestine. We used an rde-1 mutant strain that is RNAI-defective but in which rde-1 has been rescued only in the intestine (Espelt et al., 2005). As a result, these animals are refractory to RNAI in all tissues except the intestine. We crossed one of the most extensively regulated promoters, Pacdh-1::GFP, into this strain and performed RNAI on 44 of the TFs that regulate this promoter in wild-type animals. Knockdown of 28 of these TFs resulted in a reduction in intestinal GFP levels as in wild-type animals, indicating that they function cell-autonomously to regulate Pacdh-1::GFP (Figure 3A). In this experiment, the other 16 TFs, including LIN-26, did not affect this promoter (Figure 3A). Although we cannot discount differences in RNAI efficiency in wild-type and rde-1-rescued animals, this suggests that these TFs may function non-cell-autonomously to regulate intestinal acdh-1 expression.

Knockdown of some of these TFs may confer phenotypes that, in turn, induce changes in acdh-1 promoter activity. For instance, whole-animal knockdown of ceh-24 caused animals to become paralyzed and dramatically decreased acdh-1 promoter activity. However, intestine-specific ceh-24 knockdown affected neither movement nor acdh-1 promoter activity (Figure 3B). Whole-animal knockdown of ceh-24 likely affects the ability of the animal to eat and induces a starvation state, which is known to result in decreased acdh-1 promoter activity (MacNeil et al., 2013). Indeed, knockdown of ceh-24 also affected two other starvation-response reporters, Pacdh-2::GFP and Psod-3::GFP (Table S2).

Modest Overlap between Regulatory and Physical Interactions

To determine which of the regulatory interactions are likely direct, we compared the GRN to physical interactions between TFs and gene promoters detected by ChIP by the modENCODE project (Gerstein et al., 2014), by previously reported Y1H assays.
Figure 4. A Hierarchical TF Model of Information Flow in the C. elegans Intestine

(A) Overlap between regulatory and physical interactions. Green indicates TF-promoter interactions that are both physical and regulatory. Physical interactions were obtained from Y1H screens performed here and reported elsewhere and from ChIP experiments from the modENCODE project. The top bars represent overlap in all tested interactions. The bottom bars represent interactions involving only TFs that produced positive interactions in both assays. p-values (hypergeometric) are shown.

(B) Schematic depicting the principle of nested effects modeling followed by transitive reduction. Example interaction data are shown with the resulting predicted model. Note that the model includes only TFs and not the target genes. Targets of TF3 are nested within the targets of TF2, whose targets are, in turn, nested within the targets of TF1, resulting in the model shown.
(Arda et al., 2010), or by newly identified by enhanced Y1H assays (Reece-Hoyes et al., 2011; Table S2). In agreement with previous observations (Kemmeren et al., 2014; Yang et al., 2006), we found that only a small subset of physical interactions had an apparent regulatory consequence in the intestine. However, it is important to note that physical interactions may confer a regulatory effect in tissues that were not examined in this study or at developmental times not examined.

We also found that the majority of regulatory interactions detected by RNAi do not have an apparent physical basis (Figure 4A). Although missed interactions (false-negatives) in Y1H assays or ChIP may potentially explain some of this (Walhout, 2011), it is probably not the full explanation because many of the TFs involved in these regulatory interactions were detected in physical interaction screens. Instead, these observations indicate that many regulatory interactions in the GRN are indirect. Indeed, TF cascades or hierarchies have been reported in a number of systems (Jothi et al., 2009; Martinez-Antonio et al., 2012; Yu and Gerstein, 2006). In such hierarchies, the perturbation of a TF that resides high in the hierarchy would affect the expression or activity of TFs in lower tiers.

