University of Massachusetts Medical School

eScholarship@UMMS

Complex expression dynamics and robustness in C. elegans insulin networks

Ashlyn D. Ritter
University of Massachusetts Medical School

Et al.

Let us know how access to this document benefits you.

Follow this and additional works at: https://escholarship.umassmed.edu/sysbio_pubs

Part of the Genetics and Genomics Commons, and the Systems Biology Commons

Research

Complex expression dynamics and robustness in *C. elegans* insulin networks

Ashlyn D. Ritter,[1,2,3] Yuan Shen,[2,3,5] Juan Fuxman Bass,[1,3] Sankarganesh Jeyaraj,[2,3,6] Bart Deplancke,[2,3,7] Arnab Mukhopadhyay,[2,3,8] Jian Xu,[4] Monica Driscoll,[4] Heidi A. Tissenbaum[2,3,9] and Albertha J.M. Walhout[1,3,9]

[1]*Program in Systems Biology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA;* [2]*Program in Gene Function and Expression, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA;* [3]*Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA;* [4]*Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08855, USA*

Gene families expand by gene duplication, and resulting paralogs diverge through mutation. Functional diversification can include neofunctionalization as well as subfunctionalization of ancestral functions. In addition, redundancy in which multiple genes fulfill overlapping functions is often maintained. Here, we use the family of 40 *Caenorhabditis elegans* insulins to gain insight into the balance between specificity and redundancy. The insulin/insulin-like growth factor (IIS) pathway comprises a single receptor, DAF-2. To date, no single insulin-like peptide recapitulates all DAF-2-associated phenotypes, likely due to redundancy between insulin-like genes. To provide a first-level annotation of potential patterns of redundancy, we comprehensively delineate the spatiotemporal and conditional expression of all 40 insulins in living animals. We observe extensive dynamics in expression that can explain the lack of simple patterns of pairwise redundancy. We propose a model in which gene families evolve to attain differential alliances in different tissues and in response to a range of environmental stresses.

[Supplemental material is available for this article.]

Organisms must respond appropriately to physiological and environmental cues to develop, locate food, stave off toxic or pathological insults, and survive stress. The ability to adjust to the environment or to overcome stress requires the use of robust regulatory mechanisms that coordinate systems-level physiological responses. For instance, growth, reproduction, and lifespan can be adjusted in response to changes in food availability and environmental conditions such as temperature (Riddle et al. 1997; Marsh and May 2012; MacNeil et al. 2013).

Expanded gene families often accompany an increase in organismal complexity. After duplication, the resulting identical paralogs diverge in both sequence and function. Neofunctionalization occurs when one paralog adopts a novel function that is not shared by the ancestral gene (Ohno 1970), while in subfunctionalization the ancestral function(s) are divided between the two paralogs (Force et al. 1999). In addition to functional divergence, it can also be beneficial for paralogs to maintain (partial) redundancy, for instance, to provide a buffer under adverse genetic or environmental conditions (Burga et al. 2011; MacNeil and Walhout 2011).

Numerous gene families have been identified by whole-genome sequencing and gene annotation. So far, only a few families

have been analyzed at a systems level, and the balance between divergence and redundancy is only beginning to be illuminated. For example, in *C. elegans*, basic helix-loop-helix transcription factors have extensively diverged in multiple functional parameters (Grove et al. 2009). However, ETS transcription factors function redundantly in the mouse by binding and regulating the same target genes (Hollenhorst et al. 2007).

The highly conserved insulin/insulin-like growth factor signaling (IIS) pathway plays a critical role in development, lifespan, stress resistance, and metabolism (Kenyon 2011). The *C. elegans* IIS pathway includes insulin-like peptides, a receptor (DAF-2), and a downstream signaling cascade that impinges on the DAF-16 transcription factor (Kenyon 2011). Loss-of-function mutations in *daf-2* confer a range of phenotypes affecting lifespan, reproduction, metabolism, stress, and dauer formation, and a complete loss of the receptor is lethal (Gems et al. 1998; Tissenbaum and Ruvkun 1998). Dauer is a diapause stage that is evolutionarily conserved among nematodes (Sommer and Streit 2011). When growth conditions are suboptimal, *C. elegans* pheromone levels increase and can utilize the IIS pathway to enter dauer (Riddle et al. 1997). Although dauer formation has been studied extensively by modulating downstream IIS signaling components, less is known about the role of insulin-like genes in this process.

The human genome encodes 10 insulin-like peptides, including insulin, insulin-like growth factors (IGFs), and relaxins (Claeys et al. 2002). *Drosophila melanogaster* has eight insulin-like peptides, or DILPs (Colombani et al. 2012). Remarkably, the *C. elegans* genome encodes 40 insulin-like genes (Duret et al. 1998; Pierce et al. 2001; Li et al. 2003; Li and Kim 2008). Hereafter, we refer to insulin-like peptides or genes as "insulins." To date, no single insulin mutation or deletion completely recapitulates the phenotypes associated with perturbation of *daf-2* (Supplemental Table S1). In *Drosophila*, genetic interaction analyses have revealed

functional redundancy between a subset of DILPS (Broughton et al. 2008; Gronke et al. 2010). The lack of loss-of-function phenotypes for most insulins indicates that they also likely act redundantly in *C. elegans*, although to what extent is unknown.

Genes that act redundantly are often coexpressed, at least in part. Therefore, we decided to annotate spatiotemporal and conditional insulin expression as a starting point to indicate which insulins may function together. *C. elegans* are transparent, and transgenic animals carrying promoter-driving green fluorescent protein (GFP) constructs can be used to annotate spatiotemporal gene expression in living animals (Chalfie et al. 1994; Reece-Hoyes et al. 2007; Martinez et al. 2008; Grove et al. 2009). Here, we compare spatiotemporal insulin expression during development, aging, and under a variety of nutritional and stress conditions. We observed both specific and overlapping insulin expression patterns, as well as numerous dynamic expression changes. Altogether, our data indicate a potential for complex alliances between insulins.

Results

Lack of pairwise redundancy between *C. elegans* insulins in dauer formation

We used dauer diapause, one of the main outputs of the IIS pathway, to explore the putative extent of redundancy and specificity in the insulin family. Loss-of-function mutations in the DAF-2 receptor result in a dauer-constitutive (Daf-c) phenotype (Kenyon 2011). We found that neither deletions in 12 insulins nor knockdown of any of the 40 insulins results in dauers (Fig. 1A; Supplemental Table S2). As some insulins have been reported to have antagonistic effects on the IIS pathway, we asked whether loss of individual insulins would prevent dauer formation. After dauer-inducing conditions, all mutants analyzed were capable of forming dauers, indicating that they are not dauer defective (Fig. 1B). Together, these results indicate that no single insulin is the sole agonist or antagonist for coordinating dauer formation through the DAF-2 receptor. Next, we asked whether insulins may function in redundant pairs by knocking down each of the 40 insulins in the 10 insulin mutants. However, none of the combinations of insulin mutants and insulin RNAi resulted in the formation of dauers (Fig. 1A). This suggests that
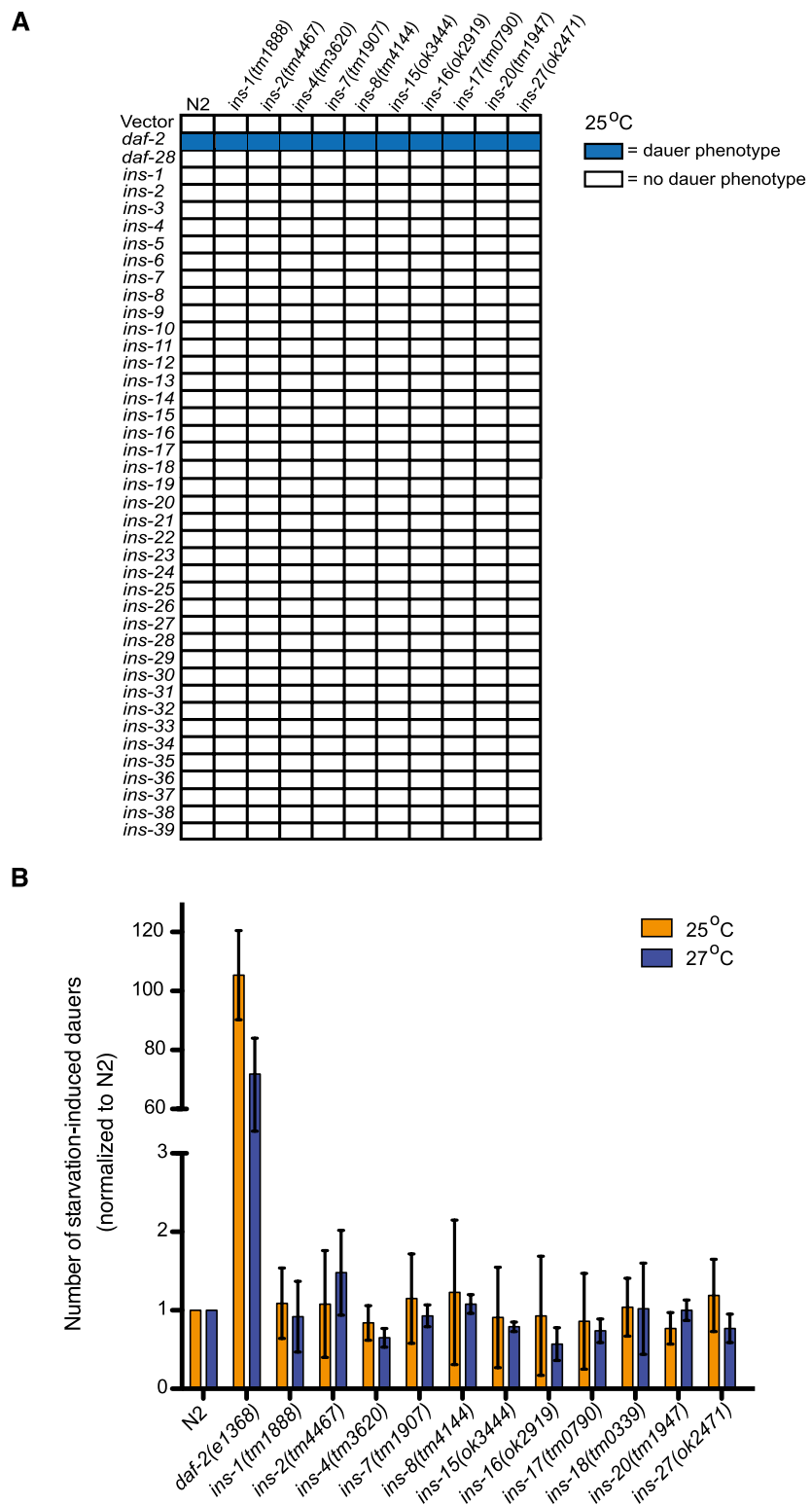
Figure 1. Examination of dauer formation upon insulin perturbation. (*A*) Neither loss of individual insulins (columns), nor pairwise insulin perturbation by combined mutation and knockdown (rows) results in dauer formation. (*B*) Individual insulin mutants form dauers in starvation-induced conditions similar to wild type (N2), indicating that they are not dauer defective. Dauer formation is illustrated as the average of duplicate experiments; data is normalized to wild-type (N2) dauer formation. Additional data can be found in Supplemental Table S2.

