Dauer fate in a *Caenorhabditis elegans* Boolean network model

Alekhya Abhiram Kandoor[1] and Janna L. Fierst[2*]

[1] Department of Biomedical Engineering, University of Virginia, Charlottesville, VA, USA

[2] Biomolecular Sciences Institute and Department of Biological Sciences, Florida International University, Miami, FL, USA

* Author for correspondence: jfierst@fiu.edu

Abstract

Cellular fates are determined by genes interacting across large, complex biological networks. A critical question is how to identify causal relationships spanning distinct signaling pathways and underlying organismal phenotypes. Here, we address this question by constructing a Boolean model of a well-studied developmental network and analyzing information flows through the system. Depending on environmental signals *Caenorhabditis elegans* develop normally to sexual maturity or enter a reproductively delayed, developmentally quiescent 'dauer' state, progressing to maturity when the environment changes. The developmental network that starts with environmental signal and ends in the dauer/no dauer fate involves genes across 4 signaling pathways including cyclic GMP, Insulin/IGF-1, TGF-β and steroid hormone synthesis. We identified three stable motifs leading to normal development, each composed of genes interacting across the Insulin/IGF-1, TGF-Beta and steroid hormone synthesis pathways. Three genes known to influence dauer fate, *daf-2*, *daf-7* and *hsf-1*, acted as driver nodes in the system. Using causal logic analysis, we identified a five gene cyclic subgraph integrating the information flow from environmental signal to dauer fate. Perturbation analysis showed that a multifactorial insulin profile determined the stable motifs the system entered and interacted with *daf-12* as the switchpoint driving the dauer/no dauer fate. Our results show that complex organismal systems can be distilled into abstract representations that permit full characterization of the causal relationships driving developmental fates. Analyzing organismal systems from this perspective of logic and function has important implications for studies examining the evolution and conservation of signaling pathways.

Introduction

Cellular states, developmental pathways and organismal phenotypes are determined by interactions between biomolecules. These biological networks are large and complex, involving genes, proteins and regulatory molecules spanning diverse pathways and levels of biological organization [1]. Data-based network inference often implicates large numbers of biomolecules through correlation-based methods [2] but network modeling and analysis indicate causal relationships are the key to controlling information flow and determining cellular fates [3]. A fundamental question is how to identify the critical, causal relationships within complex biological systems and between distant nodes in a network [4]. Here, we construct a Boolean model of a well-studied model system and use network analysis techniques to identify the core interactions responsible for the phenotype.

C. elegans is a small (~1mm), free-living nematode worm that has been used as a model organism in biological research for over 50 years [5]. Populations are androdioecious, composed of self-fertile hermaphrodites and a small number of males, resulting in straightforward laboratory culture and rapid propagation. *C. elegans* are transparent, facilitating microscope study, and eutelic, meaning that all adults have the same number of somatic cells. In *C. elegans* hermaphrodites there are 959 somatic cells while males have 1031 [6]. The entire life cycle can be completed within 3 days and these factors have led to the establishment of *C. elegans* as an important model organism for studying development, genetics and essential biological processes. *C. elegans* was the first metazoan to have its genome fully sequenced [7, 8] and the long history of *C. elegans* research has provided a wealth of genetic, genomic and developmental understanding of this small worm [9-12].

One of the most studied developmental processes in *C. elegans* is dauer [13, 14]. The small worms are vulnerable to environmental conditions including lack of food or moisture [15] and high temperature [16]. *C. elegans* senses its environment through chemoreceptors that perceive biomolecules including pheromones produced by other worms [17]. In low-stress environments the worms progress through discrete larval stages L1-L4, to adulthood and reproductive maturity [13]. Environmental triggers can activate a developmental switch at the L1 stage and lead to dauer arrest [18]. Dauer worms can move but are developmentally quiescent and survive without feeding for periods of up to 4 months.

C. elegans can be forced into the dauer fate with high pheromone concentrations, high temperature or starvation in the lab but the signals can also interact quantitatively in the worm's natural environment [17]. The dauer pheromone itself is a complex mix of ascarosides [19, 20] with natural variation in

production, composition and response between populations [21], and species [22]. *C. elegans* strains have substantial, quantitative natural variation for dauer entry [23], [24] and exit [25] but for the individual worm dauer is an irreversible developmental decision. Undergoing dauer has lasting effects on worm reproduction [26], physiology [27] and locomotion [28].

Dauer is a unique, discrete developmental switch and the molecular basis has been studied extensively. Over 100 genes across 4 signaling pathways (cyclic GMP, Insulin/IGF-1, TGF-β and steroid hormone synthesis) have been implicated as influencing dauer [13, 29]. Despite extensive descriptive characterization, we lack a robust understanding of the interplay between these pathways and the causal links that drive the system to one fate or another. Here, we study a Boolean model and identify the critical interactions determining dauer fate.

Simplified representations of complex biological networks can accurately distill systems down to essential interactions [30]. Boolean networks are graphs where nodes represent genes and edges are regulatory interactions defined as Boolean functions or logic gates, for example the operators 'AND' or 'OR' [31]. Boolean models of biological processes represent states as discrete 'on' or 'off' [32] and accurately represent the switch-like nature of dauer induction. Regulatory interactions are threshold-gated, qualitative relationships. Here, we use our Boolean network to emulate cellular behavior observed in experimental settings and systematically perturb the system to discover causal links within the larger complex network. We identify the critical interactions and combinations of gene states that determine dauer fate. Evolutionary searches of pathway conservation and divergence typically examine gene-by-gene, sequence level homology [33]. Here, we suggest that causal logic and function may be the true phenotypes underlying the evolution of signaling pathways.

Methods

Constructing the dauer network

We first extracted 85 dauer-involved genes described in the WormBook 'Dauer' chapter [13]. For each of these genes, we then searched on WormBase [53] for genes known to interact with these dauer genes and genes known to act in the same pathways as these dauer genes. We focused on regulatory interactions and included those for which we identified information regarding the direction of the interaction, for example reported relationships between effector proteins and targets. We did not include genetic interactions or genetic masking/enhancing/suppressing effects between genes. Some of

our gene pairs may have reported genetic interactions along with regulatory interactions but we focused on regulatory interactions for constructing our model.

We next surveyed the literature for regulatory interactions within this gene and pathway set and created a filtered set of dauer genes and main-effect regulatory interactions (Supplementary Table 1). For example, of the 40 insulin signaling genes in the *C. elegans* genome few have been reported to act independently in the Insulin/IGF-1 pathway [5]. We condensed the larger dauer-involved set down to 26 genes, each part of the four dauer-inducing pathways represented as nodes in our Boolean model, and connected by 34 edges. We included the nodes 'pheromone' [34] and '*cmk-1*' (a food receptor [35]) as inputs to the system. These interactions were mainly obtained from the WormBase database [53] and verified with the scientific publications cited on WormBase (a full list of nodes, interactions, and citations is given in Supplementary Table 1).