Inferring Gene Regulatory Information Flow by Nested Effects Modeling

To elucidate the flow of information between TFs, we used nested effects modeling (Markowetz et al., 2007; Figures 4B and 4C). This approach leverages high-dimensional phenotype data for a set of mutants (here changes in fluorescent protein levels in the 19 reporter strains caused by TF knockdown) to infer hierarchical relationships (here among TFs) and results in a “hierarchical TF model.” Briefly, the hierarchy is built by analyzing the overlap in effects caused by perturbation of each TF. TFs that have a larger set of target genes are assumed to be at the top of the hierarchy (Figure 4B). TFs are connected by an edge when the targets of that TF are contained or “nested” within the targets of the TF higher in the hierarchy. The modeling can only robustly place TFs that regulate two or more target genes in the hierarchy (62 of 177; Figure 4C). TFs that were indistinguishable in their regulatory interaction profile were collapsed into “super-nodes” (Figure 4C).

An important question is whether only using 19 target promoters is sufficient to obtain reliable information about regulatory information flow in the C. elegans intestine. We first addressed this question by testing the overall stability of the hierarchical TF model. We removed each individual promoter and its interactions from the GRN, regenerated the hierarchical TF model by nested effects modeling, and compared the fraction of edges maintained relative to the complete model (Figure 4D). In all cases, the majority of edges were retained, indicating that the hierarchical TF model can tolerate minor variations in the set of interactions used to construct it and that using only 19 promoters generates a robust model.

Large-scale computational methods have been used extensively to infer regulatory networks based on co-expression of a TF and its target genes. We used co-expression scores of TF pairs across a large panel of expression profiling data (Reece-Hoyes et al., 2013) to examine co-expression of TFs and target genes as well as between TFs connected in the hierarchical TF model. We did not observe significant co-expression between TFs and the target genes in the GRN. However, we did find that TFs connected by edges in the hierarchical TF model are more co-expressed than TFs that are not connected (Figure 4E; p < 0.003, rank-sum test), suggesting that the TF hierarchical model predicted meaningful relationships between TFs.

Finally, perturbation of nodes at the top of the hierarchy would be expected to have more detrimental effects than the disruption of TFs at the bottom layers. Indeed, the master regulator of intestinal development and gene expression, ELT-2 (McGhee et al., 2009), activates all 19 promoters and is placed at the top of the TF hierarchy (Figure 4C). Furthermore, TFs that do not connect to downstream TFs (the “leaves”) are depleted in lethal phenotypes (p < 0.01; Figure 4F; Table S4). Together, these observations indicate that, by using only 19 test promoters, we robustly captured the information flow between a subset of intestinal TFs.

A Complex Regulatory Path Inferred from the Hierarchical TF Model

To further validate the hierarchical TF model, we focused on the path that connects CDC-5.L to the super-node harboring NHR-68 via SBP-1 and Y111B2A.10, an uncharacterized C2H2 zinc finger TF (Figure 5A). The sbp-1 promoter was part of the 19 test promoters, and, indeed, we found that it is activated by CDC-5.L (Figure 5B). Nested effects modeling does not consider the identity of the promoters used. Therefore, this observation validates the regulatory interaction predicted by the hierarchical TF model. We further explored the phenotype of cdc-5.L given that our GRN and nested effects modeling implicated this uncharacterized gene as an important regulator of intestinal gene expression. Of the 13 promoters regulated by CDC-5.L, 11 are also regulated by SBP-1, suggesting that these TFs have similar effects on gene expression and, perhaps, physiology (Table S2).

We performed RNAi of cdc-5.L in the intestine, which resulted in thin and pale animals, phenocopying sbp-1 knockout (Figure 5C). Although cdc-5.L expression in the intestine was not reported (Table S3), its intestine-specific knockdown affects the acdh-1 promoter and results in dramatic phenotypic consequences, indicating that it does function within the gut
functions within the gut. The hierarchical TF model places promoter activity (Figure 6A), indicating that this TF does function within the gut. The hierarchical TF model places most other TFs that regulate acdh-1 in the intestine upstream of NHR-10. To test whether these TFs regulate acdh-1 indirectly by activating nhr-10, we first generated a Pnhr-10::GFP reporter strain. In agreement with the observation that NHR-10 functions within the intestine, we observed that its promoter drives GFP expression in the gut as well as in other tissues (Figure S3).

