

Research Article

Investigating the Relationship between Topology and Evolution in a Dynamic Nematode Odor Genetic Network

David A. Fitzpatrick¹ and Damien M. O'Halloran^{2,3}

¹Genome Evolution Laboratory, Department of Biology, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland

²Department of Biological Sciences, The George Washington University, 333 Lisner Hall, 2023 G Street NW, Washington, DC 20052, USA

³Institute for Neuroscience, The George Washington University, 636 Ross Hall, 2300 I Street NW, Washington, DC 20037, USA

Correspondence should be addressed to Damien M. O'Halloran, damienoh@gwu.edu

Received 17 May 2012; Revised 6 August 2012; Accepted 29 August 2012

Academic Editor: Amitabh Joshi

Copyright © 2012 D. A. Fitzpatrick and D. M. O'Halloran. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The relationship between biological network architectures and evolution is unclear. Within the phylum nematoda olfaction represents a critical survival tool. For nematodes, olfaction contributes to multiple processes including the finding of food, hosts, and reproductive partners, making developmental decisions, and evading predators. Here we examine a dynamic nematode odor genetic network to investigate how divergence, diversity, and contribution are shaped by network topology. Our findings describe connectivity frameworks and characteristics that correlate with molecular evolution and contribution across the olfactory network. Our data helps guide the development of a robust evolutionary description of the nematode odor network that may eventually aid in the prediction of interactive and functional qualities of novel nodes.

1. Introduction

For nematodes, olfaction is a central mode of survival. Olfaction contributes to the finding of food, hosts, reproductive partners, in the making of developmental decisions, and to the evasion from predators. Studies into the olfactory system of the model nematode *Caenorhabditis elegans* have yielded detailed descriptions of the molecular and cellular pathways that subserve the olfactory system [1–4]. These signaling pathways appear highly conserved across very divergent nematode species, and the sensory neurons have clear anatomical orthologs in distantly related nematodes [5]. Within the olfactory system of *C. elegans*, the ability to detect dilute volatile odors is mostly conferred by three pairs of neurons termed the Amphid Wing cells type A (AWA), Amphid Wing cells type B (AWB), and the Amphid Wing cells type C (AWC) [3, 4]. These cells are primary sensory neurons located within the sensory amphid organ of the head that forms part of an anatomically distinct subclass of amphid neurons in that they do not extend processes through the amphid opening, but instead their distal ciliated

endings terminate inside a glial sheath cell [6]. Here we describe a composite odor genetic network in *C. elegans* that encompasses all three pairs of volatile odor-detecting neurons. We used this network to identify orthologous genes of the odor network in the nematode *Pristionchus pacificus*. We chose the nematode *P. pacificus* based upon three criteria: (1) it exhibits a similar lifestyle to that of *C. elegans*, in that they are both self-fertilizing hermaphrodites that will feed on bacteria, (2) conservation of the olfactory signaling pathway of *P. pacificus* to *C. elegans* has been validated experimentally [7, 8], and (3) because it is sufficiently divergent to generate meaningful divergence data (estimated divergence between *Caenorhabditis* and *Pristionchus* is between 280 and 430 million years ago [9, 10]). One of our goals in this study was to compare selective pressures across the nodes of the odor network by comparing substitutions at silent sites to that of nonsilent sites; this analysis is best done using more divergent sequences [11]. Using this nematode odor network we searched for relationships between pathway position with divergence, diversity, and contribution across the network. Then, we designed an interaction map of the network

TABLE 1: Orthology support for each node within the odor-signaling network.

| <i>C. elegans</i> gene | <i>P. pacificus</i> ortholog | Evidence of orthology |
|---------------------------|------------------------------|-----------------------------------|
| <i>goa-1</i> | PPA23648/ <i>Ppa-go-1</i> | WormBase-Compara |
| <i>odr-3</i> | PPA14189/ <i>Ppa-odr-3</i> | InParanoid; OMA; WormBase-Compara |
| <i>egl-30</i> | PPA31690/ <i>Ppa-egl-30</i> | WormBase-Compara |
| <i>gpa-3</i> | PPA09867/ <i>Ppa-gpa-3</i> | InParanoid; OMA; WormBase-Compara |
| <i>gpa-5</i> | PPA10789 | OMA; WormBase-Compara |
| <i>gpa-13</i> | PPA31402/ <i>Ppa-gpa-13</i> | WormBase-Compara |
| <i>eat-4</i> | PPA15025/ <i>Ppa-eat-4</i> | InParanoid; WormBase-Compara |
| <i>dgk-1</i> | PPA00312/ <i>Ppa-dgk-1</i> | InParanoid; OMA; WormBase-Compara |
| <i>odr-1</i> | PPA17708 | InParanoid; OMA |
| <i>daf-11</i> | PPA14907/ <i>Ppa-daf-11</i> | WormBase-Compara |
| <i>tax-6</i> | PPA09320/ <i>Ppa-tax-6</i> | WormBase-Compara |
| <i>pkg-1</i> | PPA27475/ <i>Ppa-pkg-1</i> | InParanoid; OMA; WormBase-Compara |
| <i>tax-2</i> | PPA07436/ <i>Ppa-tax-2</i> | InParanoid; OMA; WormBase-Compara |
| <i>tax-4</i> | PPA02388/ <i>Ppa-tax-4</i> | InParanoid; OMA; WormBase-Compara |
| ¹ <i>rgs-3</i> | PPA01970/ <i>Ppa-rgs-3</i> | InParanoid; OMA; WormBase-Compara |
| ¹ <i>arr-1</i> | PPA21763/ <i>Ppa-arr-1</i> | WormBase-Compara |

¹Relationship support was confirmed by reconstructing phylogenetic trees for each gene except *rgs-3* and *arr-1* as they represent single-gene families, and therefore not possible to reconstruct phylogeny (see Figures S1–S7).

and investigated relationships between various metrics of interaction with molecular evolution, and contribution across the network. Our data is not a definitive map of odor signaling in *C. elegans* but represents a snap shot of current data, and by uncovering robust associations between network topology, evolution, and function we may ultimately design a framework that facilitates a level of predictive power over novel nodes within the network.