We used this regulatory interaction information for each gene to formulate Boolean functions (Supplementary Table 2). We combined literature-reported regulations into a preliminary model and used results obtained from simulations and perturbations to fine-tune the Boolean functions for each node until we were able to mimic the expected behavior of the system. For example, multiple articles reported that *daf-16* was activated either when *akt* was repressed [54-57] or when *daf-12* was activated [58, 59]. Simulated conditions for normal adult development should therefore result in *daf-16* repression and no dauer activation. To achieve this known output for *daf-16* we assigned an 'OR' operator for these two regulators. In some cases, we were able to extract the relationship between regulatory nodes from the literature. This approach has been similarly demonstrated in previous works to gain qualitative insights using discrete networks [60, 61].

Studying system trajectories

The initial conditions used to generate the trajectories of the system's dynamics were based on well-known environmental cues such as food availability (*cmk-1*) and dauer pheromone (pher) detected by *C. elegans* [17, 35]. Favorable conditions allowed the organism to develop into an adult ("no dauer") and unfavorable conditions into "dauer" [34]. We set the initial conditions as *cmk-1* = 1 and pher = 0 for normal development (here, dauer = 0) and *cmk-1* = 0 and pher = 1 for dauer fate (here, dauer = 1). We monitored the dynamics and convergence of the system based on the state of the 'dauer' output node (trajectories for each simulation are shown in Supplementary Figure 2). Temperature is another well-

known regulator that induces dauer formation. We chose to focus here on dauer as a response to resource availability and our model does not include temperature as an input.

We performed simulations with asynchronous updating using boolean2 software [62]. In order to explore system dynamics, the Boolean functions for each gene needs to be updated at each timestep to trace out the trajectory of the system. These transitions for the Boolean functions were carried out using an asynchronous scheme that randomly assigned the states of the nodes. Biological events do not occur at the same time and the timescale at which the events occur can be diverse. Here, we represented this with an asynchronous update scenario. In this framework a collector class averaged the states of nodes over multiple runs for simulating a system with a non-deterministic updating scheme. We performed 100 simulations for each set of initial conditions and for every iteration, the collector averaged the state of each node generated for 1000 steps. In all our simulations, the nodes reached equilibrium after approximately 15 iterations (trajectories are shown in Supplementary Figure 2). Nodes with states > 0.5 were considered 'ON' (as indicated by [62]) and nodes with states <0.5 were considered 'OFF'.

Validating the model

To validate the model, we checked whether the model has the ability to reproduce known experimental outcomes. Mutant phenotypes in *C. elegans* that lead to either adult or dauer formation are well studied [13]. We used published experimental evidence to check whether our model could reproduce similar results. For example, dauer-constitutive (Daf-c) mutants lead to dauer formation in conditions that do not favor dauer while Dauer-defective (Daf-d) mutants will not develop into dauer larvae [18, 40-42]. We knocked down known Daf-c and Daf-d genes to validate that our dauer Boolean model captured essential wild type and mutant phenotypes (Supplementary Figure 1b).

Our model simulations helped us identify the attractor space and stable motifs. A Boolean model will eventually settle down into a static state where all the nodes are either 'on' or 'off' and remain in that state through further iterations. This state of the model is called an attractor and in biological modeling attractors depict specific biological phenotypes. Within this, a set of nodes within the network forms cycles which stay in specific states, and these are called stable motifs. We can represent a network as a reduced system using the nodes that are critical for controlling the stable motifs. We can then study the

model dynamics using the reduced form and can make predictions using simpler representations of the model.

Attractors of Boolean networks represent biological phenotypes, here the 'dauer' and 'no dauer' fates. We explored the set of attractors and stable motifs for our model by iterating through all possible combinations of initial conditions and perturbations using pystablemotifs software [36, 37]. A perturbation of a node in our experiments refers to reversing the existing state of a node and specify that along with the initial conditions set. The network diagrams were created using Cytoscape [63]. Code, simulations and scripts are available at https://github.com/alekhyaa2/Dauer-Boolean-Model.

Results

Constructing a dauer Boolean network

Our exhaustive literature survey identified essential regulatory interactions among dauer pathway genes. From these we constructed a Boolean network with 26 nodes and 34 interactions (Figure 1a; Supplementary Table 1). The nodes 'pheromone' [34] and '*cmk-1*' (a food receptor [35]) were input signals that allowed the model to converge onto the dauer/no dauer output node.

Validating the model by replicating Daf-c and Daf-d phenotypes

In *C. elegans* mutant strains can be categorized as dauer-constitutive (Daf-c), genotypes which produce dauer arrest even in the absence of pheromone and presence of food, and Dauer-defective (Daf-d), genotypes which do not enter dauer arrest under typical dauer-favorable conditions (i.e., pheromone present and/or food limited) [18, 40-42]. We validated our model by knocking down known Daf-c genes (*daf-2*, *daf-7*, *daf-11*, and *daf-9*) separately under initial conditions (*cmk-1* = 1 and pher = 0) that do not typically favor dauer and found that these mutants followed the Daf-c phenotype and entered dauer arrest. We also knocked down known Daf-d mutant genes (*daf-12* and *daf-16*) under conditions that typically favor dauer (here, *cmk-1* = 0 and pher = 1) and found these replicated the Daf-d phenotype, failing to enter dauer arrest.

Identifying stable motifs

A network attractor is a set of gene states the system can stabilize in [36]. The dauer Boolean network had 10 different attractor states (Supplementary Figure 1) with 1 resulting in developmental progression to adult, 8 resulting in the dauer fate and 1 resulting in either dauer or no dauer depending on system

conditions [37]. Although an attractor represents a potential stable point, not all attractors are accessible from all states. To further analyze system stability, we divided the total network into small subgraphs [3, 38]. We identified three Stable Motifs (SMs) in the developmental progression to adult (Figure 1b-d). Each of these encompassed genes from at least two different pathways. Stable Motif 1 (SM1; Figure 1b) spanned 10 genes across the IGF-1/insulin signaling, TGF-β and Steroid hormone pathways [13]. SM2 and SM3 involved 4 and 8 genes, respectively, across the IGF-1/insulin signaling and TGF-β pathways. *daf-2*, *daf-7* and *hsf-1* (heat shock protein) were common between all 3 stable motifs with *hsf-1* acting as a crucial node where different pathways intersect. This result has been shown by previous experimental work [39]. Each of the SMs involved a different insulin signaling gene, *ins-1*, *ins-7* or *ins-18*. We did not identify any SMs for dauer formation (the A3 attractor; Supplementary Figure 1).

Identifying driver nodes in the SMs

We simulated gene perturbations to identify the driver nodes for each of our identified SMs [37]. Changing the existing state of the critical nodes shared between the stable motifs (i.e., *daf-2* or *daf-7* downregulation or *hsf-1* upregulation) prevented the system from entering into any of the SMs and pushed it to one of the dauer attractors. Perturbation analysis identified the critical causal interactions in the network as SM2, encompassing *daf-2*, *daf-7*, *ins-7* and *hsf-1*. Perturbation of *daf-2*, *daf-7* or *hsf-1* genes prohibited the system from stabilizing in the SM and the attractors corresponded to each of these nodes. We concluded that *daf-2*, *daf-7* and *hsf-1* were the driver nodes for the SMs. When we perturbed the other nodes of these three stable motifs, the system converged onto either dauer or no dauer via different stable motifs. (Supplementary Table 3). But in case of *daf-2*, *daf-7* and *hsf-1* perturbations, the system did not enter into any of the stable motifs, instead settling into the attractors of those corresponding nodes.