We knocked down the 27 TFs that regulate acdh-1 in the intestine and found that 13 TFs also activate the nhr-10 promoter (Figure 6A; Figure S3), suggesting that they may function indirectly to activate the acdh-1 promoter by regulating nhr-10 (Figure S2). One of these TFs, B0336.3, was placed with NHR-10 in a super-node but can now be placed upstream.

To test whether the 13 TFs that regulate both acdh-1 and nhr-10 require NHR-10 to regulate Pacdh-1, we crossed the Pacdh-1::GFP transgene into an nhr-10 deletion (null) mutant (tm4695). Notably, knockdown of each of these 13 TFs reduced GFP expression in the intestine in the absence of NHR-10 (Figure 6A), strongly suggesting that they function in the context of FFLs together with NHR-10 to regulate acdh-1 expression (Figure 6B). Nested effects models are not designed to identify FFLs, and we suspect that there are many instances of FFLs connecting TFs at the top of the hierarchy to other TFs as well as to the target promoters. For instance, ELT-2 physically binds and activates many intestinal gene promoters, including TFs (McGhee et al., 2007, this study) and is therefore also a component of numerous FFLs.

Terminal Nodes in the TF Hierarchy Often Correspond to Physical Interactions
TFs can regulate other TFs by directly regulating their expression but also indirectly by affecting the expression of other genes that influence the activity or stability of the TF (Figure 7A). In this study, we used gene promoters to identify regulators of gene expression, which will ultimately occur through a physical interaction with a TF. Therefore, we predicted that hierarchical TF network paths would terminate at physical interactions between the TFs and promoters.

Indeed, TF-promoter regulatory interactions of terminal nodes (defined here as the last node in a path that regulates a target gene) are 2.5 times more likely to overlap with physical protein-DNA interactions than non-terminal nodes (p = 0.01) (Figures 7A and 7B; Table S5). However, TF-TF edges within the hierarchical model do not significantly overlap with physical interactions between TFs and TF promoters. Therefore, TF-TF edges in the hierarchical model largely involve indirect, post-transcriptional regulation through other regulatory moieties, such as RNA binding proteins, kinases, or phosphatases.

Conclusions
Here we show that combining TF perturbations with promoter reporter strains and nested effects modeling enables the first-pass delineation of the structure of GRNs that control gene expression in a tissue of interest, in living animals. Applying nested effects modeling to the GRN unveiled a complex information flow between intestinal TFs that likely includes non-transcriptional regulatory mechanisms. By integrating physical TF-promoter interactions, GRN building blocks such as FFLs can be incorporated into the networks.

---

**Figure 5. Validation of the Hierarchical TF Model**

(A) A path (orange) in the predicted hierarchical TF model (gray) was selected for validation.

(B) CDC-5.L activates the sbp-1 promoter. Fractions indicate the dilution of bacteria producing double-stranded cdc-5.L RNA with bacteria producing vector alone (to circumvent lethality). DIC images are shown to the left of the fluorescent images.

(C) Systemic knockdown of cdc-5.L has dramatic effects on growth and development (left), whereas intestine-specific cdc-5.L knockdown results in a less severe phenotype that phenocopies sbp-1, consistent with cdc-5.L functioning in the intestine to regulate sbp-1 but also playing roles outside of the intestine.

(D) Knockdown of TFs in a Pnhr-68 reporter strain that expresses nuclear cherry fluorescent protein. Matching DIC photos are shown to the left of the fluorescence images.

Scale bars, 100 μm.

(Figure 3A). Indeed, a recent study has detected cdc-5.L mRNA in the intestine, although at low levels (Blazie et al., 2015).

Using Pnhr-68::H1wcherry as a reporter (Murray et al., 2012), we found that RNAi of either cdc-5.L, sbp-1, or Y111B2A.10 resulted in reduced fluorescent protein expression (Figure 5D), which, together with the above described regulation of Psbp-1::GFP by CDC-5.L, supports the regulatory path predicted by the hierarchical TF model.