2. Materials and Methods

2.1. Sequences. *Pristionchus pacificus* orthologs were located by cross-referencing matches using the orthology databases: InParanoid.7 [12], OrthoMCL database [13], and the OMA orthology matrix Browser [14]. For each node in our network WormBase (version WS231) has defined a curated ortholog using WormBase-Compara [15]; however, we corroborated these predictions using reciprocal Blast [16] searches, and, by inferring relationships by reconstructing phylogenetic trees (see Figures S1–S7 in Supplementary Materials available online at doi:10.1155/2012/548081), we outline all the evidence for orthology in Table 1. Orthologs were aligned using the multiple sequence alignment software MUSCLE v3.6 [17], and gaps were systematically stripped from all sequences after alignment, and phylogenies inferred using PhyML [18]. To determine orthologs, WormBase-Compara uses the databases TreeFam [http://www.treefam.org/], InParanoid.7, KOGs [http://www.ncbi.nlm.nih.gov/COG/], OMA. The genes used in our network (Table 1), and their corresponding ortholog in *P. pacificus* (WormBase identifier for *P. pacificus* is denoted *Ppaxxx*), were *goa-1* (G_{α} subunit protein: *Ppa-go-1*); *egl-30* (G_{α} subunit protein: *Ppa-egl30*); *dgk-1*

(DGK θ , diacylglycerol kinase theta: *Ppa-dgk1*); *eat-4* (vesicular glutamate transporter: *Ppa-eat4*); *egl-4/pkg-1* (Protein Kinase G: *Ppa-pkg1*); *tax-6* (calcineurin type A: *Ppa-tax6*); *odr-1* (receptor guanylyl cyclase: *Ppa17708*); *daf-11* (receptor guanylyl cyclase: *Ppa-daf11*); *tax-2* (cyclic nucleotide-gated channel β subunit: *Ppa-tax2*); *tax-4* (cyclic nucleotide-gated channel α subunit: *Ppa-tax4*); *odr-3* (G-protein α subunit: *Ppa-odr3*); *gpa-3* (G-protein α subunit: *Ppa-gpa3*); *gpa-5* (G protein α subunit: *Ppa10789*); *gpa-13* (G-protein α subunit: *Ppa-gpa13*); *arr-1* (Arrestin: *Ppa-arr1*); *rgs-3* (regulator of G-protein signaling: *Ppa-rgs3*). To detect *C. elegans* versus *P. pacificus* 1 : 1 candidate orthologs for the randomization study we selected orthologs from the orthology database, InParanoid.7 [12]. We only included orthologs that are represented by 100% bootstrap support, and from this approach we obtain 5,666 1 : 1 orthologs.

2.2. Analysis of Molecular Data. Synonymous (d_S) and nonsynonymous (d_N) substitution rates for orthologs were estimated using the methods of Yang and Nielsen [19] as implemented in yn00 in the PAML suite [20]. To test the null hypothesis that there is no above average selective pressure on these genes, we performed a randomization test where we determined the average d_N/d_S value for 50,000 randomly assembled networks and compared with the average d_N/d_S value for our odor network. Random networks were 16 genes in size and sampling permitted replacement. Measures of nucleotide diversity (π) were performed using DnaSP version 5 [21].

2.3. Network Analysis. Selection of vertices for the odor network was determined by mining literature databases. Network analysis was calculated using Cytoscape version 2.8.2

[22]. The network was treated as undirected and all network analyses available through Cytoscape version 2.8.2 were examined; these are average shortest path length, betweenness centrality, closeness centrality, clustering coefficient, degree, eccentricity, neighborhood connectivity, radiality, stress, and topological coefficient. The degree (k) of a node n is defined as the number of edges linked to n . The clustering coefficient (C_n) reveals how connected the neighborhood of a node is by calculating the fraction of neighboring pairs, and for a node n it is defined as

$$C_n = \frac{2e_n}{k_n(k_n - 1)}, \quad (1)$$

where k_n is the degree of n and e_n is the number of connected pairs between all neighbors of n . Betweenness centrality (C_b) is a measure of the fraction of shortest paths between node pairs (s, t) that pass through the node of interest, and for the node n it is calculated using the following formula:

$$C_b(n) = \sum_{s \neq n \neq t} \frac{\sigma_{st}(n)}{\sigma_{st}}, \quad (2)$$

where s and t are nodes different from n , σ_{st} denotes the number of shortest paths from s to t , and $\sigma_{st}(n)$ is the number of shortest paths from s to t on which n lies.