Perturbations of *daf-2*, *daf-7* and *hsf-1* resembled the Daf-c mutant system. When the model was started with dauer-favorable conditions, the system did not have any stable motifs. The same lack of stable motifs occurred with Daf-c mutant phenotypes. Under both dauer-favorable and Daf-c mutant conditions (here, related gene perturbation and/or favorable initial conditions), the system did not have any cyclic subgraphs facilitating control. Instead, the driver nodes decided the fate of the organism in expressing a wildtype or a Daf-c mutant phenotype.

Analyzing causal interactions in the dauer Boolean network

We identified a reduced network representing causal information flow [43] from environmental input to dauer fate (Figure 2a). The reduced network contained a cyclic subgraph (Figure 2b) similar to SM2, containing *daf-2*, *daf-7*, *ins-7* and *hsf-1*. However, it included a necessary inhibitory regulation of *daf-9* by *daf-2*. The dynamics of the entire system can be summarized as input to this subgraph. The full reduced network (Figure 2c) included 12 nodes and 13 edges.

Insulin signaling drives the binary dauer/no dauer fate

Gene perturbations revealed an additional 7 SMs (SM4-10; Supplementary Table 3). For each of the perturbations performed, the system could end up in one of 10 different attractors but few of these could be reached from stable conditions with active stable motifs. Activating 'dauer' conditions including low food and pheromone presence resulted in dauer fate and the A3 attractor (Figure 3a) and available food with pheromone absence resulted in the system reaching the A7 attractor through SMs 1, 2 or 3 (Figure 3b).

We identified insulin-dependent dauer arrest attractors across the perturbations. Activation of *ins-1* and *ins-18* resulted in the dauer fate through SMs 4-10 but inactivation of either pushed the system to normal development (Figure 3c). Together with *ins-7* the activation/inactivation profile of these genes pushed the system to normal development through SMs 1, 2 or 3 (Figure 1b-d). Similarly, *daf-12* activation resulted in dauer arrest with the insulin profile determining the SM the system entered. When the antagonistic insulins (*ins-1* and *ins-18*) were activated the system stabilized into SM 8 and when *ins-1* was turned off the system stabilized in either SM 2 or 3 (Figure 3d).

We found that the system must travel through SMs 1, 2 or 3 to end up in normal developmental progression and avoid the dauer fate. To enter dauer arrest the system must travel through one of SMs 4-10 with *ins-1* and *ins-18* activated and *ins-7* inactivated. In this model insulin signaling controls the transitions between normal development and dauer in the presence of gene perturbations. The insulin genes act as decision makers for the convergence of these dauer pathways.

Identifying binary states across dauer pathways

For normal development we identified one strongly correlated module, one strongly anticorrelated module and a third set of genes with quantitative relationships across the SMs (Figure 4). For dauer arrest we identified two strongly correlated modules and two strongly anticorrelated modules. These

modules span the 4 signaling pathways involved in dauer arrest and indicate co-regulation and correlative relationships between distant genes in the network.

Discussion

We have constructed a Boolean model of the dauer arrest network and identified the critical, causal interactions driving the system between dauer fate and normal development. Although the network contains 26 genes and 34 interactions, the system can be driven between the dauer/no dauer fate by controlling a few critical nodes. Importantly, the switchpoints for this system require specific combinations of gene states behaving synergistically or antagonistically. The results of our model demonstrate that complex systems can be controlled through a few critical nodes and connections, echoing both control theory [3, 44] and long-standing theoretical work on biological systems [45]. Evolutionary studies of signaling pathway divergence and conservation have focused on individual genes and sequence-level changes [33]. Our results demonstrate that diffuse, correlative interactions have no impact on system stability or developmental fate. We suggest that system-level function and dynamics are, instead, the true evolutionary targets of selection.

Many of the results of our study mirror experimental dauer investigations. For example, a systematic analysis of mutants across insulin signaling genes in *C. elegans* identified roles in lifespan, reproduction, pathogen resistance, thermotolerance and dauer entry and exit [46]. Here, by constructing a network that encompassed multiple signaling pathways we were able to analyze the role of these insulin signaling genes and robustly identify the interaction between insulin signaling, heat shock response and daf-sensitive genes that determines dauer fate. Our results demonstrate that dauer signaling pathways find different routes to transduce the signals from the environment to developmental decision.

Over 100 genes have been described as influencing the dauer fate in *C. elegans* [13, 29]. However, many of these have been identified through mutant analyses or total loss of function alleles [47]. Genes may be capable of influencing the dauer state through loss-of-function or extreme phenotypes but this may not translate to a gene's involvment in dauer fate in natural populations. We focused on the genes and interactions determined essential to the dauer network. Distilling these further into causal interactions revealed a core set of genes and proteins controlling dauer fate. Transcriptomic studies have identified large, co-regulated, correlated modules associated with dauer mutants, entry and exit [46]. Our results show that these co-regulated modules also occur in our model (Figure 4) but their expression and

function are correlated consequences of developmental decision-making occurring at a few critical points in the system.

C. elegans is a microscopic soil-dwelling organism and extremely dependent on environmental conditions. Some type of developmental arrest or metabolic stasis occurs in many small organisms with similar environmental dependence. For example, tardigrades (Phylum Tardigrada) are microscopic invertebrates closely related to nematodes and can enter both ametabolic and anhydrous states where they extrude all water and remain in suspended animation for multiple years [48]. Despite the ubiquity of developmental and metabolic stasis, there is little understanding of the genetics and regulatory systems underlying these processes [49]. Evolutionary studies of developmental arrest across nematodes have used dauer genes and pathways as a tool to construct related models. For example, the parasitic nematode *Haemonchus contortus* transitions from both its L3 and L4 stages depending on cues from the host's gut environment [50]. A dauer model based on *C. elegans* identified 61 gene homologues with varying levels of conservation, each of which may be a potential target for life cycle disruption and anti-parasitic intervention [51]. Alternatively, many of these genes may not be involved in developmental arrest in *H. contortus*. They may have pleiotropic functions in *C. elegans* or have been co-opted for other organismal phenotypes.

Multiple studies support this scenario of conserved dauer phenotype and divergent gene and sequence evolution. A study of dauer-related genes and pathways within Phylum Nematoda found a high number of *Caenorhabditis* and *C. elegans*-specific dauer genes and an evolutionarily recent co-option of TGF-β signaling into the dauer pathway [29]. Although the insulin signaling pathway is broadly conserved, outside of *Caenorhabditis* the insulin genes examined showed little evolutionary conservation. In the nematode *Pristionchus pacificus* dauer pheromones produce dauer entry in individuals of other genotypes, suggesting the worms are using dauer signaling in intraspecific competition [22]. Natural variation in *P. pacificus* dauer is partially explained by copy number variation in *dauerless*, an 'orphan' gene that lacks homologs in related species [52]. These examples of near-scale divergence suggest that although dauer itself is a conserved developmental fate, the individual genes underlying dauer may not be essential targets of selection.

