Polyphenism of a Novel Trait Integrated Rapidly Evolving Genes into Ancestrally Plastic Networks

Sofia Casasa [1], Joseph F. Biddle,[1] Georgios D. Koutsovoulos,[2] and Erik J. Ragsdale *[,1]

[1]Department of Biology, Indiana University, Bloomington, Bloomington, IN

[2]Institute National de la Recherche Agronomique, Sophia-Antipolis, PACA, France

*Corresponding author: E-mail: ragsdale@indiana.edu.

Associate editor: Ilya Ruvinsky

Abstract

Developmental polyphenism, the ability to switch between phenotypes in response to environmental variation, involves the alternating activation of environmentally sensitive genes. Consequently, to understand how a polyphenic response evolves requires a comparative analysis of the components that make up environmentally sensitive networks. Here, we inferred coexpression networks for a morphological polyphenism, the feeding-structure dimorphism of the nematode *Pristionchus pacificus*. In this species, individuals produce alternative forms of a novel trait—moveable teeth, which in one morph enable predatory feeding—in response to environmental cues. To identify the origins of polyphenism network components, we independently inferred coexpression modules for more conserved transcriptional responses, including in an ancestrally nonpolyphenic nematode species. Further, through genome-wide analyses of these components across the nematode family (Diplogastridae) in which the polyphenism arose, we reconstructed how network components have changed. To achieve this, we assembled and resolved the phylogenetic context for five genomes of species representing the breadth of Diplogastridae and a hypothesized outgroup. We found that gene networks instructing alternative forms arose from ancestral plastic responses to environment, specifically starvation-induced metabolism and the formation of a conserved diapause (dauer) stage. Moreover, loci from rapidly evolving gene families were integrated into these networks with higher connectivity than throughout the rest of the *P. pacificus* transcriptome. In summary, we show that the modular regulatory outputs of a polyphenic response evolved through the integration of conserved plastic responses into networks with genes of high evolutionary turnover.

Key words: **coexpression network, developmental plasticity, modularity, nematodes, phylogenomics, taxon-restricted genes.**

Introduction

Developmental plasticity shows the profound influence that environment can have on an ultimate phenotype, yet plasticity may itself be defined by a genetic program with its own selection pressures (Bradshaw 1965; Scheiner 1993). In the most extreme cases, plasticity results in discontinuous phenotypes, or polyphenism, likely through the alternating activation of "environmentally sensitive loci" (Via et al. 1995; Nijhout 2003; West-Eberhard 2003; Lafuente and Beldade 2019). To produce discrete phenotypic outputs, the expression of these loci is likely to be tightly organized, such that polyphenisms likely require the "on–off," step response of a regulatory network (Abouheif and Wray 2002). Further, changes to the connectivity of plastic networks, whether by rewiring or through turnover of the networks' components, have the potential to drastically change their environmental responsiveness or even outputs (West-Eberhard 2005; Aubin-Horth and Renn 2009; Schneider and Meyer 2017). Therefore, to understand how polyphenic responses and their outputs adapt, diverge, or become assimilated, it is essential to identify such networks, their evolutionary origins, and how their components change.

A critical advance toward characterizing gene regulatory networks (GRNs) has been the discovery of tangible "switch genes," transcriptional switches that activate GRNs (Davidson 2006). In some cases, developmental switches have been characterized for environmental responses, both systemic (Ogg et al. 1997; Antebi et al. 1998; Xu et al. 2015) and localized to a particular morphological trait (Ragsdale et al. 2013). Given their pervasive influence over downstream GRNs, switch-like regulators can inform how environmentally sensitive networks are modified and thus how plasticity evolves (Masel and Siegal 2009; Sommer 2020). First, a mechanistic understanding of switch genes allows us to determine how loci become annexed to an environmentally sensitive network under the control of a switch. For example, loci may even enter plasticity regulation from foreign genomes, requiring their de novo network integration downstream of plasticity modifiers (Parker and Brisson 2019). Second, laboratory manipulation of identified switch genes provides a way of delineating the downstream genes directly responsible for polyphenism, in contrast to the broader swaths of loci affected by a generic environmental response (Bui and

Ragsdale 2019). Therefore, a comparison of environmentally responsive GRNs that includes the full history of a polyphenism, as well as outgroups without polyphenism, would allow the reconstruction of how molecular targets of a plastic response have evolved to instruct alternative paths to development.

An exemplary system for making this comparison is nematodes of the family Diplogastridae, which exhibit polyphenism in their feeding morphologies. This polyphenism is best understood in *Pristionchus pacificus*, which develops into either a microbivorous ("stenostomatous," St) or a facultatively predatory ("eurystomatous," Eu) morph depending in part on population densities and the availability of bacterial food (Bento et al. 2010; Serobyan et al. 2013, 2014). Further, the polyphenism controls a morphology—moveable teeth—that like predatory feeding is a novelty of Diplogastridae. Importantly for the study of polyphenism GRNs, a switch-like genetic mechanism regulating these alternative morphs has been characterized (fig. 1): A series of enzymes with environmental sensitivity alternatively toggle the polyphenism response (Ragsdale et al. 2013; Ragsdale and Ivers 2016; Bui et al. 2018; Namdeo et al. 2018; Sieriebriennikov et al. 2018), which is ultimately controlled by two nuclear receptors, NHR-40 and NHR-1 (Kieninger et al. 2016; Sieriebriennikov et al. 2020). By perturbing multiple regulators, specifically two of the above enzymes and NHR-40, the ultimate targets of the polyphenism switch have been identified (Bui and Ragsdale 2019). Further, the targets shared by NHR-40 with NHR-1, a receptor that may act downstream of the global polyphenism switch to affect mouth morphology itself, have been likewise determined (Sieriebriennikov et al. 2020). Thus, work in this system has revealed overlapping suites of loci that putatively contribute to plastic GRNs, enabling the networks' identification. An additional feature of this system is the phylogenetic context it makes possible. Teeth and their polyphenism evolved once in an early ancestor of Diplogastridae, after which a radiation of plastic responses and mouth morphologies ensued (Susoy et al. 2015). Consequently, the histories of GRN components can be compared across a rich diversity of polyphenic lineages to advance our understanding of polyphenic GRNs and their evolution.