there is no simple pairwise redundancy between insulins in regulating dauer formation, but rather that more complex patterns of redundancy may be involved.

Complex patterns of insulin expression

To determine spatiotemporal insulin expression patterns during development and under a variety of relevant physiological and environmental conditions, we cloned the promoters of each of the 40 *C. elegans* insulins upstream of an open reading frame encoding GFP, and generated transgenic animals as described (Praitis et al. 2001). Promoter activity is often a faithful proxy of endogenous gene expression (Reece-Hoyes et al. 2007; Martinez et al. 2008; Grove et al. 2009); therefore, we refer to insulin promoter activity as "insulin expression."

We annotated insulin expression using a controlled vocabulary to enable the direct comparison between different genes and between different tissues (Martinez et al. 2008; Grove et al. 2009). In hermaphrodites, 36 of 40 insulin promoters conferred GFP expression in six major tissue types: nervous system, muscle, reproductive tissue, epithelia, the alimentary system, and coelomocytes (Fig. 2A). These were subcategorized into specific cells/tissues, totaling 19 different tissue/cell types (Fig. 2B). We found that most tissues express multiple insulins (Fig. 2B), and, conversely, that most insulins are expressed in multiple tissues (Fig. 2C). Four insulin promoters did not drive GFP expression in hermaphrodites: *ins-19*, *ins-24*, *ins-31*, and *ins-37*. However, *ins-19* and *ins-31* did exhibit GFP expression in males (Thoemke et al. 2005; Supplemental Fig. S1). Altogether, we detected activity for 38 of 40 promoters (95%).

Spatiotemporal expression patterns have previously been reported for 17 insulins (Pierce et al. 2001; Li et al. 2003; Murphy et al. 2003, 2007; Michaelson et al. 2010; Cornils et al. 2011). Our data recapitulate the expression patterns for nine of these (Supplemental Table S3): *ins-1, ins-3, ins-4, ins-7, ins-8, ins-9, ins-11, ins-18,* and *ins-33*. However, for all nine we observed expression in additional tissues. For seven insulins we found different expression patterns than reported previously. For example, we observed *ins-22* expression in head neurons, but not in the ventral nerve cord, body, and tail neurons as reported previously (Pierce et al. 2001). Differences in insulin expression may result from the inclusion or exclusion of particular regulatory sequences in the promoter fragments used. Alternatively, expression pattern annotation may not have been examined with the same resolution in different studies and weak expression may have been missed. Indeed, for two strains we observed additional expression patterns in previously published strains that were not reported (Supplemental Fig. S2).

Next, we examined the expression patterns for the 36 insulins expressed in hermaphrodites. Almost all insulins are expressed in neurons, consistent with the reported importance of neuronal insulin signaling in different organisms (Wolkow 2002). Head neurons express the greatest number of insulins ($n = 32$) (Fig. 2B). We grouped head neuron expression into three classes: sensory (including amphid neurons), pharyngeal, and ring motor/interneurons (Fig. 2D). As amphid sensory neurons are exposed to the environment, many can be visualized by DiI staining (Tong and Bürglin 2010). We stained all transgenic strains with DiI and annotated any overlap between DiI staining and GFP expression. DiI staining was also used to provide orientation for determining which insulins were expressed in each category of head neurons (Fig. 2E). In total, 25 insulins are expressed in sensory neurons, 14

of which are expressed in amphid neurons, 23 insulins in ring motor/interneurons, and 13 in pharyngeal neurons. Twenty-two insulins are expressed in two or more general types of head neurons (Fig. 2F).

Dynamic changes in insulin expression during development

To determine insulin expression changes during development, we annotated GFP expression at each larval stage (L1–L4) in young adults and mature adults. A value of 1 or 0 was assigned to each tissue/cell type to indicate the presence or absence of GFP expression, respectively (Supplemental Table S4). To capture tissue expression changes through development, we visualized the data into bipartite expression networks (Martinez et al. 2008; Grove et al. 2009). These networks contain two types of nodes: genes and cells/tissues. Edges link the gene to the tissue in which it is expressed. We visualized insulin expression patterns into networks for each of the developmental stages by using Cytoscape (Fig. 3A; Shannon et al. 2003; Supplemental Fig. S3).

Several tissues showed loss of insulin expression as the animals progress through development. For example, in 11 transgenic strains, GFP was expressed in the pharynx in early larval stages, but only three insulin promoters remained active in the adult pharynx (Fig. 3B). Most tissues showed increased insulin expression during development, particularly when animals entered the L4 stage. For example, 14 insulins are expressed in the distal tip cell at the young adult stage but only one (*ins-13*) exhibited expression early, at the L2 stage (Fig. 3B). Some tissues/cells are formed later in development, which explains later expression; for instance, the vulva, spermatheca, uterus, and gonad have not fully developed until late-L4 and adult stages. Overall, insulins exhibit dynamic stage- and tissue-specific expression patterns during development (Supplemental Fig. S4).

### Conditional dynamics of insulin expression

The IIS pathway not only modulates *C. elegans* development, it is also utilized to respond to environmental conditions (Kenyon 2011). We investigated insulin expression in young adult transgenic strains exposed to relevant conditions including heat stress, dauer diapause, starvation, aging, oxidative stress, and glucose, and visually examined changes in GFP expression (Fig. 4A; Supplemental Fig. S5A).

Neither glucose nor oxidative stress affected insulin expression. However, we did observe changes in insulin expression with the other conditions. The greatest number of insulins changed expression in dauers ($n = 14$) and with age ($n = 14$), major phenotypic outputs of IIS pathway. Except for heat stress, increased and decreased insulin expression was observed. For example, 10 insulins decreased in expression in dauer, whereas four increased. We also observed stage-specific responses in insulin expression. For instance, GFP expression increased dramatically after heat stress in L4 P*ins-4::GFP* hermaphrodites, but not in young adults (Supplemental Fig. S5B). In total, four genes, *ins-3, ins-7, ins-11,* and *ins-30*, exhibited changes in GFP expression in three conditions tested; four changed in two, 17 in one, and 14 did not change in any of the conditions tested.