Our results suggest that system-level properties responsible for developmental fate may be the true evolutionary phenotypes under selection. The small number of genes and interactions underlying stable

motifs and the causal logic backbones may be subject to higher levels of evolutionary conservation, or their system 'role' may be conserved. Laboratory genetic manipulations within *C. elegans* and across nematodes, focusing on the key genes and interactions identified here, would be a tractable, exciting avenue for future study. For example, within *C. elegans* there are at least 40 insulin signaling homologs [5]. Divergence, loss and duplication of individual genes may not be as significant as the role the homologs together play in dauer signaling.

This work has important future applications. The model can be used to study how different cellular responses work in cohort with dauer pathways by extending the system to encompass genes and interactions involved in disparate pathways including the cell cycle and immune response. Expanding this framework can provide a tractable system to test experimental hypotheses that are currently challenging, labor-intensive and beyond the scope of traditional laboratory investigations. Viewing the conservation and divergence of signaling pathways through the lens of causality and system-level function can provide a robust framework for understanding the evolution of biological networks.

Acknowledgements

JLF and AAK were supported by National Science Foundation grants EF-1921562 and DEB1941854 and NIGMS award R35GM147245 to JLF.

References

1. Weng, G., U.S. Bhalla, and R. Iyengar, *Complexity in biological signaling systems.* Science, 1999. **284**(5411): p. 92-6.
2. Stolovitzky, G., D. Monroe, and A. Califano, *Dialogue on reverse-engineering assessment and methods: the DREAM of high-throughput pathway inference.* Ann N Y Acad Sci, 2007. **1115**: p. 1-22.
3. Zanudo, J.G. and R. Albert, *Cell fate reprogramming by control of intracellular network dynamics.* PLoS Comput Biol, 2015. **11**(4): p. e1004193.
4. Lawyer, G., *Understanding the influence of all nodes in a network.* Sci Rep, 2015. **5**: p. 8665.
5. Girard, L.R., et al., *WormBook: the online review of Caenorhabditis elegans biology.* Nucleic Acids Research, 2007. **35**: p. D472-D475.
6. Corsi, A., B. Wightman, and M. Chalfie, *A Transparent Window into Biology: A Primer on Caenorhabditis elegans.* Genetics, 2015. **200**(2): p. 387-407.
7. Consortium, C.e.S., *Genome sequence of the nematode C. elegans.* Science, 1998. **282**(5396): p. 2012-8.
8. Waterston, R.H., J.E. Sulston, and A.R. Coulson, *The Genome*, in *C. elegans II*, nd, et al., Editors. 1997: Cold Spring Harbor (NY).
9. Riddle, D.L., et al., *Introduction to C. elegans*, in *C. elegans II*, nd, et al., Editors. 1997: Cold Spring Harbor (NY).
10. Kemphues, K.J. and S. Strome, *Fertilization and Establishment of Polarity in the Embryo*, in *C. elegans II*, nd, et al., Editors. 1997: Cold Spring Harbor (NY).
11. Schedl, T., *Developmental Genetics of the Germ Line*, in *C. elegans II*, nd, et al., Editors. 1997: Cold Spring Harbor (NY).
12. Schnabel, R. and J.R. Priess, *Specification of Cell Fates in the Early Embryo*, in *C. elegans II*, nd, et al., Editors. 1997: Cold Spring Harbor (NY).
13. Hu, P.J., *Dauer*. WormBook.
14. Cassada, R.C. and R.L. Russell, *Dauerlarva, a Post-Embryonic Developmental Variant of Nematode Caenorhabditis-Elegans.* Developmental Biology, 1975. **46**(2): p. 326-342.
15. Golden, J.W. and D.L. Riddle, *The Caenorhabditis-Elegans Dauer Larva - Developmental Effects of Pheromone, Food, and Temperature.* Developmental Biology, 1984. **102**(2): p. 368-378.
16. Ailion, M. and J.H. Thomas, *Dauer formation induced by high temperatures in Caenorhabditis elegans.* Genetics, 2000. **156**(3): p. 1047-1067.
17. Jeong, P.Y., et al., *Chemical structure and biological activity of the Caenorhabditis elegans dauer-inducing pheromone.* Nature, 2005. **433**(7025): p. 541-5.
18. Vowels, J.J. and J.H. Thomas, *Genetic analysis of chemosensory control of dauer formation in Caenorhabditis elegans.* Genetics, 1992. **130**(1): p. 105-23.
19. Butcher, R., et al., *Small-molecule pheromones that control dauer development in Caenorhabditis elegans.* Nature Chemical Biology, 2007. **3**(7): p. 420-422.
20. Butcher, R., et al., *A potent dauer pheromone component in Caenorhabditis elegans that acts synergistically with other components.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(38): p. 14288-14292.
21. Viney, M. and S. Harvey, *Reimagining pheromone signalling in the model nematode Caenorhabditis elegans.* Plos Genetics, 2017. **13**(11).