All chemotaxis indices (for both wildtype and mutant animals) were mined from previous publications (see references below). The chemotaxis index difference (C.I._{diff}) for each mutant within each neuron (AWA, AWB, and AWC) was determined by calculating the difference between the wildtype chemotaxis index (C.I._{wt}) and mutant chemotaxis index (C.I._{mut}). For example, a mutant that presents an average AWA chemotaxis defect C.I. = 0.5, compared to the wildtype C.I. (C.I._{wt-AWA}) of 0.9, would have a C.I._{diff-AWA} = 0.4 (i.e., subtract 0.5 from 0.9). In the case of long-term adaptation (LTA) mutants the C.I._{diff} was calculated by subtracting the C.I. value for LTA from C.I._{wt} of unadapted animals. Wildtype odortaxis: 0.9 (C.I._{wt AWA}); -0.95 (C.I._{wt AWB}); 0.85 (C.I._{wt AWC}) [4, 23]. Wildtype long-term adaptation: 0.5 (C.I._{diff AWA}); 0.65 (C.I._{diff AWC}) [24, 25]. *odr-3*: 0.6 (C.I._{diff AWA}); -0.7 (C.I._{diff AWB}); 0.5 (C.I._{diff AWC}) [23, 26]. *rgs-3*: 0.85 (C.I._{diff AWC}) [27]. *daf-11*: 0.6 (C.I._{diff AWB}); 0.5 (C.I._{diff AWC}) [28]. *dgk-1*: 0 (C.I._{diff AWC}) [29]. *eat-4*: 0.55 (C.I._{diff AWC}) [30]. *egl-30*: 0.65 (C.I._{diff AWC}) [29]. *goa-1*: 0.6 (C.I._{diff AWA}); 0.6 (C.I._{diff AWC}) [29]. *odr-1*: 0 (C.I._{diff AWA}); 0.85 (C.I._{diff AWC}) [31, 32]. *tax-2*: 0.2 (C.I._{diff AWA}); -0.8 (C.I._{diff AWB}), 0.75 (C.I._{diff AWC}) [23, 33]. *tax-4*: 0.5 (C.I._{diff AWA}); 0.75 (C.I._{diff AWC}) [23, 34]. *egl-4*: 0.8 (C.I._{diff AWA}); 0.65 (C.I._{diff AWC}) [25, 34]. The phenotype index (PI) was then calculated as follows:

$$PI = \sqrt{\left(\frac{C.I.\text{diff-mutx-AWA}}{CI_{wt-AWA}}\right)^2 + \left(\frac{C.I.\text{diff-mutx-AWB}}{CI_{wt-AWB}}\right)^2 + \left(\frac{C.I.\text{diff-mutx-AWC}}{CI_{wt-AWC}}\right)^2}. \quad (3)$$

The PI for LTA mutants was measured by calculating the relative contribution of the LTA mutant to the adaptation response. For example, if the LTA response of wildtype animals normally decreases the C.I. value to 0.3 from 0.9, then the C.I._{diff} = 0.6; if the mutant is fully required (i.e., there is no LTA response at all in the mutant) for the LTA response, then the PI is calculated as follows: 0.6/0.6 = 1; which reflects the total contribution of the gene product to LTA in wildtype animals but that which is defective in the LTA mutant. The weighted phenotype index (WPI) was calculated for each group of mutants as follows:

$$\sum_{i=1}^n \frac{(P.I.\text{mut}^i)(wmut^i)}{(wmut^i)}, \quad (4)$$

where P.I._{mut^x} is the phenotype index (P.I.) for the mutant x (mut^x), and wmut^x represents the number of odors for which mutant x (mut^x) exhibits a mutant phenotype.

3. Results

To compare divergence rate with pathway position in our network of dynamic odor transduction, we identified orthologous genes of the *C. elegans* odor network in *P. pacificus*, and divided the odor-signaling pathway members (Figure 1(a)) within this network into 3 categories: (I) regulators Class 1

(GPA-3 [35, 36], ODR-3 [26], GPA-5 [35, 36], GPA-13 [35, 36], ARR-1 [37] and RGS-3 [27]); (II) regulators Class 2 (EGL-4 [25], TAX-6 [38], ODR-1 [31], DAF-11 [28], TAX-2 [33], and TAX-4 [39]); (III) actuators (GOA-1 [40], EGL-30 [40], DGK-1 [40], and EAT-4 [41]). After binning pathway members into each category, we then calculated divergence within each group by comparing substitutions at silent sites to that of nonsilent sites and looked for patterns of correlation between topology and evolutionary rate. We did not observe a significant association between pathway position and divergence (pathway position: Spearman rank-order correlation coefficient $r_s = 0.37$, $P = 0.15$, Kendall's $\tau = 0.31$, $P = 0.13$). We also designed an orthologous intragenus network using *Caenorhabditis japonica* and compared substitutions at silent sites to that of non-silent sites and looked for a pattern of correlation with pathway position. Using this intragenus network we also did not observe significant association between pathway position and divergence (pathway position: $r_s = 0.28$, $P = 0.3$, Kendall's $\tau = 0.17$, $P = 0.4$), suggesting that the pathway position alone may not shape constraint in our odor network.

Next, we designed an interaction map of the network and tested for correlations between divergence and various topological metrics (Figure 1(b)). The interaction-based odor network is a composite map of the pathways within the three pairs of neurons that subserve volatile odor recognition