We also observed specific conditional changes in spatial expression (Supplemental Fig. S6). P*ins-27::GFP*, for example, exhibits GFP expression in muscle in all stages of development, and muscle expression decreased as the animals aged (Fig. 4B). This could reflect a loss of muscle integrity known to occur during the
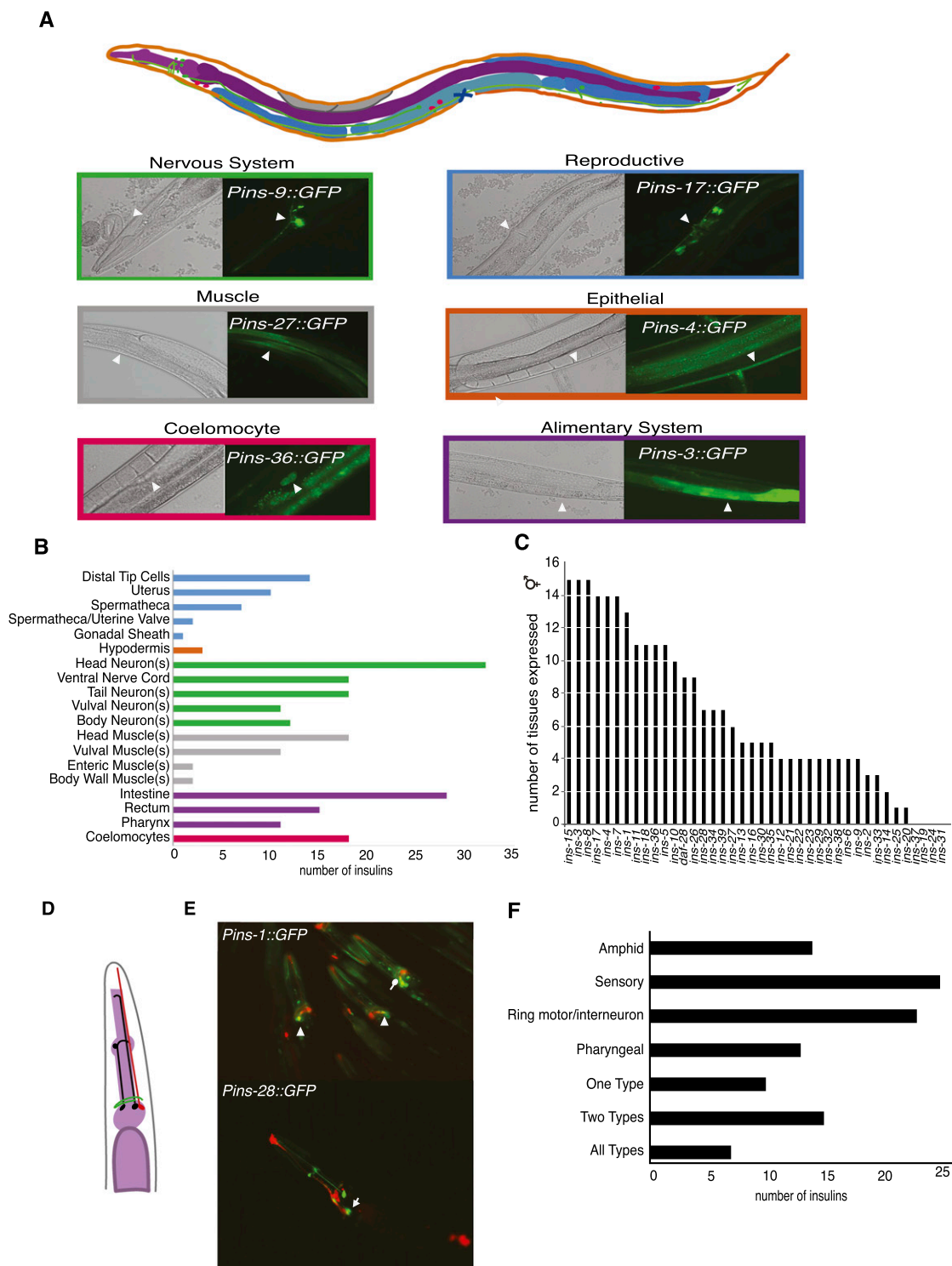
Figure 2. Spatial expression of the 36/40 insulins expressed in *C. elegans* young adult hermaphrodites. (*A, top*) Cartoon of a *C. elegans* hermaphrodite with colors illustrating insulin expression in six major tissue types: nervous system (green), reproductive tissue (blue), muscle (gray), epithelial tissue (orange), coelomocytes (pink), and alimentary system (purple). (*B*) Number of insulins expressed in each tissue/cell type. Colors correspond to the six major tissue types in *A*. (*C*) Number of tissues/cell types in which each insulin is expressed. (*D*) Cartoon of three head neuron categories: sensory (red), pharyngeal (black), ring motor/interneurons (green). (*E*) Two examples: *Pins-1::GFP* exhibits GFP expression in many head neurons including amphid sensory neurons (arrowheads) and ring interneurons (circle arrow). *Pins-28::GFP* exhibits GFP expression in pharyngeal, but not amphid neurons (arrows). Amphid neurons were visualized by DiI staining (red). Yellow indicates that GFP expression (green) occurred in amphid neurons (red). (*F*) Summary of neuronal insulin expression in three types of head neurons: sensory, ring motor/interneuron, and pharyngeal. Insulins that exhibited overlapping GFP expression and DiI staining are defined as having amphid expression.
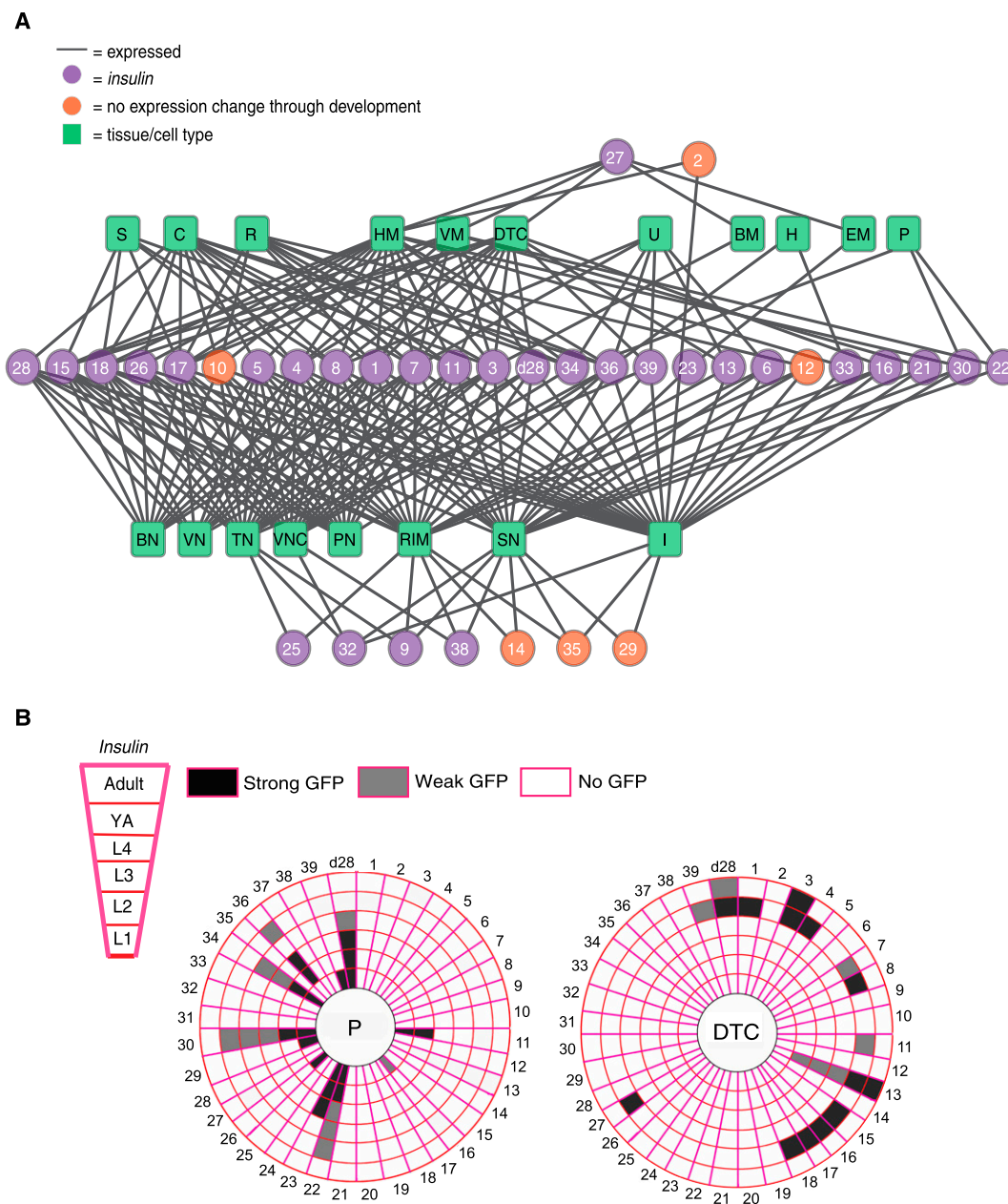
Figure 3. Insulin expression networks illustrate dynamic tissue activity. (*A*) Bipartite network connecting 35 insulins to cell/tissues at the young adult stage. *Pins-20::GFP* animals did not exhibit GFP expression at the young adult stage and were excluded. Circles represent insulins; squares represent cells/tissues. (S) Spermatheca; (C) ceolomocytes; (R) rectum; (HM) head muscle; (VM) vulva muscle; (DTC) distal tip cell; (U) uterus; (BM) body muscle; (H) hypodermis; (EM) enteric muscle; (P) pharynx; (BN) body neuron(s); (VN) vulva neuron(s); (TN) tail neuron(s); (VNC) ventral nerve cord; (PN) pharyngeal neuron(s); (RIM) ring/intermotor neuron(s); (SN) sensory neuron(s); (I) intestine. Purple circles indicate insulins that change in spatial expression through development; orange circles are insulins that do not change. The network is organized as follows: (*top* to *bottom*) insulins with no neuronal expression; non-neuronal tissues (excluding the intestine); insulins with neuronal and non-neuronal tissue expression; neuronal tissues and the intestine, insulins with only neuronal and/or intestinal expression. (*B*) "Dartboards" depicting tissue-centered view of insulin expression. Each ring represents a developmental stage, starting from the center: larval stages L1, L2, L3, and L4, young adult, and adult with eggs (~3- to 4-d-old adult). Each slice represents an insulin with strong GFP (black), weak GFP (gray), or no GFP (white) expression. (*Left*) Dartboard profile of the pharynx (P); (*right*) dartboard profile of the distal tip cell (DTC).