Here, we combined inference of gene coexpression networks with phylogenomic analysis to infer the genome-wide mechanisms of polyphenism, their origin, and the evolutionary history of their components. First, we describe the genetic basis of the *P. pacificus* polyphenism in a network context, identifying several morph-specific coexpression modules, some of which distinguish previously predicted subsets of tissue- and process-specific transcripts. Second, we assessed the extent to which these components have been co-opted from conserved processes. By independently inferring coexpression modules in a species (*Caenorhabditis elegans*) without mouth polyphenism in its evolutionary history and then comparing it to *P. pacificus*, we identified common network components, specifically those from circuits controlling the metabolic response to starvation. Similarly, we identified common network components from circuits regulating diapause in *P. pacificus*. Third, by drawing upon the phylogenetic diversity of polyphenic species in this system, we reconstructed the genome-wide history of components making up the mouth-polyphenism expression network in *P. pacificus*. To achieve this, our analyses included five new genomes representing the phylogenetic breadth of Diplogastridae and a newly resolved outgroup. Through this approach, we rooted the history of polyphenism network components and identified a role for gene family expansion and contraction in the evolution of polyphenism gene networks. In summary, we present the evolutionary history of network components instructing a novel, polyphenic trait, and we provide evidence that polyphenism gene networks evolve by integrating genes with high evolutionary turnover into deeply conserved, environmentally responsive gene circuits.

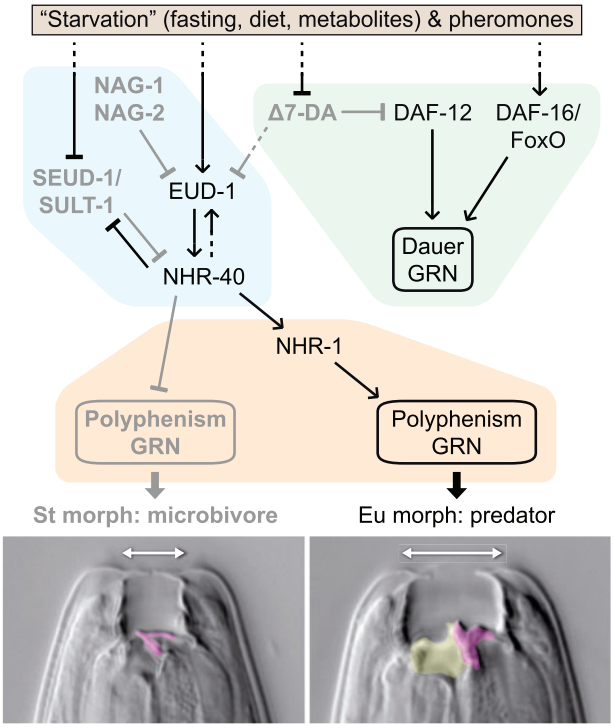

Fig. 1. Putative model for switch-mediated regulation of environmentally sensitive genes in *Pristionchus pacificus*. Two feeding-structure morphs (Eu, St) produce mouthparts that differ in their gape (arrows) and in the shape and number of their teeth (dorsal tooth of both morphs, false-colored pink; subventral tooth restricted to Eu morph, yellow), enabling different feeding strategies in response to local environmental signals (brown box). Similar, although not identical, signals likewise influence the decision to enter dauer diapause (green box) through processes with limited overlap with the mouth polyphenism. Specifically, the mouth-polyphenism switch (blue box) comprises a series of enzymes (the α-*N*-acetylglucosaminidases NAG-1 and NAG-2, sulfatase EUD-1, and sulfotransferase SEUD-1/SULT-1) that alternatively influence the activity of two nuclear receptors, NHR-40 and NHR-1. These receptors together control the development of the alternative forms, presumably through the regulation of polyphenism-specific GRNs (orange box).

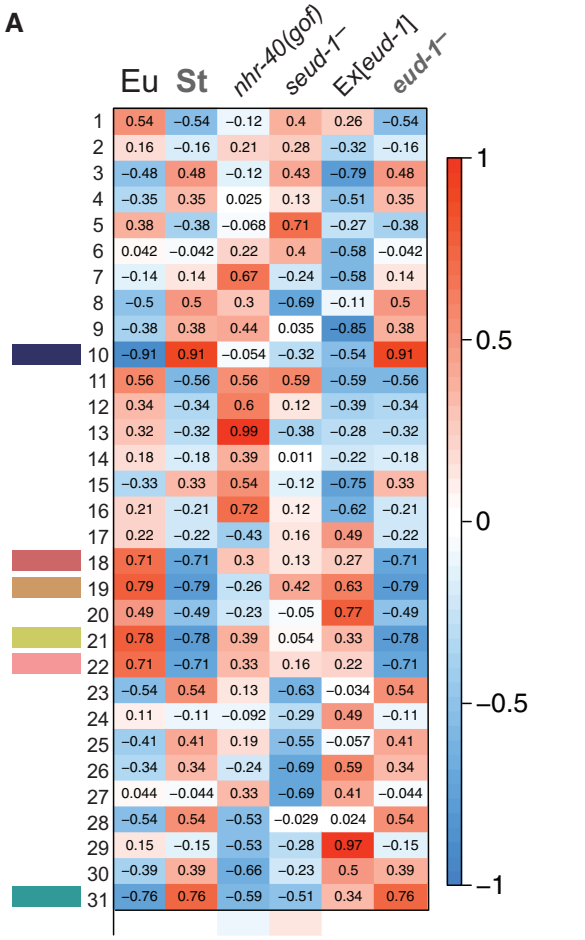

FIG. 2. Gene coexpression networks of the mouth polyphenism of *Pristionchus pacificus*. (*A*) Heat map of module–trait correlation, displaying the correlation values for each module (rows). Positive module correlation per trait (columns) is displayed in red, whereas negative is displayed in blue. Color swatches to left indicate polyphenism modules. Eu-correlated modules: 18, 19, 21, 22. St-correlated modules: 10, 31. (*B*) Visual representation of the six polyphenism modules ($P < 0.01$). Colors refer to modules in (*A*).

Results and Discussion

Coexpression Modules Organize Alternatively Activated Polyphenism Target Genes

Given the large-scale transcriptomic responses to polyphenism induction cues, it is likely that a polyphenism switch specifies traits both directly and indirectly through its molecular targets. Therefore, we first used *P. pacificus* to describe the response of a morphological polyphenism in terms of gene expression networks. Specifically, we used weighted gene coexpression network analysis (WGCNA; Langfelder and Horvath 2008) to reconstruct polyphenism networks. Because WGCNA can predict functionally relevant correlations and empirically validated GRNs (e.g., Voineagu et al., 2011; Xue et al. 2013; Walley et al. 2016), we took this approach to infer regulatory connectivity among environmentally sensitive genes. For this analysis, we used the raw transcriptomes of mutants for multiple polyphenism switch-genes (*eud-1*, *seud-1*/*sult-1*, *nhr-40* gain-of-function, and *eud-1* overexpression). From these transcripts we inferred 31 coexpression modules, six of which were polyphenism specific (fig. 2; supplementary data 1, Supplementary Material online): Four had a significant positive correlation with the predatory (Eu) morph ($P < 0.01$), whereas two had a significant negative correlation with that morph (i.e., were St-biased; $P < 0.01$). Among Eu-biased modules, one was enriched for nematode astacin (*nas*) genes, including genes differentially expressed by both *nhr-40* and *nhr-1* mutants in *P. pacificus* (Sieriebriennikov et al. 2020). Another Eu-biased module was enriched with heat shock proteins, consistent with another study that found Hsp90 to play an important role in variability of mouth morphologies (Sieriebriennikov et al. 2017). Further, one of the two St-biased modules was enriched with fatty-acid metabolism genes, recovering a putative network that was previously found through an independent approach, pathway analysis (Bui and Ragsdale 2019). Our approach thus reconstructed expression modules that are controlled by a polyphenism switch and that organize previously predicted subsets of transcripts into a putative network.