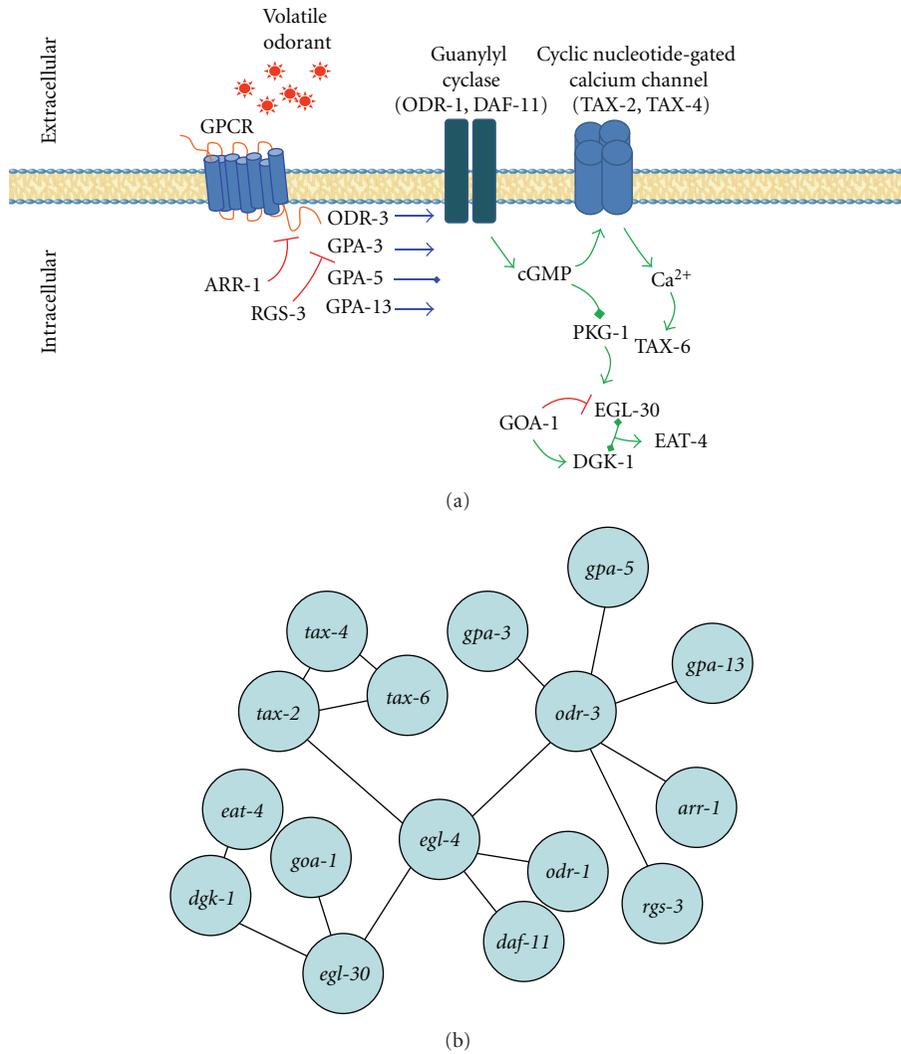


FIGURE 1: Odor-signaling pathway in *C. elegans* illustrated by pathway position and based on genetic interactions. (a) Summary of volatile odor signaling in *C. elegans*. (b) Schematic of an interaction-based odor network from *C. elegans*.

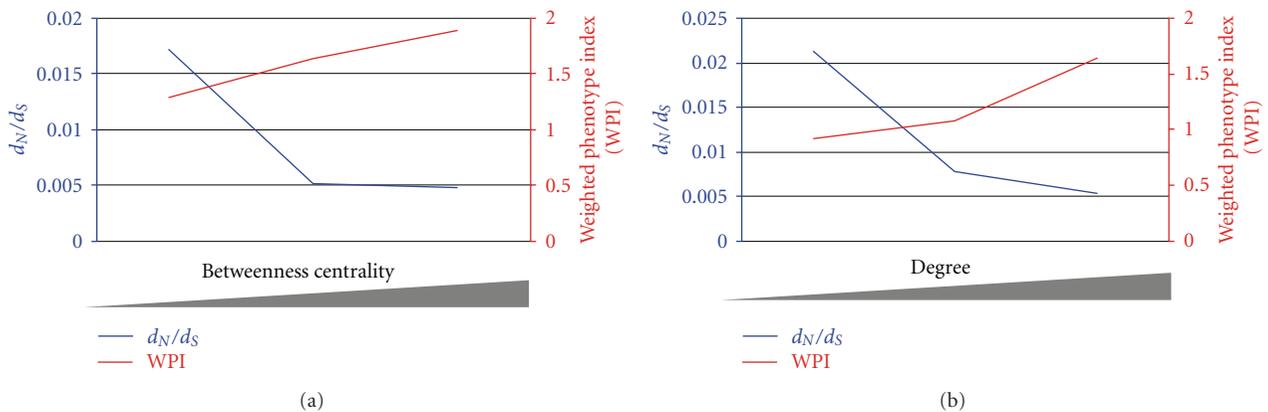


FIGURE 2: Plotting the relationship between divergence with measures of betweenness centrality or degree, and the relationship between contribution with measures of betweenness centrality or degree in a dynamic nematode odor network. (a) Divergence (left y-axis in blue) plotted against increasing measures of betweenness centrality (x-axis), and contribution (right y-axis in red) plotted against increasing measures of betweenness centrality. (b) Divergence (left y-axis in blue) plotted against increasing measures of degree (x-axis), and contribution (right y-axis in red) plotted against increasing measures of degree.