aging process (Herndon et al. 2002), and perhaps the *Pins-27::GFP* transgenic strain may provide a convenient marker to observe the aging process in living animals. Another example is in *Pins-4::GFP* animals, where GFP expression changes similarly in different conditions; it increases dramatically in the hypodermis of young adults after starvation and aging animals (Fig. 4B).
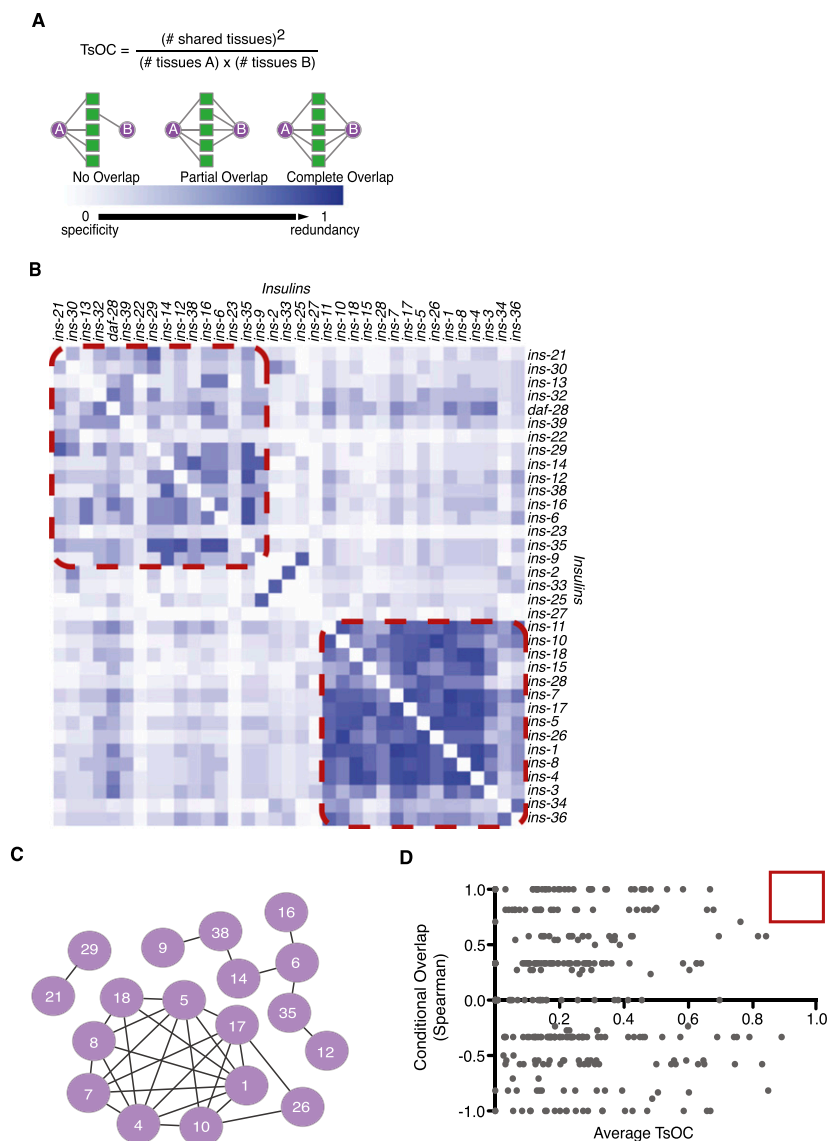
Comparing insulin expression patterns

We quantified the degree of expression overlap for pairwise combinations of insulins at each stage of development using a tissue overlap coefficient (TsOC) similarity index (Fig. 5A; Martinez et al. 2008; Grove et al. 2009). TsOC values were clustered to visually
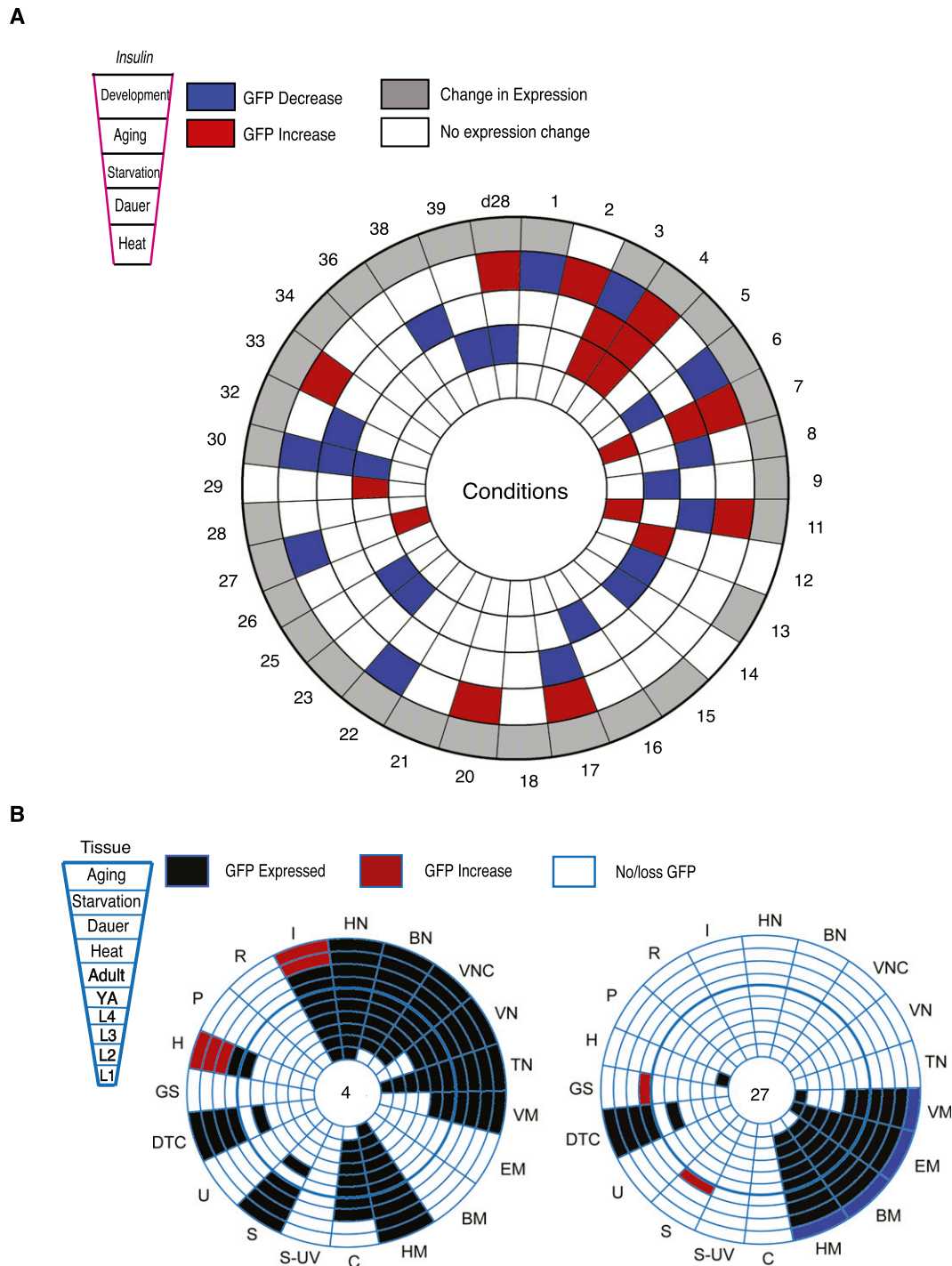
Figure 4. Conditional dynamics of insulin expression. (*A*) Dartboard depicting conditional changes in insulin expression. Each ring represents a condition; starting from the center, heat stress, dauer, starvation, aging, and development. Each slice represents a different insulin. For each condition, activity is defined by at least one tissue exhibiting an increase (red) or decrease (blue) in GFP expression. Changes in expression pattern (gray) were recorded throughout development. (*B*) Dartboards depicting gene-centered view of insulin expression. Each ring represents a developmental stage or condition, starting from the center: L1, L2, L3, and L4, young adult, and adult (~3- to 4-d-old adult), heat stress, dauer, starvation, and aging. Each slice represents a tissue with GFP expression (black), no GFP (white) expression. Compare conditional changes to the young adult stage (fifth ring). Red indicates a GFP increase; black, no change in GFP expression; white, no GFP observed. (*Left*) Dartboard of *ins-4* expression; (*right*) dartboard of *ins-27* expression.