Because at least some polyphenism-specific genes are known to be expressed in only a single cell (i.e., astacin genes in the dorsal pharyngeal gland), it is possible that our inferred modules have drawn from tissue-specific networks in the polyphenism response. To explore this possibility, we determined the location of polyphenism-specific gene-expression

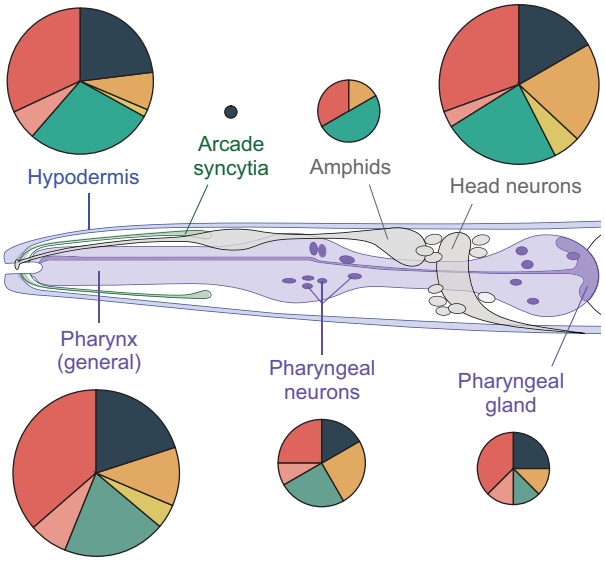

Fig. 3. Anatomical expression of homologs of *Pristionchus pacificus* mouth-polyphenism genes in the nonpolyphenic species *Caenorhabditis elegans*. Idealized nematode diagram is colored by tissue that produces dimorphic mouthparts in *P. pacificus* (purple hues, pharynx; green, arcade syncytia; blue, hypodermis/epidermis) and also includes other anterior tissues possibly involved in the polyphenism (gray). Pie charts show relative representation among six polyphenism coexpression modules (colors correspond to fig. 1*B* and *C*), with size of chart indicating relative contribution by all polyphenism modules.

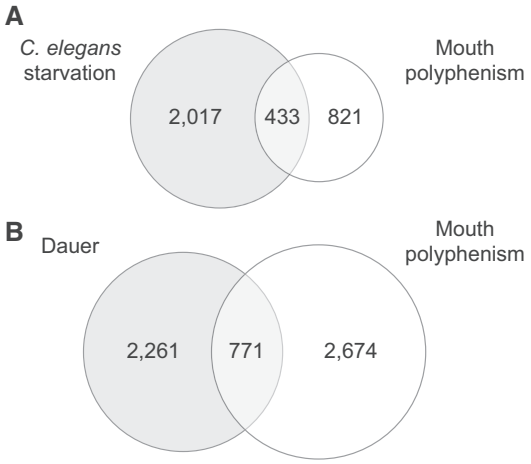

Fig. 4. Overlap of genes coexpressed in the mouth-polyphenism with environmentally sensitive loci from more ancient processes. (*A*) Coexpressed genes in the *Pristionchus pacificus* mouth polyphenism with putative homologs whose expression is induced by starvation in *Caenorhabditis elegans*. Only genes with putative homologs in both species are counted. (*B*) Coexpressed genes both in the *P. pacificus* mouth polyphenism and during *P. pacificus* dauer diapause.

in the more widely studied model *C. elegans*, a nematode whose ancestors never had the mouth polyphenism. Our search revealed that some polyphenism-specific genes are present in *C. elegans* tissues homologous with those that form polyphenic mouthparts in Diplogastridae (fig. 3). Moreover, tissue-specific expression was biased depending on module identity, such that not all modules were distributed equally across all tissues. Together, these results support two predictions. First, polyphenism development likely recruited at least some genes already present in the genome and expressed at the site of the polyphenism prior to its evolution. Second, they prompt the testable prediction that at least some genes in these modules are, as in *C. elegans*, expressed in *P. pacificus* different cell classes, which span tissues of multiple embryonic origins (Portereiko and Mango 2001; Vangestel et al. 2008) and contribute in parallel to the development of dimorphic mouthparts.

Conserved Starvation-Response Genes Are Integrated with Polyphenism Gene Networks

Having inferred polyphenism-gene expression modules, we next sought to identify their origins, particularly in ancestors before polyphenism arose. Given that the *P. pacificus* mouthpart polyphenism is in part induced by low nutrition, we hypothesized that the polyphenism may have co-opted genes ancestrally involved in a conserved starvation response. To test this hypothesis, we identified starvation-induced coexpression modules for comparison, specifically in a species outside of Diplogastridae—again, *C. elegans*—with no mouthpart polyphenism in its evolutionary history. In particular, we analyzed raw RNA-Seq data that were collected from juveniles (mid-fourth stage larvae) of this species after exposing them to food deprivation for multiple durations (Harvald et al. 2017). From the entire data set, transcripts were assigned to 12 coexpression modules, of which a single one had a positive correlation ($P < 0.01$) and one, a negative correlation ($P < 0.01$) with the starvation response (supplementary fig. S1, Supplementary Material online). We then compared genes making up the *P. pacificus* polyphenism expression network with those in *C. elegans* starvation-response modules. We identified 433 homologs that belonged to both starvation- and polyphenism-correlated modules in *C. elegans* and *P. pacificus*, respectively (fig. 4*A*). These genes formed parts of all polyphenism modules, and they comprised 34.5% of polyphenism genes with homologs in *C. elegans*, a proportion higher than expected by chance ($\chi^2 = 4.941$, df = 1, $P < 0.05$). Thus, polyphenism expression modules are significantly enriched with homologs of starvation-activated genes. Because genes in *P. pacificus* morph-biased networks were identified using mutants for polyphenism switch genes in a well-fed background (i.e., without starvation as an induction cue), we conclude that starvation-associated genes were regulated by the polyphenism switch itself. In summary, this finding indicates that these homologs had their origins in network modules ancestral to the mouth polyphenism and that these modules were repurposed to follow a polyphenism switch.