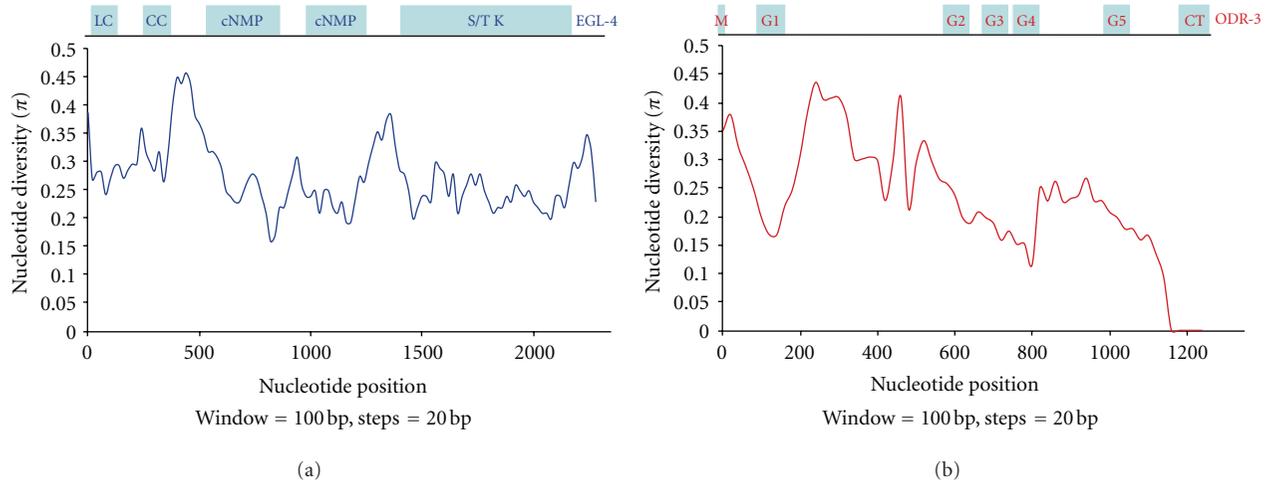


FIGURE 3: Sliding windows of nucleotide diversity within the coding regions of the genes *egl-4* and *odr-3*. Nucleotide position in the alignments (x -axis) is plotted against nucleotide diversity (π) on the y -axis. (a) Nucleotide diversity in *egl-4*. The shaded boxes on top of the graph illustrate the various functional domains within the EGL-4 protein. LC: low-complexity domain; CC: coiled-coil domain; cNMP: cyclic nucleotide-binding domain, S/TK: serine/threonine kinase domain. (b) Nucleotide diversity in *odr-3*. The shaded boxes on top of the graph illustrate the various functional domains within the ODR-3 protein. M: fatty-acid modification site; G1–G5: five alpha helices that comprise the GTPase domain; CT: C-terminal receptor interaction domain. In each case a 100-base-pair window was analyzed in 20-base-pair increments.

in *C. elegans* (these cells are the AWA, AWB, and AWC). The interaction network was designed based upon genetic interactions and localization studies. Edges may represent interactions from only one of the three odor sensing pairs of neurons, but placed in a composite framework of all three pairs of odor-sensing neurons to obtain an informative number of vertices. The network was treated as undirected. The network comprised 16 nodes with a network diameter of 5, a characteristic path length of 2.825, and an average number of neighbors equal to 2. For the interaction network we observed that a strong correlation between number of neighbors and average neighborhood connectivity (correlation = 0.857; coefficient of determination $R^2 = 0.648$). Plotting the number of neighbors against average clustering coefficient yields a strong pattern of correlation using a linear function (correlation = 0.868; coefficient of determination $R^2 = 0.753$); we observed that a decrease in clustering coefficient with an increase in number of neighbors as is typical of many networks. Network analysis was performed and the output of the analyses was tested for correlations with divergence by calculating the evolutionary rate for each node by comparing substitutions at silent sites to that of non-silent sites with the orthologous genes of the odor network in *P. pacificus*. We observed a significant association in two cases: (1) measures of betweenness centrality and (2) measures of degree. By comparing measures of betweenness or measures of degree with divergence we observed significant negative correlations (betweenness: $r_s = -0.48$, $P = 0.02$, Kendall's $\tau = -0.42$, $P = 0.03$; degree: $r_s = -0.57$, $P = 0.01$, Kendall's $\tau = -0.44$, $P = 0.03$). Comparing measures of degree to measures of betweenness we observed a strong correlation (correlation = 0.936; $R^2 = 0.877$); this is typical of networks containing nodes of high influence that increase

their influence with increases in their number of edges and is not specific for our odor network. Then, we calculated molecular diversity across the interaction network and searched for patterns of association with metrics from the following characteristics of the network: pathway position, betweenness centrality, and degree. In each case we did not observe a significant pattern of correlation (betweenness: $r_s = -0.35$, $P = 0.06$, Kendall's $\tau = -0.3$, $P = 0.13$; Pathway position: $r_s = 0.4$, $P = 0.08$, Kendall's $\tau = 0.37$, $P = 0.06$; degree: $r_s = -0.47$, $P = 0.06$, Kendall's $\tau = -0.38$, $P = 0.06$).

Next, we developed a metric of contribution within the network we termed the weighted phenotype index (WPI), and compared the total contribution of each group with divergence when organized by pathway position, betweenness, or degree. In the case of pathway position we observed a weak correlation between divergence and contribution ($r = 0.44$), and in the case of betweenness or degree we observed strong negative correlations with contribution and divergence (Figure 2(a) betweenness: $r = -0.92$; Figure 2(b) degree: $r = -0.76$). We also examined the correlation between quantities of betweenness or degree for each node in our network with contribution, and again we observed robust correlations (Figure 2(a) betweenness: $r = 0.9$; Figure 2(b) degree: $r = 0.79$). From this analysis we uncovered patterns of association between measures of betweenness, and degree with evolutionary rate and contribution.

The most highly connected nodes within our network are that of *odr-3* and *egl-4* (also called *pkg-1*). ODR-3 is a $G\alpha$ subunit protein that transduces multiple stimuli within the network [26]. EGL-4 is a protein kinase G that facilitates both primary signal transduction as well as desensitization responses within the network [25, 34]. To

examine how molecular diversity varies across these hub vertices, we performed a sliding window examination of the coding region of each gene using 100-base-pair windows in increments of 20 base pairs. In the case of *egl-4*, we observed only a few polymorphic peaks, with each peak associating with areas between domains of functional importance (Figure 3(a)); these are the N-terminal low-complexity domain (LC), the coiled-coil domain (CC), two cGMP-binding domains (cNMP), and the serine/threonine kinase domain (S/T K). In the case of *odr-3*, we observed more variability overall compared with that of *egl-4*, but again with most variable peaks associating with areas between domains of functional significance (Figure 3(b)); in particular, the five alpha helices (G1–G5) that comprise the GTPase domain, the receptor interaction C-terminal (CT) as well as the fatty acid modification site (M) (Figure 3(b)). We observed similar scales for nucleotide diversity for both *egl-4* ($\pi = 0.27$) and *odr-3* ($\pi = 0.25$) with each gene exhibiting polymorphic peaks in areas between domains of functional importance. The similar pattern of diversity is in keeping with similar trends of divergence between *egl-4* and *odr-3*: *egl-4* $d_N/d_S = 0.0041$, $d_N = 0.1183$, $d_S = 29.1777$; *odr-3* $d_N/d_S = 0.0054$, $d_N = 0.1132$, and $d_S = 21.1204$.