identify genes that share high overlap versus those with little to no overlap in expression. At the young adult stage, we observed two clusters of insulins that share high overlap (Fig. 5B). The lack of pairwise gene expression clustering correlates with the lack of pairwise phenotypic redundancy, indicating that there may be more complex patterns of redundancy. We next investigated to what degree expression overlap changes through development. Of 630 possible insulin pairwise comparisons (of 36 insulins

Figure 5. Comparing insulin gene expression illustrates specificity and overlap in expression. (A) Tissue overlap coefficients (TsOCs) define the expression overlap between all pairwise insulins (A,B). Insulins (circles) exhibit GFP expression (edges) in *C. elegans* tissues (squares). A TsOC of 1 indicates complete overlap while a TsOC of 0 indicates no overlap in expression. (B) Tissue overlap matrix using TsOC scores from the young adult stage. Red boxes highlight visually delineated clusters of highly overlapping insulins. (C) Coexpression network of insulins (circles) that share the top 5% average TsOC scores through six larval/adult stages. An edge represents an average TsOC ≥ 0.68; maximum average TsOC = 0.89 (*ins-4* and *ins-5*). (D) No gene pairs (red box) share the same expression in both conditional (as measured by a Spearman rank test) and developmental tissue expression. TsOC is presented as an average of the six TsOC values for each developmental stage.

To capture the overall degree of overlap of the insulin expression patterns throughout the lifetime of the animal, we averaged the TsOCs between all stages. The top 5% TsOC values (TsOC ≥ 0.68) were distributed among 18 genes (Fig. 5C). Some insulins (e.g., *ins-4*) shared high overlap with many other genes in this group, while others (e.g., *ins-26*) shared high overlap with only two other genes.

Next, we quantified similarity in conditional expression. We first assigned the following values for each condition: −1 for a GFP decrease/loss, 0 for no changes, and +1 for a GFP increase/additional tissue expression. Then, using a Spearman rank test, we compared all pairwise combinations to identify insulins that correlate in their response. A conditional overlap of 1, for example, indicates that two insulins responded the same under every condition tested. In contrast, a score of −1 indicates that two insulins responded oppositely. Of 325 pairwise combinations (among 26 insulins that exhibit a conditional change), 23 insulin pairs, consisting of 18 genes, shared a score of −1, thus responding oppositely. Thirty-six pairs, consisting of 18 genes, shared a score of +1, thus responding the same under the conditions tested. Finally, no pair exhibited high correlation both during development and under different conditions (Fig. 5D, red box). This suggests that no insulins are completely coexpressed over all conditions and stages.
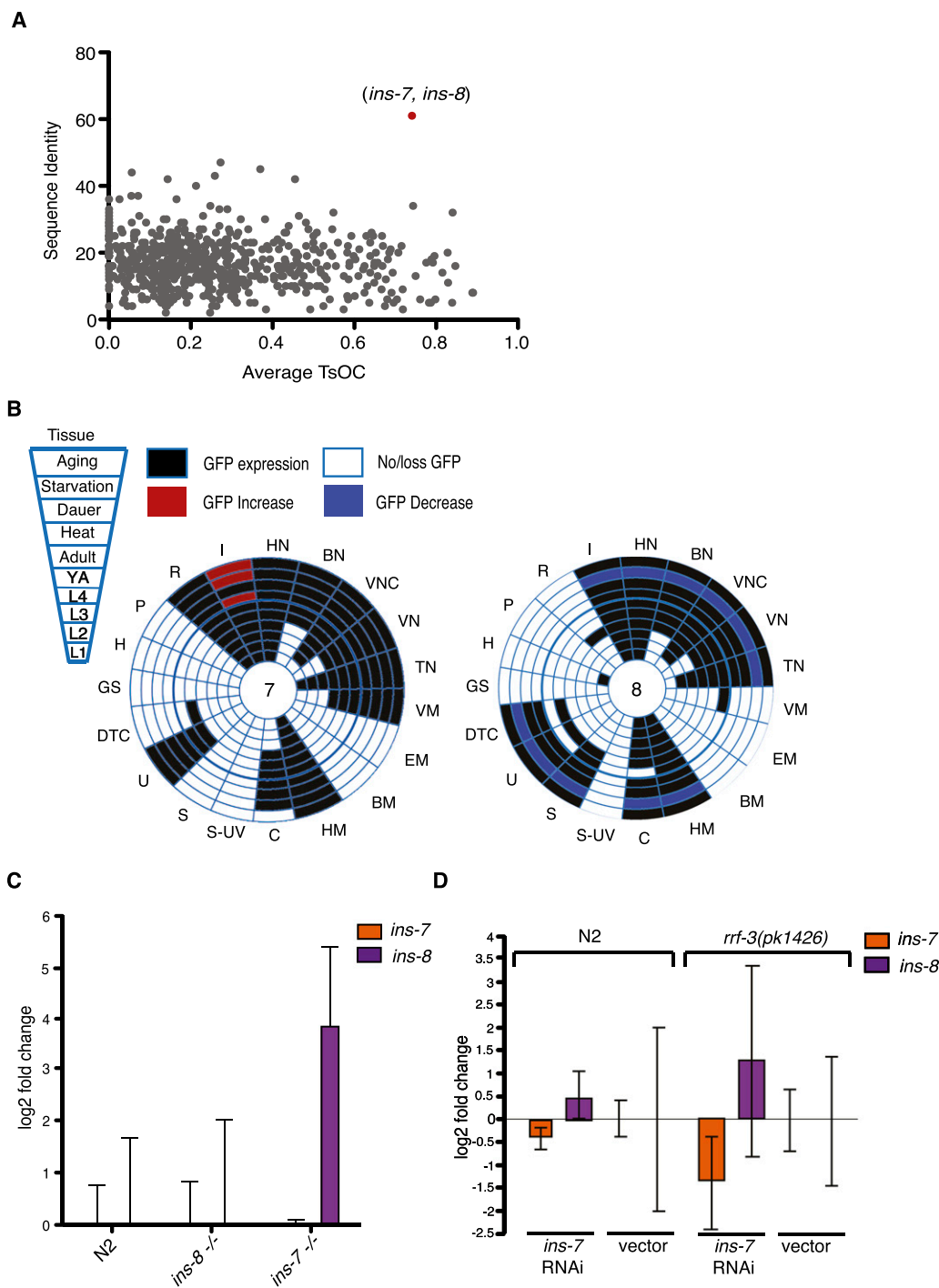
ins–8 expression increases upon loss of *ins–7*

Intuitively, one might expect that close homologs that share high-expression overlap may exhibit redundancy. When we compared pairwise insulin protein sequence identity with their average TsOC scores, we did not observe a correlation between sequence and expression similarity (Fig. 6A).

One outlying pair of insulins, *ins-7* and *ins-8*, is both similar in sequence and in spatiotemporal expression (Fig. 6A,B). To test putative redundancy between these insulins, we first examined developmental timing and dauer formation in single *ins-7(tm1907)* or *ins-8(tm4144)* mutants and found no differences compared to wild-type animals (Fig. 1; data not shown). Close paralogs can compensate each other's loss by an increase in expression (Raj et al. 2010; Burga et al. 2011). We examined endogenous *ins-7* expression in an *ins-8* mutant and vice versa and found that while loss of *ins-8* did not result in a change in *ins-7* expression, loss of *ins-7* resulted in a ~15-fold increase in *ins-8* expression (Fig. 6C). In a recent study it was reported that endogenous *ins-7* is

examined), 37 pairs, consisting of 22 genes, had a TsOC of zero throughout development, illustrating their mutually exclusive expression. There was no insulin that exhibited a unique expression pattern (i.e., with a TsOC = 0 with every other insulin). Further, although 11 insulin pairs consisting of 19 insulins exhibited complete overlap (TsOC = 1) in a single larval stage, none exhibited complete overlap throughout development. Even though the degree of overlap between insulins changed from one larval stage to the next, most remained consistently higher or lower in their degree of expression overlap (Supplemental Table S5).