Dauer Development Pathways Were Co-opted Downstream of a Derived Polyphenism Switch

The developmentally arrested, dauer diapause stage of *P. pacificus* is induced by a combination of pheromones and starvation, which coincide with cues required for mouth

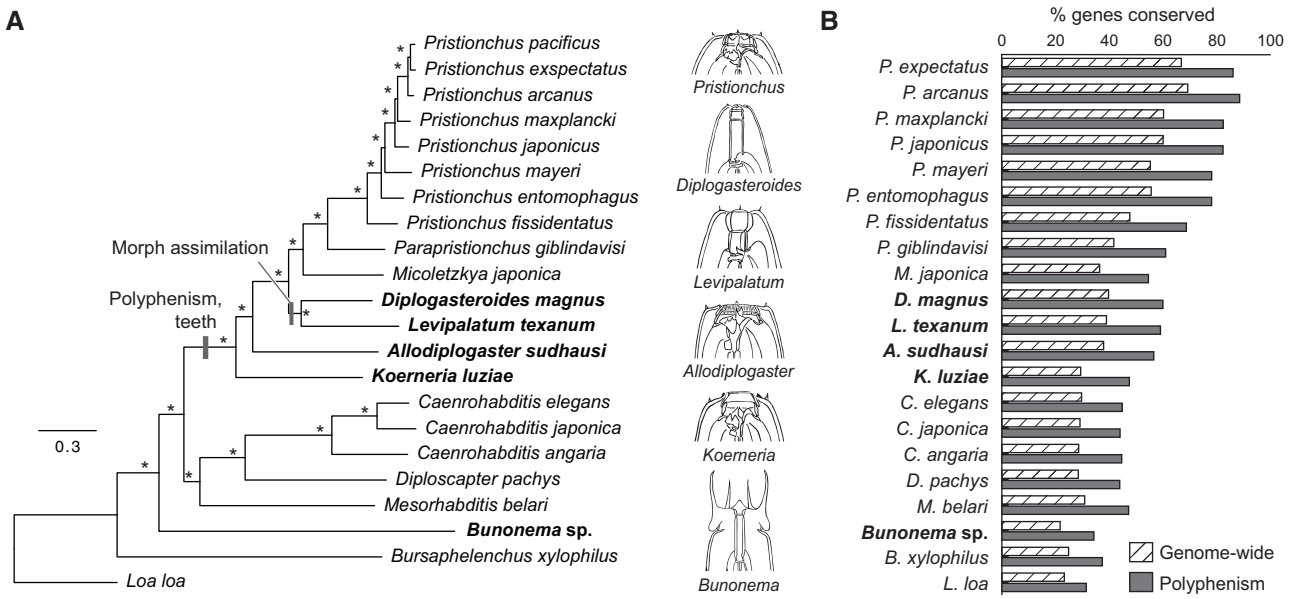

FIG. 5. Phylogenetic context for Diplogastridae and genome-wide comparisons of taxon-restricted genes. Boldface font indicates species with original genome sequences. (*A*) Species tree for Diplogastridae inferred from 778 genome-wide orthologs. Diagrams to right show mouthpart diversity of *Pristionchus* and genera with species sequenced here, including nematodes with polyphenism (Eu morph shown) and those with a single assimilated morph. In contrast, *Bunonema* has relatively simple internal mouthparts, which are also the ancestral state for the sister group of Diplogastridae (Rhabditidae). Asterisks indicate 100% node support (both BS and, for tree inferred under a coalescent model, LPP). Representations of diplogastrid genera after Susoy et al. (2015). (*B*) Numbers of strict orthologs (as defined by reciprocal best BLAST), both genome-wide and of genes in *P. pacificus* polyphenism coexpression networks, conserved between *P. pacificus* and other species.

polyphenism (Ogawa et al. 2009; Bose et al. 2012). However, dauer diapause and mouth polyphenism only partially overlap in their regulatory mechanisms (Bento et al. 2010; Ogawa et al. 2011; fig 1), leaving it unclear as to what extent dauer gene-expression networks were also co-opted in the evolution of polyphenism. To address this question, we analyzed a different set of transcriptomes from *P. pacificus*, namely those for developmental life-stages (Baskaran et al. 2015), to infer an independent set of gene coexpression networks, including those associated with the dauer stage. Transcripts in this data set were assigned to 31 coexpression modules, of which three had a positive correlation with dauer response ($P < 0.01$) and two had a negative correlation ($P < 0.01$; supplementary fig. S2, Supplementary Material online). We then asked whether dauer development genes were more likely to contribute to polyphenism development than expected by chance. We found that 22.4% of polyphenism-specific genes, from across all modules, were also genes that regulate *P. pacificus* dauer development, strongly supporting our hypothesis ($\chi^2 = 110.5$, df = 1, $P < 1 \times 10^{-8}$; fig. 4*B*). Because dauer development is an older, more widespread feature of nematodes than mouth polyphenism, it is likely that conserved dauer development modules were co-opted in the evolution of the mouth polyphenism. Further, their co-option may have been independent of upstream dafachronic acid (Δ7-DA) signaling, which also influences the mouth polyphenism (Bento et al. 2010), given that the modules we identified are controlled by the mouth-polyphenism switch and not directly by DAF-12, the known receptor of Δ7-DA. Together with our findings on starvation networks above, these results suggest that the evolution of polyphenism depended on the co-option of genes anciently involved in environmentally sensitive processes.

New Nematode Genomes Root the Origins of Genes in Polyphenism Networks

Despite the involvement of conserved processes, the mouth polyphenism is a derived developmental feature controlling a novel morphological trait. Therefore, we hypothesized that polyphenism expression modules also must have been built from more than repurposed expression networks. To test this hypothesis, we established a comparative framework for Diplogastridae, using original genome sequences for several nematode species, to track changes among genes in these modules throughout polyphenism evolution. First, we sequenced the genomes of four diplogastrid species: two (*Koerneria luziae*, *Allodiplogaster sudhausi*) representing the deepest known splits between lineages with mouth polyphenism relative to *P. pacificus*, and two (*Diplogasteroides magnus*, *Levipalatum texanum*) belonging to a lineage that secondarily lost the polyphenism (Susoy et al. 2015). We also sequenced *Bunonema* sp. (strain RGD898), a potential but previously unresolved outgroup to either Diplogastridae or a more inclusive clade (Kiontke et al. 2007; van Megen et al. 2009; Susoy et al. 2015). From these sequences, we inferred a species phylogeny using genome-wide orthologs (778 genes) across these species, other Diplogastridae, "Rhabditidae" (i.e., Eurhabditis sensu Sudhaus 2011, which includes *C. elegans*), and two outgroup species (fig. 5*A*; supplementary fig. S3; supplementary data 4, Supplementary Material online). Our inference, made from both 1) a concatenated matrix