Overall, nodes within the odor network are undergoing purifying selection, with an average d_N/d_S value across the network of 0.012 for *P. pacificus* versus *C. elegans*. To place this value in a context of global divergence rates between *C. elegans* and *P. pacificus* we examined divergence across 5,666 pairs of 1 : 1 orthologs, then generated randomized data sets (50,000 in total) comprising an equal number of nodes as our odor network (i.e., 16), and examined the frequency of mean d_N/d_S values in each case (Figure 4, blue bars). From this analysis we found an average d_N/d_S value = 0.14 which is more than an order of magnitude higher ($\sim 11.6X$) than what we observed for our dynamic odor network. To control for the number of $G\alpha$ subunit proteins within our odor network we also conducted a randomization trial whereby d_N/d_S averages from 50,000 random data sets of 16 nodes that contain 6 $G\alpha$ subunit genes in each case were generated (Figure 4, red bars). In this control experiment we still observed a significantly higher average d_N/d_S value (d_N/d_S average = 0.106) than for our odor network. Taken together, this suggests that the dynamic odor-signaling network is shaped by a more pronounced purifying selection than what guides global constraint across the genome.

4. Discussion

Positional rate variation (PRV) has been tested within numerous biosynthetic pathways where it has been demonstrated that upstream genes (perhaps on account of greater pleiotropy) undergo more selective constraint than downstream genes [42–44]. However, it is unlikely that this effect holds true for all signaling pathways as input and network architecture will vary widely. In the case of sensory signaling pathways that subservise an environment versus perception “arms race,” the idea of upstream genes exhibiting greater diversity may be more applicable. This trend was observed in a study of the *Drosophila* innate immune system

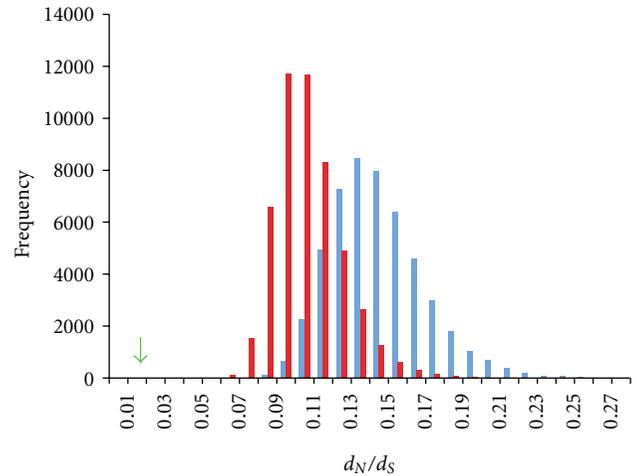


FIGURE 4: Frequency distribution of average d_N/d_S values from 50,000 randomization sets comprising 16 node groupings from 5,666 1 : 1 orthologs between *C. elegans* and *P. pacificus* (blue bars). Red bars indicate frequency distribution of average d_N/d_S values from 50,000 randomization sets of 16 node groupings that include 6 $G\alpha$ subunit proteins in each; this control group was examined to investigate potential bias in our network by having 6 $G\alpha$ subunit proteins within our odor genetic network. The x -axis represents binned divergence categories plotted against the frequency on the y -axis. The green arrow indicates the actual average d_N/d_S value of the nematode odor network.

where the authors demonstrated that recognition genes undergo higher levels of positive selection than immune effector genes [45]. In our network we did not observe a significant correlation between pathway position and divergence, which was in keeping with previous observations from data on intragenus networks of odor signaling in *Caenorhabditis* [46]. However, by comparing measures of degree or centrality with divergence, we did observe significant negative correlations. This may suggest that pathway position is not the only factor in shaping constraint or change in our network, but rather network characteristics such as connectivity may play major roles in guiding molecular evolution. Negative correlations between divergence with node relevance, and positive correlations between essentiality with node relevance have been reported previously and may represent a general trend of many biological networks [47, 48].

By comparing global levels of divergence between *C. elegans* and *P. pacificus* we found a pervasive theme of constraint across the genome; however, through our randomization study (Figure 4) we found that global levels of constraint are an order of magnitude lower than the level of constraint across the odor network. This suggests that the odor-signaling pathway represents a fixed circuit whose network properties are preserved by strong purifying selection. Low levels of divergence have been reported within *Caenorhabditis* for members of the olfactory pathway [46], and previously we have found that a large component in Regulators Class 1 and 2 has undergone extensive nematode-specific gene duplication events, namely, the $G\alpha$ subunit proteins and guanylyl cyclase proteins [49, 50]. Many of

these genes are either exclusively or highly expressed in primary sensory neurons [26, 28, 31, 35, 36, 51]. This genetic expansion facilitates multiple capacities that shape developmental and survival strategies through intercellular and intracellular processing of polymodal sensory input. Many of these duplicates are present in all four major clades of the phylum nematoda [49] suggesting that the odor-signaling network arose early in nematode evolution and has undergone neofunctionalization events that are preserved by strong functional constraint.

Acknowledgments

The authors are very grateful to Rolf Sommer and Christian Rodelsperger for providing them with *Pristionchus pacificus* coding sequence files and to Scott Hamilton for critical reading of this manuscript. They also wish to thank their two anonymous reviewers for their excellent suggestions.