22. Mayer, M. and R. Sommer, *Natural variation in Pristionchus pacificus dauer formation reveals cross-preference rather than self-preference of nematode dauer pheromones.* Proceedings of the Royal Society B-Biological Sciences, 2011. **278**(1719): p. 2784-2790.
23. Harvey, S., A. Shorto, and M. Viney, *Quantitative genetic analysis of life-history traits of Caenorhabditis elegans in stressful environments.* Bmc Evolutionary Biology, 2008. **8**.
24. Viney, M., M. Gardner, and J. Jackson, *Variation in Caenorhabditis elegans dauer larva formation.* Development Growth & Differentiation, 2003. **45**(4): p. 389-396.
25. Bubrig, L.T., J.M. Sutton, and J.L. Fierst, *Caenorhabditis elegans dauers vary recovery in response to bacteria from natural habitat.* Ecology and Evolution, 2020. **10**(18): p. 9886-9895.
26. Ow, M.C. and S.E. Hall, *Starvation memory resulting in reproductive plasticity is conserved in some Caenorhabditis elegans wild isolates.* MicroPubl Biol, 2020. **2020**.
27. Ow, M.C., A.M. Nichitean, and S.E. Hall, *Somatic aging pathways regulate reproductive plasticity in Caenorhabditis elegans.* Elife, 2021. **10**.
28. Pradhan, S., et al., *Environmental Programming of Adult Foraging Behavior in C. elegans.* Curr Biol, 2019. **29**(17): p. 2867-2879 e4.
29. Gilabert, A., et al., *Expanding the view on the evolution of the nematode dauer signalling pathways: refinement through gene gain and pathway co-option.* BMC Genomics, 2016. **17**: p. 476.
30. Bornholdt, S., *Less is more in modeling large genetic networks.* Science, 2005. **310**(5747): p. 449-+.
31. Bornholdt, S., *Boolean network models of cellular regulation: prospects and limitations.* Journal of The Royal Society Interface, 2008. **5**(Suppl 1): p. S85.
32. Wang, R., A. Saadatpour, and R. Albert, *Boolean modeling in systems biology: an overview of methodology and applications.* Physical Biology, 2012. **9**(5).
33. Pires-daSilva, A. and R. Sommer, *The evolution of signalling pathways in animal development.* Nature Reviews Genetics, 2003. **4**(1): p. 39-49.
34. Golden, J.W. and D.L. Riddle, *The Caenorhabditis elegans dauer larva: developmental effects of pheromone, food, and temperature.* Dev Biol, 1984. **102**(2): p. 368-78.
35. Neal, S.J., et al., *Feeding state-dependent regulation of developmental plasticity via CaMKI and neuroendocrine signaling.* Elife, 2015. **4**.
36. Rozum, J.C., et al., *Parity and time reversal elucidate both decision-making in empirical models and attractor scaling in critical Boolean networks.* Sci Adv, 2021. **7**(29).
37. Rozum, J.C., et al., *pystablemotifs: Python library for attractor identification and control in Boolean networks.* Bioinformatics, 2021.
38. Zanudo, J.G. and R. Albert, *An effective network reduction approach to find the dynamical repertoire of discrete dynamic networks.* Chaos, 2013. **23**(2): p. 025111.
39. Barna, J., et al., *Heat shock factor-1 intertwines insulin/IGF-1, TGF-beta and cGMP signaling to control development and aging.* BMC Dev Biol, 2012. **12**: p. 32.
40. Vowels, J.J. and J.H. Thomas, *Multiple chemosensory defects in daf-11 and daf-21 mutants of Caenorhabditis elegans.* Genetics, 1994. **138**(2): p. 303-16.
41. Gottlieb, S. and G. Ruvkun, *daf-2, daf-16 and daf-23: genetically interacting genes controlling Dauer formation in Caenorhabditis elegans.* Genetics, 1994. **137**(1): p. 107-20.
42. Albert, P.S. and D.L. Riddle, *Mutants of Caenorhabditis elegans that form dauer-like larvae.* Dev Biol, 1988. **126**(2): p. 270-93.

43. Maheshwari, P. and R. Albert, *A framework to find the logic backbone of a biological network.* BMC Syst Biol, 2017. **11**(1): p. 122.
44. Zanudo, J.G.T., G. Yang, and R. Albert, *Structure-based control of complex networks with nonlinear dynamics.* Proc Natl Acad Sci U S A, 2017. **114**(28): p. 7234-7239.
45. Riedl, R., *SYSTEMS-ANALYTICAL APPROACH TO MACRO-EVOLUTIONARY PHENOMENA.* Quarterly Review of Biology, 1977. **52**(4): p. 351-370.
46. Fernandes de Abreu, D.A., et al., *An insulin-to-insulin regulatory network orchestrates phenotypic specificity in development and physiology.* PLoS Genet, 2014. **10**(3): p. e1004225.
47. Fay, D., *Genetic mapping and manipulation: chapter 1--Introduction and basics.* WormBook, 2006: p. 1-12.
48. Keilin, D., *The problem of anabiosis and latent life: history and current concept.* Proceedings of the Royal Society Series B, 1959. **150**: p. 149-191.
49. Kalimari, M., et al., *Comparative transcriptomics suggest unique molecular adaptations within tardigrade lineages.* BMC Genomics, 2019. **20**: p. 607.
50. Gibbs, H., *Hypobiosis in parasitic nematodes— an update.* Advances in Parasitology, 1986. **25**: p. 129-74.
51. Ma, G., et al., *Dauer signalling pathway model for Haemonchus contortus.* Parasites & Vectors, 2019. **12**.
52. Mayer, M., et al., *The Orphan Gene dauerless Regulates Dauer Development and Intraspecific Competition in Nematodes by Copy Number Variation.* PLoS Genetics, 2015. **11**(6): p. e1005146.
53. Harris, T.W., et al., *WormBase: a modern Model Organism Information Resource.* Nucleic Acids Research, 2019. **48**(D1): p. D762-D767.
54. Henderson, S.T. and T.E. Johnson, *daf-16 integrates developmental and environmental inputs to mediate aging in the nematode Caenorhabditis elegans.* Curr Biol, 2001. **11**(24): p. 1975-80.
55. Lin, K., et al., *Regulation of the Caenorhabditis elegans longevity protein DAF-16 by insulin/IGF-1 and germline signaling.* Nat Genet, 2001. **28**(2): p. 139-45.
56. Lee, R.Y., J. Hench, and G. Ruvkun, *Regulation of C. elegans DAF-16 and its human ortholog FKHRL1 by the daf-2 insulin-like signaling pathway.* Curr Biol, 2001. **11**(24): p. 1950-7.
57. Paradis, S. and G. Ruvkun, *Caenorhabditis elegans Akt/PKB transduces insulin receptor-like signals from AGE-1 PI3 kinase to the DAF-16 transcription factor.* Genes Dev, 1998. **12**(16): p. 2488-98.
58. Matyash, V., et al., *Sterol-derived hormone(s) controls entry into diapause in Caenorhabditis elegans by consecutive activation of DAF-12 and DAF-16.* PLoS Biol, 2004. **2**(10): p. e280.
59. Dowell, P., et al., *Convergence of peroxisome proliferator-activated receptor gamma and Foxo1 signaling pathways.* J Biol Chem, 2003. **278**(46): p. 45485-91.
60. Hopfensitz, M., et al., *Attractors in Boolean networks: a tutorial.* Computational Statistics, 2013. **28**(1): p. 19-36.
61. Bornholdt, S., *Boolean network models of cellular regulation: prospects and limitations.* J R Soc Interface, 2008. **5 Suppl 1**: p. S85-94.
62. Albert, I., et al., *Boolean network simulations for life scientists.* Source Code Biol Med, 2008. **3**: p. 16.

63. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome Res, 2003. **13**(11): p. 2498-504.