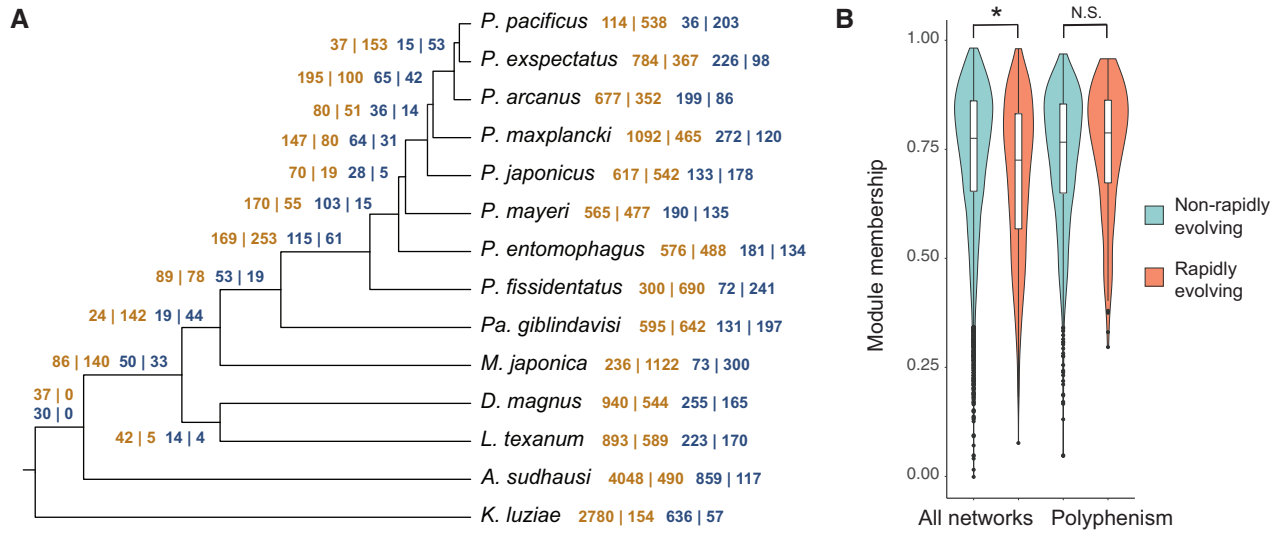

FIG. 6. Gene-family evolution of loci contributing to coexpression networks of the *Pristionchus pacificus* mouth polyphenism. (*A*) Gene-family evolution across Diplogastridae. Values show numbers of gene-family expansions (left of each vertical bar) and contractions (right of bar) next to each internal branch and, for terminal branches, to right of taxon name. Yellow values, genome-wide events; blue values, events including homologs of polyphenism genes. (*B*) Network connectivity, measured as module membership (*MM*), of genes in all coexpression networks and networks specific to the mouth polyphenism.

and 2) under a multispecies coalescent model, resulted in two principal findings. First, it fully supported (bootstrap support, BS = 100%; local posterior probability, LPP = 100%) all inferred relationships within Diplogastridae, congruent with previous inferences from more limited data. Second, it completely resolved (BS = 100%; LPP = 100%) *Bunonema* as an outgroup to Diplogastridae and sequenced Rhabditidae.

In parallel to our phylogenetic study, we also used the new genomes to identify homologs of genes belonging to polyphenism expression networks, such that we could infer the origins of these genes relative to the appearance of polyphenism in Diplogastridae. We found a consistent decrease of genome-wide homologs with increasing phylogenetic distance, whether or not they belong to polyphenism networks (fig. 5*B*). Additionally, the number of recovered orthologs shared with *P. pacificus* tended to be higher for polyphenism networks than for genes throughout the genome, consistent with the overrepresentation of conserved-process genes in the polyphenism. To directly test whether taxon-restricted genes disproportionately contributed to the evolution of polyphenism, we identified which genes in *P. pacificus* were diplogastrid-restricted, as defined by their absence in both Rhabditidae and *Bunonema*. We found the number of diplogastrid-restricted, polyphenism-specific genes to be lower (19.6%) than the proportion of polyphenism-specific genes throughout the genome (23.2%; $\chi^2 = 18.959$, df = 1, $P < 10^{-4}$). This result suggests that taxon-restricted genes were not pervasive in forming polyphenism networks, consistent with our finding above that the repurposing and rewiring of previously existing networks played a major part in the origin of polyphenism-biased gene expression. Finally, we found that only 23 homologs of *P. pacificus* polyphenism-specific genes were lost in the ancestor exclusive to *L. texanum* + *D. magnus*, showing no significant pattern associated with polyphenism loss. Together, our results provide a first comparative framework for diplogastrid-outgroup contrasts, including the origins of genes in polyphenism expression networks.

Polyphenism Networks Are Overrepresented by Rapidly Evolving Gene Families

Using genomes across Diplogastridae and outgroup species, we then sought to determine whether genome-wide patterns, specifically those of gene-family contraction and expansion, have contributed to the regulation of a novel, plastic trait. We wanted to distinguish between several scenarios possibly explaining the evolution of polyphenism expression networks: Did polyphenism-targeted networks simply evolve through the rewiring of networks from conserved, environmentally sensitive processes? Or did polyphenism evolve through integrating genes from conserved processes, as described above, into new networks including genes that are rapidly evolving? Alternatively, does the rapid turnover of conserved-process genes themselves characterize polyphenism-specific networks? To discern among these possibilities, we first identified which polyphenism-specific transcripts belonged to rapidly changing gene families (fig. 6*A*). Of the polyphenism transcripts, 145 (12.62%) fell within this category, proportionally more than the 7.27% in this category among genome-wide transcripts ($\chi^2 = 55.126$, df = 1, $P < 1 \times 10^{-12}$). Further, we found that the proportion of polyphenism genes that belong to rapidly evolving families was not enriched with starvation genes ($\chi^2 = 0.266$, df = 1, $P = 0.606$) or dauer genes ($\chi^2 = 2.874$, df = 1, $P = 0.09$), suggesting that genes with high turnover make up parts of the network specific to novel mouthpart development per se. In summary,

our results show that polyphenism gene networks repurposed otherwise conserved, environmentally sensitive processes by connecting them to genes involved in rapid gene-family evolution.