Figure 6. Divergence, overlap, and compensation in expression between *ins-7* and *ins-8*. (*A*) There is no clear positive correlation between the degree of overlap in expression through development (average TsOC) and sequence similarity of insulins. *ins-7* and *ins-8* (red dot) are the only pair that share both high-sequence similarity and high TsOC. TsOC is presented as an average of the six TsOC values for each developmental stage. (*B*) *ins-7* and *ins-8* exhibit high overlap in expression under standard laboratory conditions (first six rings, from L1 to adult), but exhibit different conditional changes in expression (four outermost rings). (Red) GFP increase; (blue) GFP decrease; (black) no change in GFP expression; (white) no GFP observed. (*C*) In the absence of *ins-7* (orange), *ins-8* (purple) expression is increased. (*D*) *ins-8* increases in both wild-type (N2) (*P* < 0.0001) and *rrf-3(pk1426)* mutants (*P* = 0.054) upon RNAi knockdown of *ins-7*. mRNA abundance was measured by qRT-PCR for *C* and *D*. Triplicate repeats from biological duplicate samples were measured. One biological sample is illustrated here. The error bar indicates variation between triplicates. *P*-values were calculated using a Student's *t*-test.

expressed at higher levels than *ins-8* throughout development (Baugh et al. 2011). *ins-7* and *ins-8* are less than 1 kb apart in the genome; therefore, in the *ins-7(tm1907),* the *ins-8* gene resides in closer proximity to the (stronger) *ins-7* promoter (data not shown). To test whether this could explain the observed increase in expression, we examined *ins-8* expression upon *ins-7* RNAi in both

wild-type (N2) and RNAi hypersensitive *rrf-3(pk1426)* mutants (Simmer et al. 2002). Again, we observed an increase in *ins-8* expression upon perturbation of *ins-7*, albeit to a lesser extent (Fig. 6D). Therefore, the up-regulation of *ins-8* in the *ins-7* mutant background is not due to the close proximity of the two genes. This data also illustrates that RNAi targeting *ins-7* does not elicit a complete knockdown.

ins-7 expression increases upon dauer formation (Fig. 6B). However, *ins-7* mutants do not have any defects in dauer formation (Fig. 1; data not shown). Since *ins-8* expression increases upon perturbation of *ins-7*, we asked whether *ins-8* functions redundantly with *ins-7* in modulating dauer formation. The genomic proximity of the two genes precludes the generation of double mutants. However, neither RNAi of *ins-7* in an *ins-8* mutant, or vice versa, either in wild-type or in *rrf-3(pk1426)* mutants conferred a dauer formation phenotype or any other phenotypes including lethality, L1 arrest, or developmental delay compared with vector control RNAi or wild-type animals (data not shown).

Discussion

In this study, we report widespread and dynamic spatiotemporal expression of *C. elegans* insulins during development and under different physiological and environmental conditions. Not all expression patterns may completely capture endogenous insulin expression for several reasons. First, the promoter fragments used may lack regulatory elements required for expression in some tissues or to repress expression in others. Second, expression was annotated to the resolution of individual cells in some but not all tissues. Annotating at the level of individual cells is challenging, especially in neurons where insulins are abundantly expressed. Thus, the overlap in expression for neuronal insulins may be lower than estimated here. Third, we did not observe GFP expression in the germline, most likely because of transgene silencing (Kelly et al. 1997). Finally, insulins are secreted peptides that may act cell-non-autonomously (Billing et al. 2011; Bai et al. 2012). Nonetheless, we provide both a first comprehensive data set as well as a set of transgenic strains, both of which enable future studies of the functional dynamics of the entire *C. elegans* insulin gene family in vivo.

Several examples illustrate how the insulin expression patterns provide a resource for the derivation of functional hypotheses. First, the expression patterns of individual insulins can likely be used to predict specific functions. For instance, we found that *ins-3* is expressed in the distal tip cell, spermatheca, and the spermatheca-uterine valve. Previously, it has been shown that perturbations of *ins-3* result in defects in germline proliferation, and that this is due to the function of *ins-3* in the soma (Michaelson et al. 2010). Second, changes in expression of individual insulins can likely be used to infer specific functions as well. For example, it has been shown that *daf-28* is involved in dauer formation, and we find that its expression is turned off in dauers (Cornils et al. 2011; this study). Third, the expression data can likely be used to derive hypotheses about complex functional redundancies. For instance, 14 insulins change in expression as the animals age; eight of which increase and six of which decrease in expression. Future studies of which insulins are receptor agonists versus antagonists will be important to derive hypotheses regarding combinatorial insulin function in lifespan regulation.

Remarkably, pairwise combinations of insulin loss/knockdown did not result in defects in dauer formation. This could be due to inefficient insulin knockdown in neuronal tissues required

for dauer formation (Hu 2007). Alternatively, there may be more complex patterns of redundancies between more than two insulins. The lack of 1:1 insulin relationships is supported by the observation of two clusters in the expression network, rather than simple pairwise coexpression patterns.

Two insulins, *ins-7* and *ins-8*, share high-sequence similarity and expression overlap through development. We find that *ins-8* levels increase in the absence of *ins-7*, but did not detect compensatory changes in *ins-7* expression upon loss of *ins-8*. Knockdown of *ins-7* in an *ins-8* mutant or vice versa did not affect dauer formation. However, loss of *ins-7* slightly extends lifespan (Murphy et al. 2007; Supplemental Fig. S7). This suggests that, while *ins-8* increases upon loss of *ins-7*, it cannot fully compensate for the loss of *ins-7* in modulating longevity, further illustrating the lack of 1:1 relationships between insulins. Additionally, while loss of *ins-7* extends lifespan, it does not recapitulate the dramatic lifespan extension caused by a loss-of-function in *daf-2*. Together, this data suggests that *ins-7* and *ins-8* may genetically interact with other insulins. Alternatively, it is conceivable that the overexpression of *ins-8* in the *ins-7(tm1907)* rather than the loss of *ins-7* may be responsible for the observed lifespan extension of *ins-7* mutants, which would suggest that *ins-8* is an antagonist, whereas *ins-7* is an agonist. However, dauer formation was not observed upon loss of *ins-7* nor were *ins-8(tm4144)* mutants dauer defective (Fig. 1). This illustrates that *ins-8* is not likely an antagonist.

In the wild, *C. elegans* is exposed to a variety of environmental conditions. Food and temperature serve as modulators of dauer formation (Riddle et al. 1997). By entering dauer, development ceases until a favorable environment becomes available. Intuitively, differential insulin activity could ensure that a whole-organism response, such as dauer formation, only occurs under specific conditions and not in response to any type of stress. Remarkably, we found that insulins that share high expression overlap during development are distinct from those that share similarities under different conditions (Fig. 7A). This may reflect a need for specificity in coordinating development versus responding to different stresses. The expansion of the insulin family and its genetic wiring may provide the animal with a complex and highly organized repertoire of responses that can be tailored according to developmental or environmental need. In biological systems, different components can be interchangeable under certain conditions, but perform distinct functions under other conditions. Such redundancy, or degeneracy, contributes to the robustness, complexity, and evolvability of a system (Edelman and Gally 2001; Kitano 2004). We propose that the *C. elegans* insulin family has evolved to attain a type of "block design," a term used in combinatorial mathematics. An evolved block design can lead to the formation of different functional alliances in different cells for different developmental versus environmental conditions (Fig. 7B). Cooperative action between insulins may provide a robust framework for interpreting physiological and environmental cues and ensuring that the DAF-2 receptor initiates the appropriate response. For instance, *daf-28* and *ins-6* cooperate in modulating dauer formation. However, each insulin also functions specifically in dauer entry or exit, respectively (Cornils et al. 2011).