References

- [1] S. Ward, "Chemotaxis by the nematode *Caenorhabditis elegans*: identification of attractants and analysis of the response by use of mutants," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 70, no. 3, pp. 817–821, 1973.
- [2] D. B. Dusenbery, "Countercurrent separation: a new method for studying behavior of small aquatic organisms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 70, no. 5, pp. 1349–1352, 1973.
- [3] C. I. Bargmann and H. R. Horvitz, "Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in *C. elegans*," *Neuron*, vol. 7, no. 5, pp. 729–742, 1991.
- [4] C. I. Bargmann, E. Hartwig, and H. R. Horvitz, "Odorant-selective genes and neurons mediate olfaction in *C. elegans*," *Cell*, vol. 74, no. 3, pp. 515–527, 1993.
- [5] A. O. W. Stretton, R. E. Davis, and J. D. Angstadt, "Neural control of behaviour in *Ascaris*," *Trends in Neurosciences*, vol. 8, no. 6, pp. 294–300, 1985.
- [6] L. A. Perkins, E. M. Hedgecock, J. N. Thomson, and J. G. Culotti, "Mutant sensory cilia in the nematode *Caenorhabditis elegans*," *Developmental Biology*, vol. 117, no. 2, pp. 456–487, 1986.
- [7] S. M. Kroetz, J. Srinivasan, J. Yaghoobian, P. W. Sternberg, and R. L. Hong, "The cGMP signaling pathway affects feeding behavior in the necromenic nematode *Pristionchus pacificus*," *PLoS ONE*, vol. 7, no. 4, Article ID e34464, 2012.
- [8] R. L. Hong, H. Witte, and R. J. Sommer, "Natural variation in *Pristionchus pacificus* insect pheromone attraction involves the protein kinase EGL-4," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 22, pp. 7779–7784, 2008.
- [9] C. Dieterich, S. W. Clifton, L. N. Schuster et al., "The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism," *Nature Genetics*, vol. 40, no. 10, pp. 1193–1198, 2008.
- [10] R. J. Sommer, "*Pristionchus pacificus*," *WormBook*, vol. 14, pp. 1–8, 2006.
- [11] S. Kryazhimskiy and J. B. Plotkin, "The population genetics of dN/dS," *PLoS Genetics*, vol. 4, no. 12, Article ID e1000304, 2008.
- [12] G. Östlund, T. Schmitt, K. Forslund et al., "Inparanoid 7: new algorithms and tools for eukaryotic orthology analysis," *Nucleic Acids Research*, vol. 38, no. 1, pp. D196–D203, 2009.
- [13] L. Li, C. J. Stoeckert Jr., and D. S. Roos, "OrthoMCL: identification of ortholog groups for eukaryotic genomes," *Genome Research*, vol. 13, no. 9, pp. 2178–2189, 2003.
- [14] A. Schneider, C. Dessimoz, and G. H. Gonnet, "OMA Browser—exploring orthologous relations across 352 complete genomes," *Bioinformatics*, vol. 23, no. 16, pp. 2180–2182, 2007.
- [15] K. Yook, T. W. Harris, T. Bieri et al., "WormBase 2012: more genomes, more data, new website," *Nucleic Acids Research*, vol. 40, no. D1, pp. D735–D741, 2012.
- [16] S. F. Altschul, T. L. Madden, A. A. Schäffer et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Research*, vol. 25, no. 17, pp. 3389–3402, 1997.
- [17] R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, vol. 32, no. 5, pp. 1792–1797, 2004.
- [18] S. Guindon and O. Gascuel, "A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood," *Systematic Biology*, vol. 52, no. 5, pp. 696–704, 2003.
- [19] Z. Yang and R. Nielsen, "Estimating synonymous and non-synonymous substitution rates under realistic evolutionary models," *Molecular Biology and Evolution*, vol. 17, no. 1, pp. 32–43, 2000.
- [20] Z. Yang, "PAML: a program package for phylogenetic analysis by maximum likelihood," *Computer Applications in the Biosciences*, vol. 13, no. 5, pp. 555–556, 1997.
- [21] P. Librado and J. Rozas, "DnaSP v5: a software for comprehensive analysis of DNA polymorphism data," *Bioinformatics*, vol. 25, no. 11, pp. 1451–1452, 2009.
- [22] P. Shannon, A. Markiel, O. Ozier et al., "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.
- [23] E. R. Troemel, B. E. Kimmel, and C. I. Bargmann, "Reprogramming chemotaxis responses: sensory neurons define olfactory preferences in *C. elegans*," *Cell*, vol. 91, no. 2, pp. 161–169, 1997.
- [24] H. A. Colbert and C. I. Bargmann, "Odorant-specific adaptation pathways generate olfactory plasticity in *C. elegans*," *Neuron*, vol. 14, no. 4, pp. 803–812, 1995.
- [25] N. D. L'Etoile, C. M. Coburn, J. Eastham, A. Kistler, G. Gallegos, and C. I. Bargmann, "The cyclic GMP-dependent protein kinase EGL-4 regulates olfactory adaptation in *C. elegans*," *Neuron*, vol. 36, no. 6, pp. 1079–1089, 2002.
- [26] K. Roayaie, J. G. Crump, A. Sagasti, and C. I. Bargmann, "The $G\alpha$ protein ODR-3 mediates olfactory and nociceptive function and controls cilium morphogenesis in *C. elegans* olfactory neurons," *Neuron*, vol. 20, no. 1, pp. 55–67, 1998.
- [27] D. M. Ferkey, R. Hyde, G. Haspel et al., "*C. elegans* G protein regulator RGS-3 controls sensitivity to sensory stimuli," *Neuron*, vol. 53, no. 1, pp. 39–52, 2007.
- [28] D. A. Birnby, E. M. Link, J. J. Vowels, H. Tian, P. L. Colacurcio, and J. H. Thomas, "A transmembrane guanylyl cyclase (DAF-11) and Hsp90 (DAF-21) regulate a common set of chemosensory behaviors in *Caenorhabditis elegans*," *Genetics*, vol. 155, no. 1, pp. 85–104, 2000.
- [29] M. Matsuki, H. Kunitomo, and Y. Iino, "Go α regulates olfactory adaptation by antagonizing G q α -DAG signaling in *Caenorhabditis elegans*," *Proceedings of the National Academy*