Highly Connected Genes in Polyphenism Networks Undergo Rapid Turnover

To understand how polyphenism gene networks became integrated in terms of connectivity, we investigated whether the speed of gene family evolution was related to position within a module. Theory from genome-wide patterns predicts that peripheral genes are generally less constrained to evolve than their highly connected counterparts (Hahn and Kern 2005). At the same time, gene duplications fundamentally allow daughter genes to take on replacement functions while inheriting high connectivity (Wagner 2011). Thus, we tested whether genes from rapidly evolving families were in a more peripheral position in the polyphenism gene network compared with those in nonrapidly evolving families. We used module membership (*MM*, also known as k_{ME}), which measures correlation between gene expression and module eigengene, or the first principal component of variation (Langfelder and Horvath 2008), thereby quantifying how close a gene is to a module. This measure can be used to identify highly connected genes and therefore determine the gene's position in the networks, that is, peripheral versus highly connected among module members (Filteau et al. 2013). As predicted, genes in rapidly evolving families had lower *MM* than those in nonrapidly evolving families when considering all *P. pacificus* gene networks ($t = -8.163$, $df = 966.9$, $P < 1 \times 10^{-14}$; fig. 6*B*). However, this difference did not hold for polyphenism genes ($t = 1.644$, $df = 270.1$, $P = 0.101$). Instead, polyphenism genes in rapidly changing families had an *MM* value similar to those genes in nonrapidly evolving families (rapid, 0.766; nonrapid, 0.748). This finding indicates that, despite predictions from genome-wide patterns, polyphenism gene networks comprising otherwise conserved components show strong integration of genes with dynamic turnover.

Evolution of Polyphenism by Network Integration of Conserved and Rapidly Evolving Targets

Molecular mechanisms of developmental polyphenism have been revealed for numerous animal systems, especially in terms of its genome-wide effects (Beldade et al. 2011; Brisson and Davis 2016; Projecto-Garcia et al. 2017; Lafuente and Beldade 2019; Yang and Pospisilik 2019). Fewer studies have examined polyphenism mechanisms in a comparative framework, although such studies have provided insight into how its regulators are born and ultimately work together to achieve polyphenism (Sieriebriennikov et al. 2018; Bhardwaj et al. 2020; Biddle and Ragsdale 2020). As such, the evolutionary rooting of the rest of polyphenism's molecular "basis" can distinguish between conserved and novel components that contribute to both environmental sensitivity and the polyphenic morphologies themselves. Here, we have investigated the molecular origins and evolution of developmental polyphenism—specifically a novel trait, nematode teeth—through a comparative analysis of its regulatory targets. We provide evidence that pathways regulating two major environmentally sensitive processes, specifically starvation and dauer responses, were repurposed to form part of the mouth polyphenism expression network. Further, we found that genes with rapid turnover throughout the history of the polyphenism were, when compared with the rest of the genome, highly integrated within this network. Together, our findings explain how the regulatory outputs of a polyphenism switch arise.

Molecular studies of nutritionally sensitive polyphenisms have indicated that at least some of their components may be generalizable features. For example, nutritionally sensitive polyphenisms show the pervasive deployment of insulin/insulin growth-like factor signaling across animal phyla (Kimura et al. 1997; Emlen et al. 2012; Xu et al. 2015; Casasa and Moczek 2018). Nevertheless, the particular roles played by this shared signaling pathway became cooperative with other developmental processes through independent evolutionary events (Nijhout and McKenna 2018). An understanding of how new uses evolve from more widely conserved nutritional responses can be informed by the phylogenetic histories of their components. Here, we identify components from nutritionally responsive networks shared between two taxa (*C. elegans* and *P. pacificus*) that differ in the presence of a derived polyphenism (mouthpart dimorphism). Importantly, these networks do not show a strict one-to-one correspondence but suggest that rearrangements between the two lineages have taken place. Further, we find that network composition has changed through high rates of turnover in genes at nodes of high connectivity. Thus, a novel environmental response—in this case, the nutritional sensitivity of a morphological trait—does not require novel components any more than expected by chance, but it can rather repurpose expression networks through dynamic turnover of members with preexisting connections.

Another implication of our findings is for the evolution of polyphenism from continuous plasticity. It is conceivable that transcriptional responses for continuous plasticity, which also show modularity and distinct GRN structure (Aubin-Horth et al. 2005; Schneider et al. 2014; Kenkel and Matz 2017), often serve as the substrate for multimodal plasticity (Wagner et al. 2007; Jones and Robinson 2018). This may be through channeling responses through discontinuous environments (Brakefield and Reitsma 1991; Nijhout 2003), steepening the threshold between alternatives through hormone sensitivity (Emlen and Nijhout 1999; van Bergen et al. 2017), or possibly the counteraction of antagonistic signaling pathways (Kijimoto and Moczek 2016). Alternatively, it is possible that a preexisting switch can be repurposed for a new trait. Unlike groups whose species differ in their discontinuity of plasticity, such as horned beetles or butterflies with eyespots (Brakefield et al. 1996; Emlen et al. 2005), the common ancestor of Diplogastridae likely had mouth dimorphism with only rare intermediates (Susoy et al. 2015). Our results suggest that an early switch mechanism may have borrowed from GRNs already responsive to a hard switch. Like the mouth polyphenism, dauer induction is a strictly binary decision (Hu

2007), the molecular basis of which is at least in part conserved between *C. elegans* and *P. pacificus* (Ogawa et al. 2009, 2011). Thus, it is possible that dauer-controlled GRNs were either used in the evolutionary origin of the polyphenism or were a later genetic accommodation of polyphenism initially achieved through another induction mechanism. Ancestral state reconstruction of the polyphenism switch in the ancestor of polyphenic nematodes, such as by genetic perturbations in the species we present here, can ultimately test these possibilities.

In addition to rooting the history of environmentally sensitive loci in the *P. pacificus* mouth polyphenism, our comparative analysis has enabled us to investigate the signatures of genetic assimilation (loss of plasticity) in a lineage that has lost the polyphenism (including two of the species sequenced). Specifically, we have tracked the turnover of homologs of loci in coexpression networks as inferred for *P. pacificus*. Although the initial trigger for assimilation could have been a deleterious mutation in a single locus of major effect, such as in a switch gene, changes to loci that originally modified that plasticity are additionally expected to complete the assimilation (Schlichting and Pigliucci 1993; Lande 2009). Because these changes should involve interactions among large regulatory networks, genome-wide signatures following a trait's assimilation may distinguish these effects (Renn and Schumer 2013; Ehrenreich and Pfennig 2016). In our study of two species with a common ancestor which assimilated a single form (*D. magnus*, *L. texanum*), we found no indication of genome-wide gene loss, namely of homologs of polyphenism-network components, following the assimilation of a single form. This result may be explained in three ways. First, it is possible that the polyphenism GRN was fundamentally different between *P. pacificus* and that other ancestor that lost polyphenism. In this case, ancestral state reconstruction of polyphenism GRNs themselves, additionally using extant representatives (e.g., *K. luziae*, *A. sudhausi*) could test this idea. Alternatively, the relatively few components of the polyphenism GRN that were lost in assimilated species possibly include conserved, major effectors specific to polyphenic development. Functional genetic analysis of these components in *P. pacificus* could feasibly test their role in regulating or instructing alternative morphologies. Finally, if we assume some conservation of GRNs, as supported by our comparison of *P. pacificus* to *C. elegans*, we can hypothesize that network loss or rewiring, rather than decay of environmentally sensitive loci themselves, accompanied the canalization of a single morph. This hypothesis would be consistent with our finding that many *P. pacificus* polyphenism network components are involved in essential, conserved processes such as metabolic responses and dauer development.