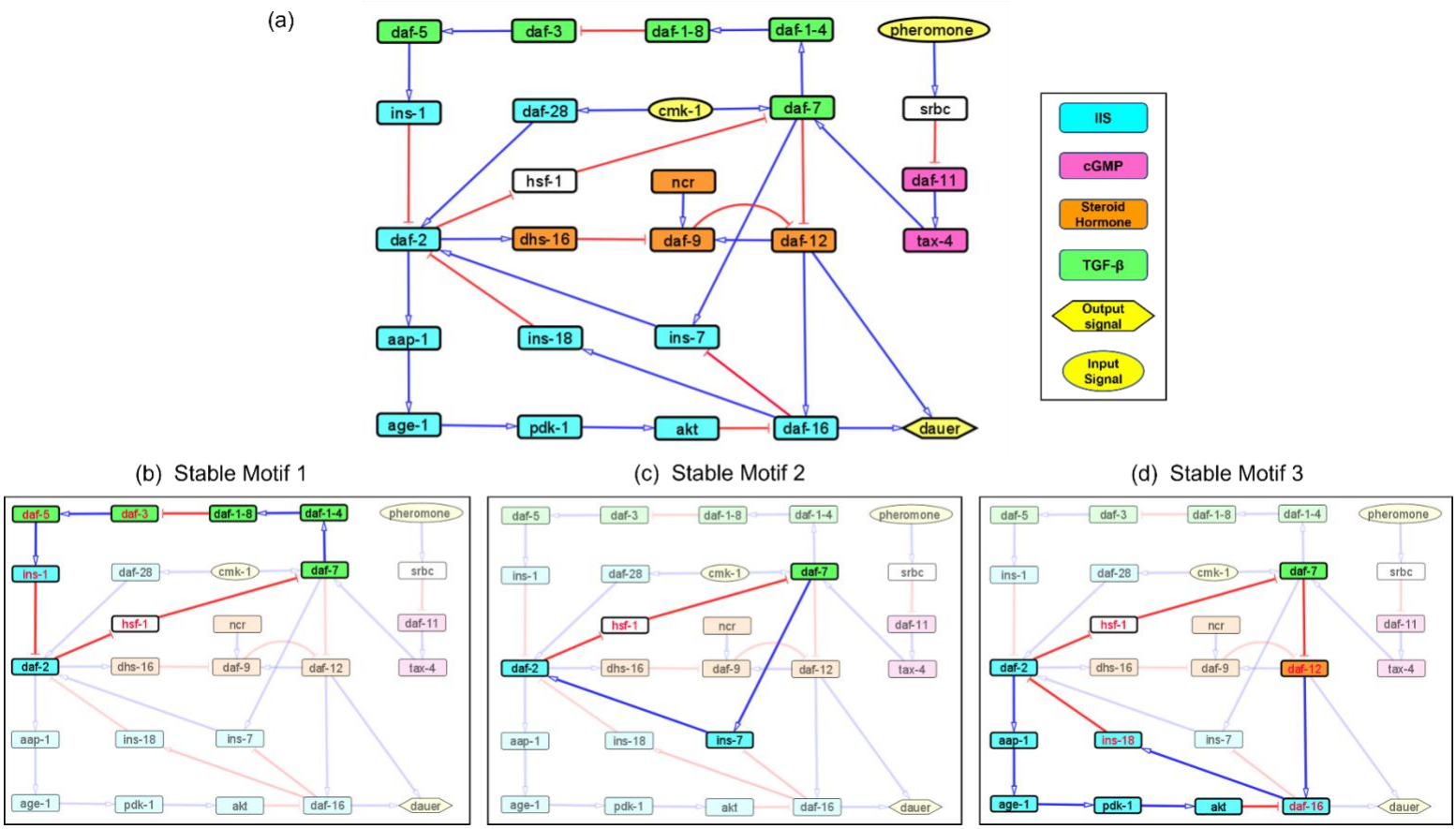


Figure 1 Models representing the four major dauer inducing pathways and the stable motifs of this network. (a) The dauer Boolean network included 26 nodes and 34 edges. Each node represented a molecular entity (gene, receptor, protein, etc.) involved in signal transduction pathways. Edges indicated regulatory interactions between the nodes that activated (signified by blue arrow heads) or inhibited (signified by red 'T' edges). The four dauer inducing pathways were represented by four different colors. Here, *cmk-1* and pheromone were input signals which when turned on and off, decided the dauer fate. Nodes *srbc* (pheromone receptor) and *hsf-1* (heat shock protein) did not belong to any of the four dauer pathways. (b), (c) and (d) Subgraphs that were part of the cyclic SMs 1, 2 and 3 respectively, into which the dauer Boolean network stabilizes. The nodes in red text were inhibited and nodes in black text were activated.

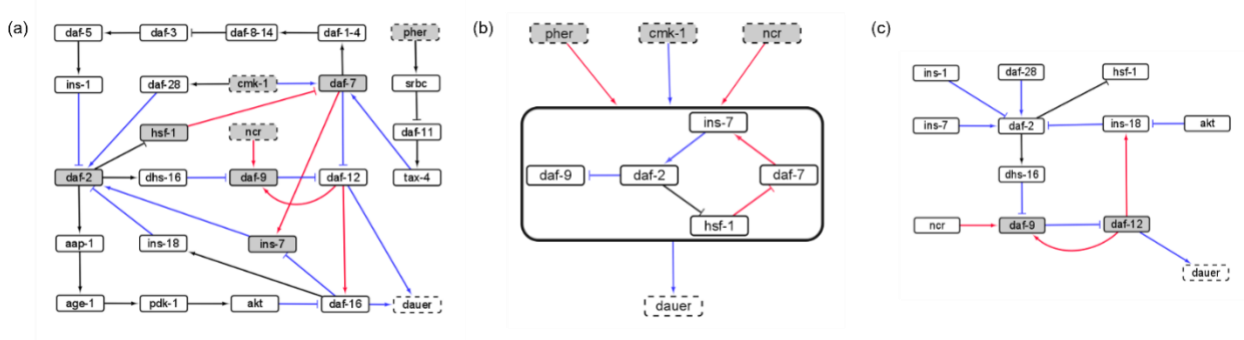


Figure 2 The causal logic backbone of the dauer Boolean network. (a) The edges were assigned causal logic (blue = necessary, red = sufficient and black = sufficient and necessary). Nodes filled grey with dashed borders were input signals where grey filled nodes were part of cyclic subgraphs (SMs) and solid borders indicated output nodes. (b) The logic backbone derived from identifying subgraphs from input to motif nodes and motif nodes to output signals. Activation of pher and *ncr* was sufficient to reach the motif and *cmk-1* was necessary. The motif stabilization was sufficient to control dauer fate. (c) The reduced dauer logic network consisted of 12 nodes and 13 edges. *Daf-9* and *daf-12* entered a cyclic feedback loop. The insulin factors involved in different endocrine pathways were important points of the network regulating *daf-2* activity and present in the reduced network.