**NHR-10 Functions in the Context of Feedforward Loops**

Next we focused on NHR-10, which directly binds to and activates Pacdh-1 (Arda et al., 2010; MacNeill et al., 2013, this study). Intestine-specific nhr-10 knockdown reduced acdh-1 promoter activity (Figure 6A), indicating that this TF does function within the gut. The hierarchical TF model places...
We could only place the TFs that regulate multiple promoters in the model. Therefore, we anticipate that future experiments with additional transgenes will further expand and refine both the GRN and the TF hierarchy. Future RNAi screens with additional types of regulators such as signaling molecules or RNA binding proteins will provide more insight into the detailed mechanisms of indirect regulation for the tested genes. RNAi in living animals with specific sets of gene promoters, combined with nested effects modeling, is a powerful method to derive tissue-specific GRNs and TF hierarchies that connect specific regulatory elements to TF functionality.

**EXPERIMENTAL PROCEDURES**

**TF RNAi Library Construction**

The *C. elegans* genome encodes 934 TFs (Reece-Hoyes et al., 2005) as well as 30 uDBPs (Deplancke et al., 2006), proteins that have been retrieved in Y1H assays but that do not possess a recognizable DNA binding domain. For simplicity, we refer to these combined as TFs. We generated a comprehensive TF RNAi library containing 891 TFs (95%) and all uDBPs. For simplicity, we refer to these combined as TFs. We generated a comprehensive TF RNAi library containing 891 TFs (95%) and all uDBPs.

The TF RNAi library was constructed using open reading frame (ORF) clones from the *C. elegans* ORFeome (Reboul et al., 2003), supplemented with cloned ORFs that were generated in-house. ORFs were cloned into the L4440-Dest-RNAi vector by gateway LR cloning (Walhout et al., 2000). Plasmids containing TF ORFs were purified and transformed into *E. coli* HT115 bacteria, grown in Luria broth (LB) + 15 µg/ml tetracycline + 50 µg/ml ampicillin, and frozen in 5% glycerol in 96-well plates. 71 clones were obtained from the Arringer RNAi library (Kamath et al., 2003), and five were obtained from the ORFeome RNAi library (Rual et al., 2004). GFP- or mCherry-containing clones (positive controls) and empty L4440 vector (negative control) were included in each plate for screening.

**C. elegans Strains**

We used 19 transgenic strains that express a fluorescent protein in the intestine as a starting point for in vivo GRN mapping. These strains harbor gene promoters that drive expression in the intestine. Some reporters are predominantly intestinal, whereas others exhibit a broader expression pattern that includes the gut (Figure S1). Pbb-3::GFP, Pnhr-10::GFP, Ppgd-3::GFP, and Pacs-19::GFP were constructed by gateway cloning of promoterome clones into pDEST-DOD4 (Dupuy et al., 2004). Ppgd-3 is the promoter for *mai-1, gpd-2*, and *gpd-3*, which reside together in an operon. The final constructs express GFP from the promoter of interest and contain a wild-type unc-119 ORF. These constructs were introduced into *unc-119(ed2)* worms by biolistic bombardment (Pravis et al., 2001). The *Pnhr-10::GFP* transgene was integrated into the genome by UV irradiation (Evans, 2006) and outcrossed three times with N2. *nhr-101(gk586) and nhr-114(gk849)* were obtained from the Caenorhabditis Genetics Center (CGC) and outcrossed three times with N2. Following outcrossing, both alleles were crossed to VL749 and VL750 to generate VL1126 [*nhr-101(gk586); wwIs24[Pacdh-1::GFP + unc-119(+)]] and VL1125 [*nhr-114(gk849); wwIs24[Pacdh-2::GFP + unc-119(+)], nhr-68(gk708); wwIs24[Pacdh-1::GFP + unc-119(+)]] and was described previously (Watson et al., 2013). RW10732, which carries the *nhr-68::H1-wCherry* transgene (ssts10507[nhr-68::H1-wCherry + unc-119(+)]), was obtained from the CGC. All other strains have been described elsewhere.