Finally, our results provide a framework for determining how the molecular interactions among environmentally sensitive loci affect the ultimate morphologies produced. Specifically, it has highlighted the types of targets that polyphenism expression networks influence. At a system-wide level, these have apparently not drawn on novel genes, at least not more than expected by chance, despite the potential of such genes to drive network changes (Zhang et al. 2015) and otherwise rich substrate for the selection of novel traits (Kaessmann 2010; Tautz and Domazet-Lošo 2011; Chen et al. 2013). This is consistent with findings in other systems in which a novel trait is regulated as a plastic response. For example, facultative eusociality in bumble bees is apparently achieved through GRNs enriched for conserved genes, which simply required putative rearrangements in connectivity (Kapheim et al. 2020). Similarly, in the case of a morphological novelty, head horns in dung beetles, degree of plasticity is correlated with the number of its inferred regulatory components, which are not enriched for novel genes (Casasa et al. 2020). Other examples suggest that ancestral stress responses, beyond facilitating the regulation of a novel trait, may even cause the initial appearance of that trait (Swafford and Oakley 2019; Wagner et al. 2019). If already part a switch-like regulatory response, for instance, that of dauer formation, the result might be binary development of a novel morphology. Although it is additionally possible that novel loci of large effect may contribute to the evolution of novel, plastic traits, our results suggest that a polyphenic innovation relied in large part on the repurposing of ancient environmentally sensitive genes. Using the comparative platform we establish here, functional tests may distinguish how GRN modules of their components have individually contributed to the evolution of mouth morphologies in Diplogastridae.

In summary, we have examined genome-wide mechanisms of a morphological polyphenism in a gene coexpression network context. We provide evidence that a polyphenic developmental response evolved by recruiting and repurposing components from two ancient plastic responses, starvation and a diapause switch, suggesting that recruitment of ancient plastic networks may be a general mechanism to facilitate the origin of a novel plastic response. We also found that this process involved their integration with genes with rapid turnover, which in principle contribute to the features that distinguish the novel response. We predict that functional-genetic queries will reconstruct how such networks are recruited in the first place and are later accommodated into a fully functional, novel phenotype.

Materials and Methods

Gene Coexpression Network Inference

To infer gene networks regulated in the *P. pacificus* polyphenism, we used transcriptomic count data previously collected for alternative mouth-morphs, specifically using fully penetrant polyphenism mutants (i.e., constitutively Eu and St lines) under standardized genetic and environmental conditions (Bui and Ragsdale 2019). From the raw counts, we constructed a signed gene coexpression network, using the WGCNA package in R (Langfelder and Horvath 2008). Modules were defined using default settings and then correlated with morph (Eu, St) and mutant background (*eud-1* mutant, constitutively St; *eud-1* overexpression, *nhr-40* mutant, and *seud-1*/*sult-1* mutant lines, constitutively Eu). We similarly inferred coexpression networks for *C. elegans* using transcriptomic counts data from Harvald et al. (2017). For this inference, module–trait correlation was performed using the

data set's hours (0–16) of exposure to starvation. To compare results across species, we identified conserved genes, which were defined by having at least a one-way best BlastP match between for *P. pacificus* and *C. elegans*. To construct the developmental stage-specific coexpression network for *P. pacificus*, we used transcriptomic counts generated by Baskaran et al. (2015), which were correlated here with dauer diapause, early development (pooled J1-J3 stages), and late development (pooled J3, J4, and adult stages) as traits. For this analysis, expression data from nondauer life-stages were binned into "early" and "late" categories according to how samples grouped in a hierarchical cluster analysis (supplementary fig. S2*A*, Supplementary Material online), performed using the "fastcluster" package in R (Müllner 2013). Networks were visualized (fig. 2*B*) using Cytoscape (Shannon et al. 2003).

To determine tissue-specific expression of polyphenism-gene homologs in a nonpolyphenic species (*C. elegans*), we downloaded data from the *C. elegans* Gene Expression database (www.gfpworm.org; last accessed June 26th 2020) for eight tissues: pharyngeal gland, pharyngeal neurons, amphids (sensory organs), head neurons, pharynx, hypodermis (epidermis), and arcade syncytia. We then identified *C. elegans* homologs of polyphenism genes among these sets using BlastP and found their distribution across each of the polyphenism-specific modules.

Nematode Strains and Inbreeding

To make phylogenetic comparisons of genes in *P. pacificus* polyphenism networks, we selected four species from across Diplogastridae (*A. sudhausi*, *D. magnus*, *K. luziae*, *L. texanum*) and a possible outgroup species (*Bunonema* sp.), as informed by previous phylogenetic studies, in addition to using published genomes for several species closer to *P. pacificus* (Rödelsperger et al. 2014; Prabh et al. 2018) and for several Rhabditidae. Prior to sequencing, we systematically inbred lines from three of the above species (*A. sudhausi* strain SB413, *D. magnus* TMG144, *L. texanum* TMG5), all of which are androdioecious (i.e., with males and self-fertilizing hermaphrodites), to increase homozygosity throughout their genomes. During this process, we isolated five individual hermaphrodites (morphological females) at each generation, with the subsequent generation being established from the most fecund pair. We inbred these lines under standard environmental conditions, namely at room temperature on nematode growth medium agar plates seeded with 400 µl *Escherichia coli* OP50 in L-broth. Systematic inbreeding was unsuccessful for the gonochoristic (i.e., obligately outcrossing) species *K. luziae* (NKZ323), and it was not performed for *Bunonema* sp. (RGD898), which is parthenogenic and hence was presumed to be highly homozygous. Nematodes of *K. luziae* were maintained in cultures as above; *Bunonema* sp. was reared on agar plates (1:1, agar and sterilized water), with 5–6 chunks of nematode growth medium placed on top of the agar medium, and kept at room temperature.