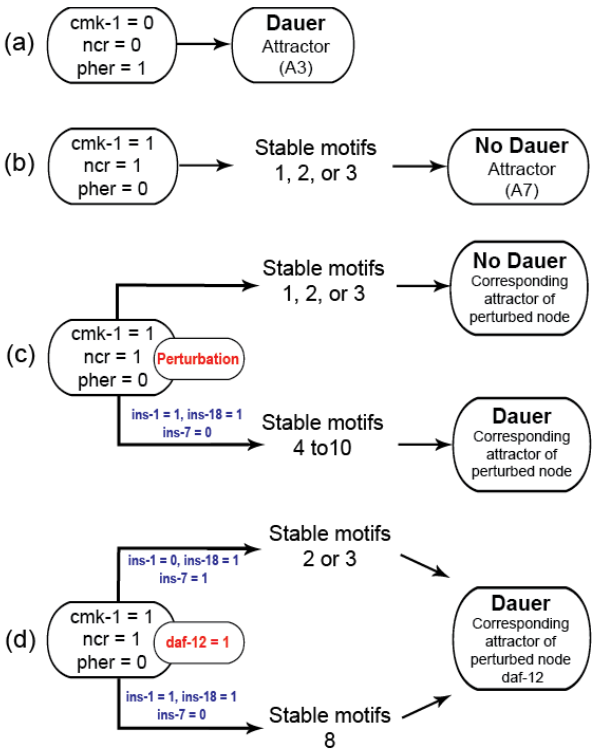


Figure 3 Model transitions from initial conditions to attractors. (a) Under unfavourable conditions, the system eventually reached the "Dauer" attractor A3 without settling in any of the cyclic motifs. (b) Under favorable conditions (*cmk-1* is on and pheromone is off), the network stabilized in either of the three cyclic SMs which led to the attractor A7 and the "No dauer" phenotype. (c) When a node of SM 1, 2 or 3 was perturbed the system stabilized in different SMs and converged on either dauer arrest or normal development. (d) *daf-12* perturbation resulted in dauer arrest through SMs 2, 3 or 8. *daf-12* presence, irrespective of initial conditions, was sufficient for dauer formation. SMs 4-10 and the attractors of the perturbed node simulations are given in the Supplementary Materials.

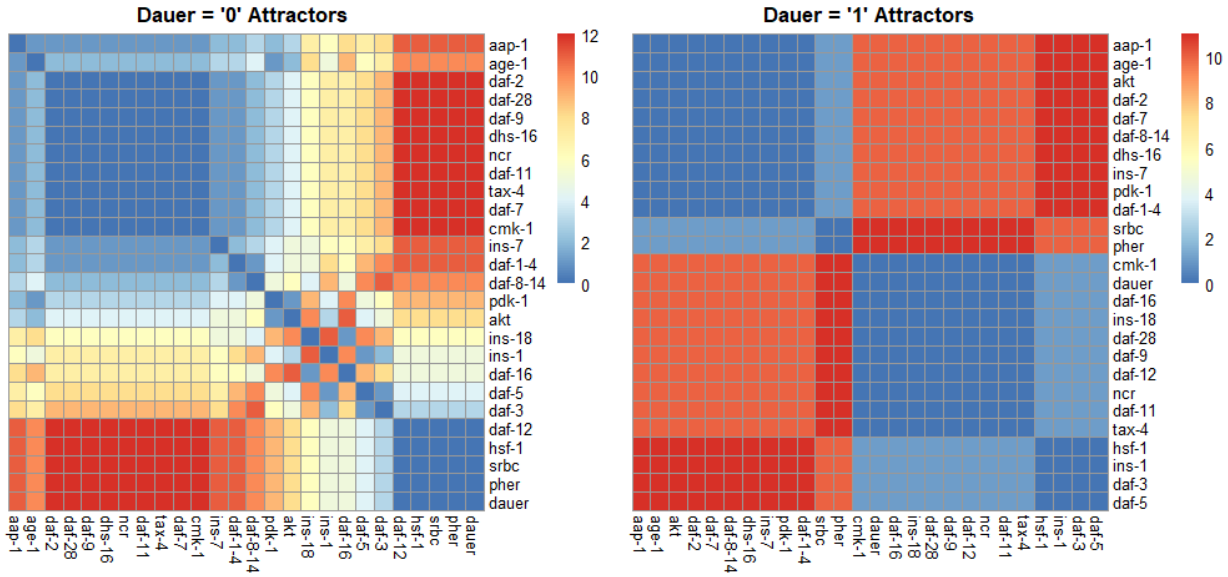


Figure 4 Correlation plots as a function of hamming distance for every node of the attractors converging to either "dauer" or "no dauer" when the nodes of the SMs were perturbed. Hamming distance = 0 implied that the two nodes were in the same state and as the distance value increased the states in which these two nodes exist in the attractors varied. From the dauer=1 correlation plot we observed clusters of nodes that existed in same state across the four dauer inducing pathways. *Ins-1* and *ins-18* were positively and negatively correlated when attractors converged on dauer and no dauer phenotypes, respectively.

538 **Supplementary Table 1** Regulatory interactions of the dauer Boolean network.

Upstream node	Downstream node	Interaction	PubMed ID
aap-1	*age-1*	Activating	12393910
age-1	*pdk-1*	Activating	10364160
*akt**	*daf-16*	Inhibitory	11747825, 11381260, 11747821, 9716402
cmk-1	*daf-28*	Activating	26335407
cmk-1	*daf-7*	Activating	26335407
daf-1	*daf-8-14**	Activating	10887089
*daf-11**	*tax-4*	Activating	8893027
daf-12	*daf-16*	Activating	15383841, 12966085
daf-12	*daf-16*	Activating	15383841
daf-12	*daf-9*	Activating	15084461
daf-12	dauer	Activating	15383841
daf-16	*ins-18*	Activating	22683638
daf-16	*ins-7*	Inhibitory	12845331
daf-16	*ins-7*	Inhibitory	19489741
daf-16	dauer	Activating	15383841
daf-2	*aap-1*	Activating	12393910
daf-2	*dhs-16*	Activating	22505847
daf-2	*hsf-1*	Inhibitory	23116063
daf-28	*daf-2*	Activating	12654727, 27125673
daf-3	*daf-5*	Activating	14681186
daf-5	*ins-1*	Activating	21533078
daf-7	*daf-1-4*	Activating	10887089, 8413626, 8570636
daf-7	*ins-7*	Activating	19489741
*daf-8-14**	*daf-3*	Inhibitory	20081192, 14992718
dhs-16	*daf-9*	Inhibitory	22505847
hsf-1	*daf-7*	Inhibitory	23116063
ins-1	*daf-2*	Inhibitory	21343369, 11274053
ins-18	*daf-2*	Inhibitory	22683638
ins-7	*daf-2*	Activating	12845331
*ncr**	*daf-9*	Activating	15509773

pdk-1	*akt**	Activating	10364160
*pher**	*srbc**	Activating	19797623
*srbc**	*daf-11**	Inhibitory	19797623
tax-4	*daf-7*	Activating	16903785