Strains for intestine-specific RNAi were generated by crossing *Pacdh-1::GFP* males with VP303 rde-1(ne219);kbIs7[nhx-2p::rde-1 + rol-6(su1006)] hermaphrodites (Espelt et al., 2005). To select for the rde-1 mutation, F1 animals were placed on lin-26 RNAi.lin-26 RNAi results in completely penetrant larval arrest and lethality when a wild-type copy of *rde-1* is present in the hypodermis. Therefore, viable GFP-positive adult rollers selected in the F2 generation are animals that carry the rde-1 mutation and also carry the intestine-specific *rde-1* as marked by the rol-6-dominant mutant transgene. *Pacdh-1::GFP* and the *rde-1*;rol-6 array were homozygous. The resultant strain, VL1097 [cbIs7[nhx-2p::rde-1 + rol-6(su1006)]];wwIs24[Pacdh-1::GFP + unc-119(+)], was again verified for resistance to lin-26 RNAi.

**RNAi Screening**

We knocked down each individual TF in all strains in three independent experiments. We visually scored increases or decreases in fluorescent protein expression in the intestine at the L4 larvae/young adult stages. Therefore, we retested interactions that were retested were kept in the final dataset. Interactions that were found two or three times were retested, and only interactions that were retested were kept in the final dataset.

Bacteria harboring TF RNAi clones were inoculated into LB + 50 µg/ml ampicillin in 96-well deep-well dishes and grown overnight at 37°C. The following day, fresh cultures were inoculated in 96-well deep-well dishes with 50 µl of overnight culture in 1 ml of LB + ampicillin. Cultures were grown for 6 hr (an OD600 of approximately 0.8), and bacteria were pelleted and resuspended in 1/10 of the original volume in M9 buffer. 10 µl of the bacterial suspension was added to each well of a 96-well nematode growth medium (NGM) agar plate containing 5 mM isopropyl β-D-thiogalactopyranoside (IPTG) and 50 µg/ml ampicillin. Plates were dried and seeded with 20-30 eggs/well. When wells contained a mix of L4 and young adults, animals were screened visually on a fluorescence-enabled dissection microscope for increases or decreases in fluorescence in the intestine. Knockdowns were scored as positive when most animals in the well displayed a change in intestinal GFP or mCherry expression. Changes in expression in other tissues were not recorded. Each transgenic strain was screened three times independently. For each individual screen, bacterial cultures were grown independently, and wells were scored in an unbiased manner; i.e., without knowledge of the results of previous screens. All interactions detected in at least two of three independent screens were tested a fourth time in 24-well plates using 50 animals/well. Clones that failed to pass this retest were not considered true hits. Following the first round of screening, we...
collected RNAi clones that regulated two or more transgenes and rescreened all 19 transgenic strains following knockdown of these factors in 24-well dishes to generate a high-confidence dataset. All RNAi clones included in the final dataset were sequence-verified.

Using eggs as a starting point in the screens enabled us to circumvent embryonic lethality. Nevertheless, knockdown of several TFs resulted in larval arrest or severe sickness. To include these TFs in our screens, we performed serial dilution of bacteria carrying double-stranded RNA (dsRNA) delivery vectors for elt-2, sbp-1, hel-1, let-607, hbl-1, nhr-23, bed-3, cdc-5.1, and lin-26 with bacteria containing the vector alone. By doing so, we were able to observe changes in GFP expression in adult animals.

Enhanced Y1H (eY1H) assays were performed as described previously (Reece-Hoyes et al., 2011). ELT-2 binding to each of the 19 gene promoters was assayed by transforming haploid yeast bait strains with a construct encoding an AD-ELT-2 prey because we found that GATA factors are frequently missed in mating-based eY1H assays.