- of Sciences of the United States of America*, vol. 103, no. 4, pp. 1112–1117, 2006.
- [30] W. M. Nuttley, S. Harbinder, and D. van der Kooy, “Regulation of distinct attractive and aversive mechanisms mediating benzaldehyde chemotaxis in *Caenorhabditis elegans*,” *Learning and Memory*, vol. 8, no. 3, pp. 170–181, 2001.
- [31] N. D. L’Etoile and C. I. Bargmann, “Olfaction and odor discrimination are mediated by the *C. elegans* guanylyl cyclase ODR-1,” *Neuron*, vol. 25, no. 3, pp. 575–586, 2000.
- [32] J. Y. Sze and G. Ruvkun, “Activity of the *Caenorhabditis elegans* UNC-86 POU transcription factor modulates olfactory sensitivity,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 16, pp. 9560–9565, 2003.
- [33] C. M. Coburn and C. I. Bargmann, “A putative cyclic nucleotide-gated channel is required for sensory development and function in *C. elegans*,” *Neuron*, vol. 17, no. 4, pp. 695–706, 1996.
- [34] S. A. Daniels, M. Ailion, J. H. Thomas, and P. Sengupta, “egl-4 Acts through a transforming growth factor- β /SMAD pathway in *Caenorhabditis elegans* to regulate multiple neuronal circuits in response to sensory cues,” *Genetics*, vol. 156, no. 1, pp. 123–141, 2000.
- [35] G. Jansen, K. L. Thijssen, P. Werner, M. van der Horst, E. Hazendonk, and R. H. A. Plasterk, “The complete family of genes encoding G proteins of *Caenorhabditis elegans*,” *Nature Genetics*, vol. 21, no. 4, pp. 414–419, 1999.
- [36] H. Lans, S. Rademakers, and G. Jansen, “A network of stimulatory and inhibitory G α -subunits regulates olfaction in *Caenorhabditis elegans*,” *Genetics*, vol. 167, no. 4, pp. 1677–1687, 2004.
- [37] A. Palmitessa, H. A. Hess, I. A. Bany, Y. M. Kim, M. R. Koelle, and J. L. Benovic, “*Caenorhabditis elegans* arrestin regulates neural G protein signaling and olfactory adaptation and recovery,” *The Journal of Biological Chemistry*, vol. 280, no. 26, pp. 24649–24662, 2005.
- [38] A. Kuhara, H. Inada, I. Katsura, and I. Mori, “Negative regulation and gain control of sensory neurons by the *C. elegans* calcineurin TAX-6,” *Neuron*, vol. 33, no. 5, pp. 751–763, 2002.
- [39] H. Komatsu, I. Mori, J. S. Rhee, N. Akaike, and Y. Ohshima, “Mutations in a cyclic nucleotide-gated channel lead to abnormal thermosensation and chemosensation in *C. elegans*,” *Neuron*, vol. 17, no. 4, pp. 707–718, 1996.
- [40] K. G. Miller, M. D. Emerson, and J. B. Rand, “G(o) α and diacylglycerol kinase negatively regulate the G(q) α pathway in *C. elegans*,” *Neuron*, vol. 24, no. 2, pp. 323–333, 1999.
- [41] R. Y. N. Lee, E. R. Sawin, M. Chalfie, H. R. Horvitz, and L. Avery, “EAT-4, a homolog of a mammalian sodium-dependent inorganic phosphate cotransporter, is necessary for glutamatergic neurotransmission in *Caenorhabditis elegans*,” *Journal of Neuroscience*, vol. 19, no. 1, pp. 159–167, 1999.
- [42] M. D. Rausher, R. E. Miller, and P. Tiffin, “Patterns of evolutionary rate variation among genes of the anthocyanin biosynthetic pathway,” *Molecular Biology and Evolution*, vol. 16, no. 2, pp. 266–274, 1999.
- [43] H. Ramsay, L. H. Rieseberg, and K. Ritland, “The correlation of evolutionary rate with pathway position in plant terpenoid biosynthesis,” *Molecular Biology and Evolution*, vol. 26, no. 5, pp. 1045–1053, 2009.
- [44] D. Waxman and J. R. Peck, “Pleiotropy and the preservation of perfection,” *Science*, vol. 279, no. 5354, pp. 1210–1213, 1998.
- [45] T. B. Sackton, B. P. Lazzaro, T. A. Schlenke, J. D. Evans, D. Hultmark, and A. G. Clark, “Dynamic evolution of the innate immune system in *Drosophila*,” *Nature Genetics*, vol. 39, no. 12, pp. 1461–1468, 2007.
- [46] R. Jovelin, J. P. Dunham, F. S. Sung, and P. C. Phillips, “High nucleotide divergence in developmental regulatory genes contrasts with the structural elements of olfactory pathways in *Caenorhabditis*,” *Genetics*, vol. 181, no. 4, pp. 1387–1397, 2009.
- [47] S. Wuchty, “Evolution and topology in the yeast protein interaction network,” *Genome Research*, vol. 14, no. 7, pp. 1310–1314, 2004.
- [48] H. B. Fraser, A. E. Hirsh, L. M. Steinmetz, C. Scharfe, and M. W. Feldman, “Evolutionary rate in the protein interaction network,” *Science*, vol. 296, no. 5568, pp. 750–752, 2002.
- [49] D. M. O’Halloran, D. A. Fitzpatrick, G. P. McCormack, J. O. McInerney, and A. M. Burnell, “The molecular phylogeny of a nematode-specific clade of heterotrimeric G-protein α -subunit genes,” *Journal of Molecular Evolution*, vol. 63, no. 1, pp. 87–94, 2006.
- [50] D. A. Fitzpatrick, D. M. O’Halloran, and A. M. Burnell, “Multiple lineage specific expansions within the guanylyl cyclase gene family,” *BMC Evolutionary Biology*, vol. 6, article 26, 2006.
- [51] S. Yu, L. Avery, E. Baude, and D. L. Garbers, “Guanylyl cyclase expression in specific sensory neurons: a new family of chemosensory receptors,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 7, pp. 3384–3387, 1997.