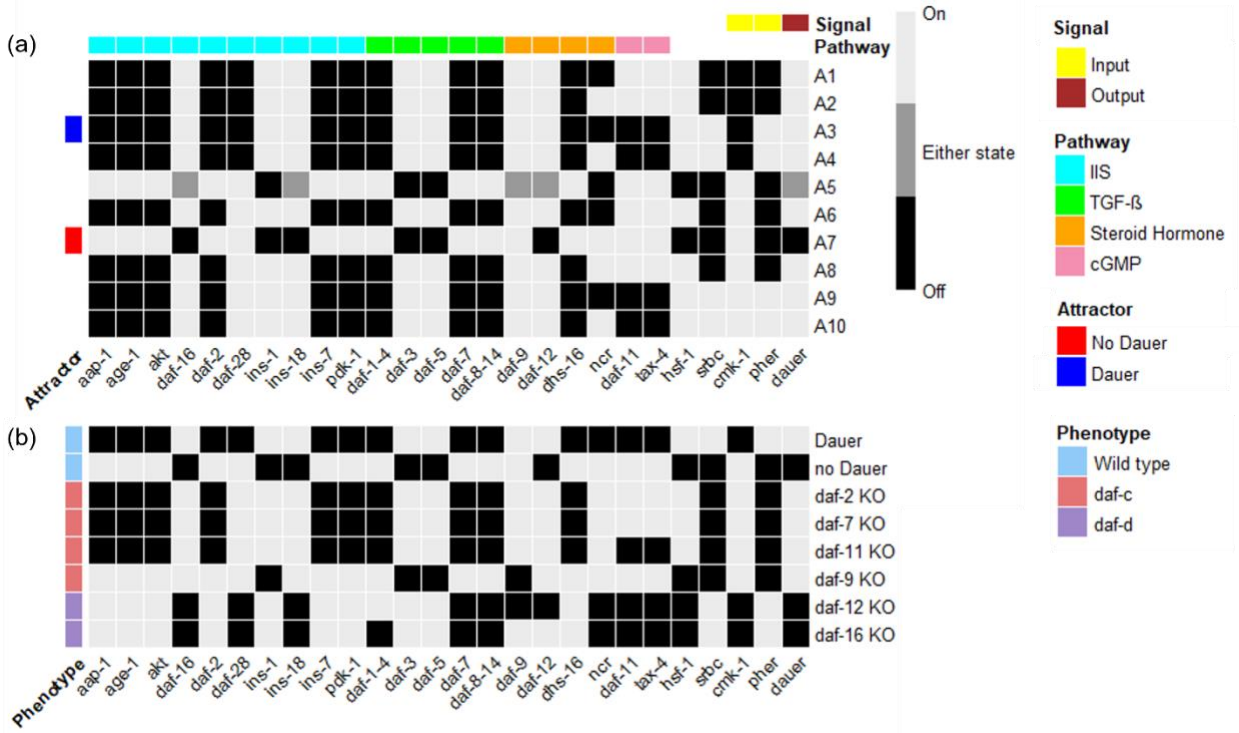
*****The nodes names used in literature have been modified for convenience. Here, *daf-11* = cGMP, pher = Pheromone, *ncr* = *ncr-1/2, daf-8-14* = *daf-8* and *daf-14*, *akt* = *akt-1/2*, *srbc* = *srbc-64* and *srbc-66*.

Supplementary Table 2 The Boolean rules of the dauer network. Initial conditions for dauer were *cmk-1* = 0, *ncr* = 0 and pher = 1. Initial conditions for normal development were *cmk-1* = 1, *ncr* = 1 and pher = 0. The state of each node at time 't+1' was calculated according to the states of upstream regulator nodes at time 't'.

Downstream node *= Upstream regulators
*aap-1 *= daf-2*
*age-1 *= aap-1*
*akt *= pdk-1*
*daf-11 *=* not *srbc*
*daf-12 *=* not *daf-9* or not *daf-7*
*daf-1-4 *= daf-7*
*daf-16 *=* not *akt* or *daf-12*
*daf-2 *= (ins-7* and *daf-28)* or *(*not *ins-1* or not *ins-18)*
*daf-28 *= cmk-1*
*daf-3 *=* not *daf-8-14*
*daf-5 *= daf-3*
*daf-7 *= (tax-4* and *cmk-1)* and not *hsf-1*
*daf-8-14 *= daf-1-4*
*daf-9 *=* not *dhs-16* or *ncr* or *daf-12*
*dauer *= daf-16* and *daf-12*
*dhs-16 *= daf-2*
*hsf-1 *=* not *daf-2*
*ins-1 *= daf-5*
*ins-18 *= daf-16*
*ins-7 *= daf-7* or not *daf-16*
*pdk-1 *= age-1*
*srbc *= pher*
*tax-4 *= daf-11*

Supplementary Table 3 Stable motifs in the dauer Boolean network.

Name	Stable motif	Perturbation
SM 4	*{'ins-18': 1, 'pdk-1': 0, 'daf-5': 1, 'akt': 0, 'age-1': 0, 'daf-16': 1, 'aap-1': 0, 'hsf-1': 1, 'daf-3': 1, 'daf-8-14': 0, 'daf-2': 0, 'ins-1': 1, 'daf-7': 0, 'daf-1-4': 0}*	*ins-7*=0
SM 5	*{'ins-18': 1, 'daf-5': 1, 'daf-12': 1, 'daf-16': 1, 'hsf-1': 1, 'daf-3': 1, 'daf-8-14': 0, 'daf-2': 0, 'ins-1': 1, 'daf-7': 0, 'daf-1-4': 0}*	*ins-7*=0
SM 6	*{'pdk-1': 0, 'daf-5': 1, 'akt': 0, 'age-1': 0, 'daf-16': 1, 'aap-1': 0, 'hsf-1': 1, 'daf-3': 1, 'daf-8-14': 0, 'daf-2': 0, 'ins-1': 1, 'daf-7': 0, 'ins-7': 0, 'daf-1-4': 0}*	*ins-18*=1
SM 7	*{'daf-5': 1, 'daf-12': 1, 'daf-16': 1, 'hsf-1': 1, 'daf-3': 1, 'daf-8-14': 0, 'daf-2': 0, 'ins-1': 1, 'daf-7': 0, 'ins-7': 0, 'daf-1-4': 0}*	*ins-18*=1
SM 8	*{'daf-5': 1, 'daf-12': 1, 'daf-16': 1, 'hsf-1': 1, 'daf-3': 1, 'daf-8-14': 0, 'daf-2': 0, 'ins-1': 1, 'daf-7': 0, 'ins-7': 0, 'daf-1-4': 0}*	*daf-12*=1, *daf-16*=1, *akt*=0, *pdk-1*=0, *age-1*=0, *aap-1*=0
SM 9	*{'ins-18': 1, 'daf-12': 1, 'daf-16': 1, 'daf-2': 0, 'daf-7': 0, 'ins-7': 0, 'hsf-1': 1}*	*ins-1*=1, *daf-5*=1, *daf-3*=1, *daf-8-14*=0, *daf-1-4*=0
SM 10	*{'ins-18': 1, 'pdk-1': 0, 'akt': 0, 'age-1': 0, 'daf-16': 1, 'hsf-1': 1, 'daf-2': 0, 'daf-7': 0, 'ins-7': 0, 'aap-1': 0}*	*ins-1*=1, *daf-5*=1, *daf-3*=1, *daf-8-14*=0, *daf-1-4*=0



Supplementary Figure 1 Trajectories of attractors and phenotypes. (a) A heatmap that shows the states of every node of the system in its steady state for 10 different attractors (A1 – A10) identified for this dauer network. Attractors A3 and A7 represent the two wild type phenotypes of this system i.e. “dauer” and “no dauer”. The nodes part of the four different pathways and those that are output/input signals are annotated. (b) A similar heatmap for different phenotypes exhibited by the system.