ChIP data were obtained from the modENCODE website (http://www.modencode.org), and interactions were included as positive when the ChIP peak region observed was within the promoter sequence used in our analysis.

**Intestinally Expressed TFs**
The list of TFs expressed in the intestine was compiled using available data (McGhee et al., 2007; Pauli et al., 2006). From the McGhee data, we selected genes with two or more serial analysis of gene expression (SAGE) counts. Additional intestinal expression information was obtained by manual curation of expression patterns available in WormBase (http://www.wormbase.org) and other published reports. TFs were considered intestinally expressed when they were retrieved in either gene expression dataset or when an expression pattern was reported that included the intestine. TFs were considered “not intestinal” when an expression pattern was reported that did not include the intestine.

**Lethal Phenotypes**
Lethal phenotypes were curated based on data reported by WormBase (http://www.wormbase.org). Genes annotated with the following phenotypes were considered to confer a lethal phenotype: larval lethal, lethal, and embryonic lethal.

**Generation of Hierarchical TF Models**
Models were generated using the nested effects modeling algorithm (Markowetz et al., 2007) using a false positive (FP) rate of 0.01 and a false negative (FN) rate of 0.469. These rates were determined based on the reproducibility of interactions across individual screens. Briefly, we considered as “gold standard” those RNAi interactions that were found in at least two of the three replicate screens. These gold standard interactions could then be used to estimate the FP and FN rate for each individual RNAi screen. Specifically, interactions that were only detected in one of the replicates were considered FP, and interactions that were missed in one of the replicates but found in the other two were designated FN. These gold standard interactions could then be used to estimate the FP and FN rate for each individual RNAi screen. Specifically, interactions that were only detected in one of the replicates were considered FP, and interactions that were missed in one of the replicates but found in the other two were designated FN. The FP and FN estimates were then averaged across the three screens to derive average per-screen FP and FN rates. To estimate the FP and FN rates in the final dataset, i.e., after requiring the interactions to appear in two of the three initial screens as well as the retest (see above), we calculated the probability of observing false negatives or false positives based on the per-screen quality estimates, assuming independence of the three screens and retest, which resulted in final FP and FN rates of 0.0005 and 0.469, respectively. It is important to note that the true, biological false negative rate is likely to be lower than this estimate because the final retest involved a matrix experiment in which each transgenic strain was tested versus RNAi of...
each the 63 TFs that regulate at least two promoters (see above). We used a conservative FP estimate of 0.01 for the nested effects modeling to make the modeling process more flexible. Additionally, the “triples threshold” parameter was set to 0.95 to ensure a small degree of flexibility during modeling. We performed a transitive reduction of edges to facilitate the visual interpretation of the resulting hierarchical network model. The model was generated by using data from all TFs with at least two RNAi interactions (except hel-1, which likely regulates RNA export) (MacMorm et al., 2003). Importantly, the resulting model is highly robust because the global structure is largely maintained at a range of different error rates (i.e., more than 85% of the edges are included). Furthermore, at different FN rates (0.45, 0.40, 0.30, and 0.20), we consistently observed enriched co-expression between TFs connected by an edge in the hierarchical models as well as an enrichment of lethal phenotypes in non-terminal TFs. Therefore, our results are not sensitive to the selected parameter settings used in the modeling.

**Terminal Node Analysis**

We hypothesized that TFs lower in the hierarchy may be more likely to affect the expression of reporters by direct physical binding, whereas TFs placed higher in the hierarchy would more likely affect the target genes indirectly through a cascade of interactions with TFs in lower levels. To test this hypothesis, we contrasted the overlap of RNAi interactions with protein-DNA interactions between terminal and non-terminal TFs of the intestine-specific TF hierarchy. For each of the reporters, we defined a sub-graph of the TF model consisting of TFs that reported an RNAi interaction with that reporter. In each sub-graph, we defined as terminal nodes those that did not have any outgoing edge (i.e., the leaves of the sub-graph) and the remaining nodes as non-terminal (Table S3). Sub-graphs containing only a single node were not included. Next, we calculated the total number of RNAi interactions that overlapped with physical interactions for terminal and non-terminal nodes across the subgraphs and performed a hypergeometric test to evaluate whether the overlap was significantly higher among terminal nodes.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes three figures and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.cels.2015.08.003.