C. elegans likely utilize 40 insulins to achieve both specificity and redundancy to maintain fitness in a complex environment. The total possible number of combinations in different sets, for instance with three, four, or even more, is immense, and potentially provides the animal with a large and flexible repertoire of possibilities to recognize and respond appropriately to many environmental stimuli. Future studies will illuminate the generality of this observation for insulins in other nematodes, as well as
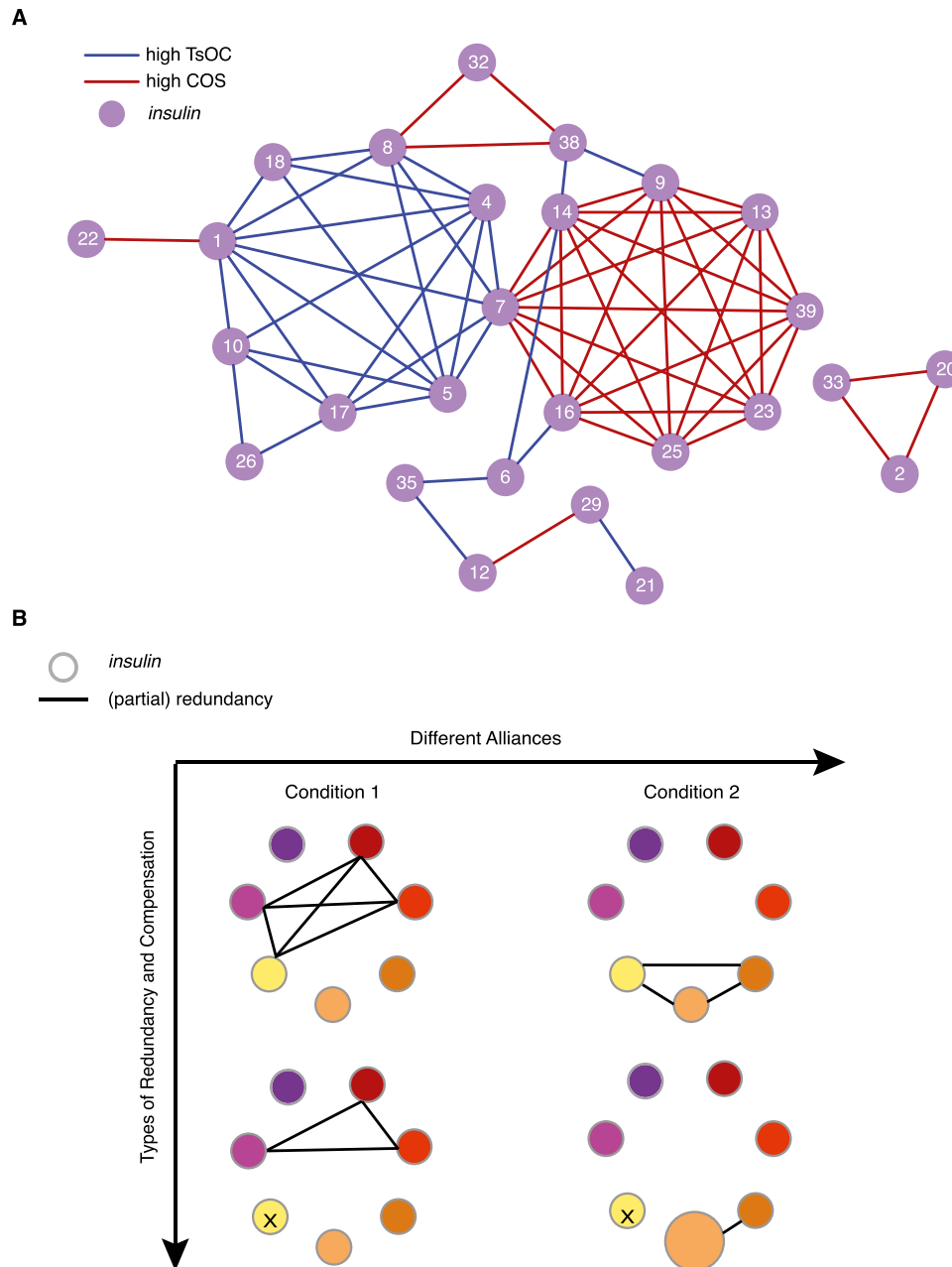
Figure 7. Model for complexity and robustness in gene families. (*A*) A coexpression network of the insulins (circles) that share the top 5% average TsOC through six larval/adult stages (blue edge) and the top 5% of conditional overlap (COS; red edge). Distinct clusters are observed whereby no insulins are both connected through high TsOC and high COS. (*B*) A model of "block design" or "alliances" in insulins under a given condition. A colored node represents one insulin. An edge represents overlapping function. Different alliances allow for the appropriate response to different conditions (*top* row). Although different conditions likely require different insulins, some (yellow node) may always be shared between conditions. Loss of single insulin activity does not cause a phenotypic affect due to redundancy (*bottom, left*) and/or compensation (*bottom, right*) with other insulins.

other gene families in *C. elegans* and other complex multicellular organisms such as humans.

Methods

Generation of *Pins::GFP* transgenic animals

We used gene predictions available in WormBase WS150 (http://ws150.wormbase.org). In total, 22 of the 40 promoter constructs were obtained from the Promoterome collection (Dupuy et al.

2004). The remaining promoters were PCR amplified and cloned as described (Dupuy et al. 2004). Primer sequences are provided in Supplemental Table S6. Transgenic animals were generated by microparticle bombardment as described (Reece-Hoyes et al. 2007). For each *Pins::GFP* reporter construct, two to six independent lines were obtained. At least two lines were examined for GFP expression, and the line that showed the highest transmission was used further. Nineteen of 40 *Pins::GFP* strains harbored the GFP construct integrated in the genome. For all nonintegrated lines, 50–100 animals were analyzed to capture each component of the

expression pattern. An increase in expression was recorded if the level of GFP dramatically increased and/or if expression in additional tissues was observed, while an expression decrease was defined by a dramatic reduction in GFP levels and/or a complete loss of expression in at least one cell/tissue type. All strains are available from the *C. elegans* Genetics Center, or upon request.

Pins::GFP expression pattern annotation

GFP expression was examined in 40–50 animals by fluorescence microscopy using a Zeiss Axioscope 2 *plus* microscope as described previously (Reece-Hoyes et al. 2007; Martinez et al. 2008; Grove et al. 2009). Temporal expression patterns were classified into six stages: four larval stages (L1–L4), young adult (no eggs), and ~3- to 4-d-old adult with eggs. Developmental stages were defined by the length and number of cells in the gonad (wormatlas.org). Spatial expression patterns were classified into 19 categories that correspond to tissues, cell types, organs, and individual cells. Temporal and spatial expression was standardized into a binary code, where "1" represents expression detected and "0" indicates no expression detected.

Tissue overlap analysis

Tissue overlap coefficient (TsOC) analysis was performed as described (Vermeirssen et al. 2007; Martinez et al. 2008; Grove et al. 2009), with minor modifications. Nodes were defined as insulins and the cell/tissue in which they are expressed. TsOC was analyzed for each pair of nodes by a geometric formula (Goldberg and Roth 2003).

Conditional expression analysis

Each stress was determined sufficient by testing with genes known to respond to the respective condition (see Supplemental Fig. S5). Oxidative stress: 1 mL of 200 mM paraquat was added to 10 mL of OP50-seeded NGM plates. Young adults were added and incubated at 20°C for 6 h. Verified using *Pgst-4::GFP* transgenics (Tawe et al. 1998). Glucose: 1 mL of 2% glucose was added to 10 mL of OP50-seeded NGM plates. Young adults were added and incubated at 20°C for 1, 6, 12, and 24 h. Verified using *daf-16a::GFP* transgenics (Lee et al. 2009). Heat stress: L4 larvae and young adults were placed on OP50-seeded NGM plates and incubated at 30°C for 6 h. Plates were not stacked to assure that heat was equally distributed among all plates. Verified using *Phsp-4::GFP* transgenics (Heschl and Baillie 1990). Starvation: L4 and young adult animals were washed twice using 1X M9 buffer then placed on peptone-free plates without bacteria. Animals were incubated at 20°C for 24 h. Verified using *Pacdh-1::GFP* transgenics (MacNeil et al. 2013). Dauer: We obtained dauers by allowing plates to starve. Strains were grown at 20°C until no food remained. Dauers were picked 3–4 d after plates were starved. Aging: 1-d-old adults were placed on OP50-seeded 5-fluorodeoxyuridine (FuDR) plates. GFP expression was recorded on days 5, 10, and 15. In all experiments, the presence of GFP expression was translated into numbers where "1" represents a GFP increase, "0" represents no change in GFP, and "−1" represents a GFP decrease. Correlations (i.e., conditional overlap) were determined using Spearman's rank test.

Methods, including strain maintenance, genotyping, RNA interference, dauer analysis, and quantitative RT–PCR can be found in the Supplemental Material.

Acknowledgments

References

Bai H, Kang P, Tatar M. 2012. *Drosophila* insulin-like peptide-6 (dilp6) expression from fat body extends lifespan and represses secretion of *Drosophila* insulin-like peptide-2 from the brain. *Aging Cell* **11:** 978–985.

Baugh LR, Kurhanewicz N, Sternberg PW. 2011. Sensitive and precise quantification of insulin-like mRNA expression in *Caenorhabditis elegans*. *PLoS ONE* **6:** e18086.

Billing O, Kao G, Naredi P. 2011. Mitochondrial function is required for secretion of DAF-28/insulin in *C. elegans*. *PLoS ONE* **6:** e14507.

Broughton S, Alic N, Slack C, Bass T, Ikeya T, Vinti G, Tommasi AM, Driege Y, Hafen E, Partridge L. 2008. Reduction of DILP2 in *Drosophila* triages a metabolic phenotype from lifespan revealing redundancy and compensation among DILPs. *PLoS ONE* **3:** e3721.