**AUTHOR CONTRIBUTIONS**

L.T.M. and A.J.M.W. designed the study. L.T.M. and E.A. generated subgraphs and performed the PacDh-1::GFP screen. G.G. integrated the Phnr-10::GFP transgene into the genome. All other experiments were performed by L.T.M. C.P. performed all modeling under supervision from C.L.M. L.T.M. and A.J.M.W. wrote the paper. All authors discussed the results and edited the manuscript prior to submission.

**ACKNOWLEDGMENTS**

We thank members of the A.J.M.W. laboratory and Sander van den Heuvel for discussions and critical reading of the manuscript. Some nematode strains were provided by the CGC, which is funded by the NIH Office of Research Infrastructure Programs (P40 OD010440). This work was supported by NIH grant GM082971 (to A.J.M.W. and C.L.M.) and a CIHR post-doctoral fellowship (to L.T.M.).

Received: April 17, 2015
Revised: June 12, 2015
Accepted: August 4, 2015
Published: August 26, 2015

**REFERENCES**


Transcription Factor Activity Mapping of a Tissue-Specific In Vivo Gene Regulatory Network

Lesley T. MacNeil, Carles Pons, H. Efsun Arda, Gabrielle E. Giese, Chad L. Myers, and Albertha J.M. Walhout
Supplemental Tables and Figures

**Table S1:** Description of *C. elegans* strains and RNAi clones in the library. See also Figure 1.

**Table S2:** List of all regulatory and physical interactions with target promoters. See also Figure 2.

**Table S3:** Summary of TF expression data. NR indicates that the transcript was not reported in intestinal expression profiling data and no transcriptional reporter strain has been reported. See also Figure 3.

**Table S4:** Lethal phenotypes reported following perturbation of TFs in the hierarchical TF model. See also Figure 4.

**Table S5:** Interactions that are both physical and regulatory in the hierarchical TF model. Nodes are categorized as terminal or non-terminal based on their position in the subnetworks for each individual target gene. See also Figure 7.

**Figure S1:** Expression pattern of transgenes used for identification of regulatory interactions. Green box indicates expression observed. Expression in neurons is marked as positive if expression was observed in any neuron. See also Figure 1.

**Figure S2:** TFs in the GRN are enriched for intestinal expression. Pie charts indicate fraction of TFs in the GRN with intestinal expression. See also Figure 2.

**Figure S3:** GFP expression following knockdown of indicated factors in *Pnhr-10::GFP* transgenic animals. See also Figure 6.
MacNeil et al., Figure S1

<table>
<thead>
<tr>
<th></th>
<th>Intestine</th>
<th>Muscle</th>
<th>Hypodermis</th>
<th>Pharynx</th>
<th>Neurons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psbp-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pmdl-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacdh-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacdh-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PF35E12.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pbli-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pgst-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plet-7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ptrap-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peft-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pmir-71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pmir-63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacs-19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ppqm-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pgpd-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phsp-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phsp-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psod-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psod-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pnhr-178</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MacNeil et al., Figure S2

- All TFs: 49% Intestinal, 59% Not intestinal, 177 total
- All hits: 59% Intestinal, 49% Not intestinal, 177 total
- k-out >1: 78% Intestinal, 22% Not intestinal, 63 total

Legend: Green = Intestinal, Blue = Not intestinal, Gray = Not reported
MacNeil et al., Figure S3

Pnhr-10::GFP

vector

B0261.1

B0336.3

mxl-3

sbp-1
cdc-5.L

hmg-4

100 µM

F43C11.7

F43G9.12

let-607

nhr-17

100 µM