Burga A, Casanueva MO, Lehner B. 2011. Predicting mutation outcome from early stochastic variation in genetic interaction partners. *Nature* **480:** 250–253.

Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC. 1994. Green fluorescent protein as a marker for gene expression. *Science* **263:** 802–805.

Claeys I, Simonet G, Poels J, Van Loy T, Vercammen L, De Loof A, Vanden Broeck J. 2002. Insulin-related peptides and their conserved signal transduction pathway. *Peptides* **23:** 807–816.

Colombani J, Andersen DS, Leopold P. 2012. Secreted peptide Dilp8 coordinates *Drosophila* tissue growth with developmental timing. *Science* **336:** 582–585.

Cornils A, Gloeck M, Chen Z, Zhang Y, Alcedo J. 2011. Specific insulin-like peptides encode sensory information to regulate distinct developmental processes. *Development* **138:** 1183–1193.

Dupuy D, Li QR, Deplancke B, Boxem M, Hao T, Lamesch P, Sequerra R, Bosak S, Doucette-Stamm L, Hope IA, et al. 2004. A first version of the *Caenorhabditis elegans* Promoterome. *Genome Res* **14:** 2169–2175.

Duret L, Guex N, Peitsch MC, Bairoch A. 1998. New insulin-like proteins with atypical disulfide bond pattern characterized in *Caenorhabditis elegans* by comparative sequence analysis and homology modeling. *Genome Res* **8:** 348–353.

Edelman GM, Gally JA. 2001. Degeneracy and complexity in biological systems. *Proc Natl Acad Sci* **98:** 13763–13768.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151:** 1531–1545.

Gems D, Sutton AJ, Sundermeyer ML, Albert PS, King KV, Edgley ML, Larsen PL, Riddle DL. 1998. Two pleiotropic classes of *daf-2* mutation affect larval arrest, adult behavior, reproduction and longevity in *Caenorhabditis elegans*. *Genetics* **150:** 129–155.

Goldberg DS, Roth FP. 2003. Assessing experimentally derived interactions in a small world. *Proc Natl Acad Sci* **100:** 4372–4376.

Gronke S, Clarke DF, Broughton S, Andrews TD, Partridge L. 2010. Molecular evolution and functional characterization of *Drosophila* insulin-like peptides. *PLoS Genet* **6:** e1000857.

Grove CA, De Masi F, Barrasa MI, Newburger DE, Alkema MJ, Bulyk ML, Walhout AJ. 2009. A multiparameter network reveals extensive divergence between *C. elegans* bHLH transcription factors. *Cell* **138:** 314–327.

Herndon LA, Schmeissner PJ, Dudaronek JM, Brown PA, Listner KM, Sakano Y, Paupard MC, Hall DH, Driscoll M. 2002. Stochastic and genetic factors influence tissue-specific decline in ageing *C. elegans*. *Nature* **419:** 808–814.

Heschl MF, Baillie DL. 1990. The HSP70 multigene family of *Caenorhabditis elegans*. *Comp Biochem Physiol B* **96:** 633–637.

Hollenhorst PC, Shah AA, Hopkins C, Graves BJ. 2007. Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. *Genes Dev* **21:** 1882–1894.

Hu PJ. 2007. Dauer. In *Wormbook* (ed. The *C. elegans* Research Community). doi: 10.1895/wormbook.1.144.1. http://www.wormbook.org.

Kelly WG, Xu S, Montgomery MK, Fire A. 1997. Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146:** 227–238.

Kenyon C. 2011. The first long-lived mutants: Discovery of the insulin/IGF-1 pathway for ageing. *Philos Trans R Soc Lond B Biol Sci* **366:** 9–16.

Kitano H. 2004. Biological robustness. *Nat Rev Genet* **5:** 826–837.

Lee SJ, Murphy CT, Kenyon C. 2009. Glucose shortens the life span of *C. elegans* by downregulating DAF-16/FOXO activity and aquaporin gene expression. *Cell Metab* **10:** 379–391.

Li C, Kim K. 2008. Neuropeptides. In *Wormbook* (ed. The *C. elegans* Research Community). doi: 10.1895/wormbook.1.142.1. http://www.wormbook.org.

Li W, Kennedy SG, Ruvkun G. 2003. *daf-28* encodes a *C. elegans* insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway. *Genes Dev* **17:** 844–858.

MacNeil LT, Walhout AJ. 2011. Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. *Genome Res* **21:** 645–657.

MacNeil LT, Watson E, Arda EH, Zhu LJ, Walhout AJM. 2013. Diet-induced developmental acceleration independent of TOR and insulin in *C. elegans*. *Cell* **153:** 253–266.

Marsh EK, May RC. 2012. *Caenorhabditis elegans*, a model organism for investigating immunity. *Appl Environ Microbiol* **78:** 2075–2081.

Martinez NJ, Ow MC, Reece-Hoyes JS, Barrasa MI, Ambros VR, Walhout AJ. 2008. Genome-scale spatiotemporal analysis of *Caenorhabditis elegans* microRNA promoter activity. *Genome Res* **18:** 2005–2015.

Michaelson D, Korta DZ, Capua Y, Hubbard EJ. 2010. Insulin signaling promotes germline proliferation in *C. elegans*. *Development* **137:** 671–680.

Murphy CT, McCarroll SA, Bargmann CI, Fraser A, Kamath RS, Ahringer J, Li H, Kenyon C. 2003. Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* **424:** 277–283.

Murphy CT, Lee SJ, Kenyon C. 2007. Tissue entrainment by feedback regulation of insulin gene expression in the endoderm of *Caenorhabditis elegans*. *Proc Natl Acad Sci* **104:** 19046–19050.

Ohno S. 1970. *Evolution by gene duplication*. Springer-Verlag, New York.

Pierce SB, Costa M, Wisotzkey R, Devadhar S, Homburger SA, Buchman AR, Ferguson KC, Heller J, Platt DM, Pasquinelli AA, et al. 2001. Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev* **15:** 672–686.

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

Raj A, Rifkin SA, Andersen E, van Oudenaarden A. 2010. Variability in gene expression underlies incomplete penetrance. *Nature* **463:** 913–918.

Reece-Hoyes JS, Shingles J, Dupuy D, Grove CA, Walhout AJ, Vidal M, Hope IA. 2007. Insight into transcription factor gene duplication from *Caenorhabditis elegans* Promoterome-driven expression patterns. *BMC Genomics* **8:** 27.

Riddle DL, Blumenthal T, Meyer BJ, Priess JR. 1997. *C. elegans II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* **13:** 2498–2504.

Simmer F, Tijsterman M, Parrish S, Koushika SP, Nonet ML, Fire A, Ahringer J, Plasterk RH. 2002. Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. *Current Biol* **12:** 1317–1319.

Sommer RJ, Streit A. 2011. Comparative genetics and genomics of nematodes: Genome structure, development, and lifestyle. *Annu Rev Genet* **45:** 1–20.

Tawe WN, Eschbach ML, Walter RD, Henkle-Duhrsen K. 1998. Identification of stress-responsive genes in *Caenorhabditis elegans* using RT-PCR differential display. *Nucleic Acids Res* **26:** 1621–1627.

Thoemke K, Yi W, Ross JM, Kim S, Reinke V, Zarkower D. 2005. Genome-wide analysis of sex-enriched gene expression during *C. elegans* larval development. *Dev Biol* **284:** 500–508.

Tissenbaum HA, Ruvkun G. 1998. An insulin-like signaling pathway affects both longevity and reproduction in *Caenorhabditis elegans*. *Genetics* **148:** 703–717.

Tong YG, Bürglin TR. 2010. Conditions for dye-filling of sensory neurons in *Caenorhabditis elegans*. *J Neurosci Methods* **188:** 58–61.

Vermeirssen V, Barrasa MI, Hidalgo CA, Babon JA, Sequerra R, Doucette-Stamm L, Barabasi AL, Walhout AJ. 2007. Transcription factor modularity in a gene-centered *C. elegans* core neuronal protein-DNA interaction network. *Genome Res* **17:** 1061–1071.

Wolkow CA. 2002. Life span: Getting the signal from the nervous system. *Trends Neurosci* **25:** 212–216.

Complex expression dynamics and robustness in *C. elegans* insulin networks

Ashlyn D. Ritter, Yuan Shen, Juan Fuxman Bass, et al.

Supplemental Material	http://genome.cshlp.org/content/suppl/2013/04/16/gr.150466.112.DC1.html
References	This article cites 45 articles, 24 of which can be accessed free at: http://genome.cshlp.org/content/23/6/954.full.html#ref-list-1
Creative Commons License	This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at http://creativecommons.org/licenses/by-nc/3.0/.
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to *Genome Research* go to:
http://genome.cshlp.org/subscriptions