DNA and RNA Extraction, Preparation, and Sequencing

For each species, we rinsed nematodes from 8 to 10 mixed-stage cultures into M9 buffer in a 45-ml collection tube, after which nematodes were washed and shaken with 0.9% NaCl with ampicillin (50 µg/µl) overnight. This process was repeated, with the addition of chloramphenicol (2 µg/µl) to the wash solution for the second night. Washed nematodes were pelleted by slow centrifugation at 1,300 rpm for 4 min, with the pellet thereafter moved to a 1.5-ml tube and stored at −80 °C until extraction. DNA was extracted using the Epicentre MasterPure Complete DNA & RNA Purification Kit, following the manufacturer's instructions. DNA libraries were prepared using the TruSeq Stranded mRNA HT Sample Prep Kit (Illumina) according to the manufacturer's protocol, and 8-nt barcodes were added for multiplexing. Barcoded libraries were verified using a Qubit3 fluorometer (ThermoFisher Scientific) and a 2200 TapeStation bioanalyzer (Agilent Technologies) and sequenced on a NextSeq 500 (Illumina) with NextSeq300 Mid Output v2 kit (Illumina).

For RNA extraction, we collected nematodes as above. To remove microbial contaminants, we added 40 µl of ampicillin (50 µg/µl) and 40 µl of chloramphenicol (25 µg/µl) to a 40-ml solution of 0.9% NaCl containing the collected nematodes, which were then washed with gentle shaking for 2 h at room temperature. Washed nematodes were stored in 1 ml of TRIzol and frozen at −80 °C until extraction. Frozen nematode pellets (in TRIzol) were then frozen and thawed three times in liquid nitrogen, with each freeze and thaw lasting 5 min. Total RNA was then extracted using a Direct-zol RNA Miniprep Kit (Zymo Research). RNA libraries were prepared using the Illumina TruSeq RNA Library Prep Kit v2, after which libraries were sequenced on an Illumina HiSeq 3000.

Genome Assembly and Annotation

We generated ∼100–150 million paired-end reads of 150- and 500-bp insert sizes for each of the five nematode lines described above. Assemblies for *A. sudhausi*, *D. magnus*, and *L. texanum* were as follows. Reads were trimmed with Cutadapt (Martin 2011) and error-corrected with Reckoner (Długosz and Deorowicz 2017). We made a preliminary single-end assembly using Minia (Drezen et al. 2014) to check the insert size, and we identified possible contaminants in this assembly with Blobtools (Kumar et al. 2013). We then fully assembled the nuclear genomes for these species using SPAdes (Bankevich et al. 2012) and removed scaffolds that were identified as contamination, at low coverage (<1×), and < 500 bp in length. We masked repeats using RepeatModeler and RepeatMasker (Smit et al. 2015). We annotated these genomes using the BRAKER pipeline (Hoff et al. 2019), including the BAM file created by aligning RNA-Seq reads to the genome with STAR (Dobin et al. 2013). We independently assembled the mitochondrial genomes for these species with plasmidSPAdes (Antipov et al. 2016) and annotated them using MITOS2 on the MITOS Web Server (Bernt et al. 2013). To assemble the transcriptomes of these species, we trimmed RNA-Seq reads using Trimmomatic

(Bolger et al. 2014), followed by assembly with Shannon (Kannan et al. 2016). We used Kallisto (Bray et al. 2016) to quantify expression levels.

For genome assemblies of *K. luziae* and *Bunonema* sp., raw reads were first adapter-trimmed and quality-filtered using Trimmomatic, requiring a minimum base quality score of 20 averaged across a sliding window of 3 bases (parameters, *LEADING:20 TRAILING:20 SLIDINGWINDOW:3:20*). Reads shorter than 20 bases posttrimming were discarded. Cleaned reads were assembled using SPAdes, with default parameters, as for the genome assemblies described above. Some of the larger scaffolds with relatively very high coverage were compared with the rest suggested possible mitochondrial and bacterial contamination. From the assembled contigs, we extracted 10-kb fragments (sliding over 5 kb) and computed 4-mer frequencies for each fragment. These oligonucleotide frequency data were mapped to 3D space using Rtsne (van der Maaten and Hinton 2008) and clustered using DBSCAN (Ertöz et al. 2003). Each scaffold was assigned to the cluster with the highest number of fragments derived from that scaffold. Scaffold sequences from each cluster were aligned to NCBI nt database using BLAST to obtain the corresponding taxonomic associations, which distinguished nematode sequences from those of bacterial contaminants, that latter of which were filtered from the assemblies. Using a Burrows–Wheeler aligner (Li and Durbin 2009), the cleaned reads were mapped back to the assembly scaffolds. Gene predictions were annotated using Augustus (Stanke and Morgenstern 2005) with protein sequences from *P. pacificus* and *C. elegans* used as training set. Mitochondrial genomes were annotated using MITOS2. To assemble the transcriptomes of *K. luziae* and *Bunonema* sp., raw RNA-Seq reads were trimmed as described above for DNA reads. Stranded transcriptome assemblies were generated from the cleaned reads using RnaSPAdes (Bushmanova et al. 2019).

We assessed the completeness of all genome assemblies by a comparison with the Nematoda odb9 BUSCO data set using HMMER (Johnson et al. 2010). Genome assembly statistics are given in supplementary table S1, Supplementary Material online.

Species-Tree Inference

To enable a historical reconstruction of polyphenism network components, we inferred the phylogeny for Diplogastridae and its potential outgroups. For this inference, we identified single-copy orthologs using Orthofinder (Emms and Kelly 2015), selecting 928 clusters on the basis of having at least 75% species present with a single protein in the cluster. An evolutionary model was selected automatically for each cluster as implemented in IQ-TREE (Nguyen et al. 2015). We selected single-copy subtrees with at least ten species and 50% BS using PhyloTreePruner (Kocot et al. 2013), which resulted in 778 trimmed clusters. We then took two approaches to phylogenetic inference. First, we concatenated the conceptually translated amino acid sequences into a supermatrix using FASconCAT-G (Kück and Longo 2014), after which we inferred the phylogeny under the maximum likelihood (ML) criterion and invoking a GTR+G substitution model, as implemented in RAxML-NG (Kozlov et al. 2019). In parallel, we performed individual ML inferences for each trimmed gene-cluster in IQ-TREE, following which we inferred the species tree under a multiple-species coalescent model, as implemented in ASTRAL (Mirarab et al. 2014).

Identification of Conserved and Diplogastrid-Restricted Genes

To identify homologs of *P. pacificus* genes in polyphenism-specific modules, we performed a genome-wide, reciprocal best-hit BlastP search against each of the species in figure 5A, after which we searched for genes that belonged to polyphenism-specific modules. To identify diplogastrid-restricted genes, we first identified *P. pacificus* homologs using BlastP (with either direction considered as a hit) against rhabditid species (*C. elegans*, *C. japonica*, *C. angaria*, *Diploscapter pachys*, *Mesorhabditis belari*) and *Bunonema* sp. We then searched for *P. pacificus* genes that were absent in this set of conserved genes and defined them as diplogastrid specific. To identify *P. pacificus* genes that had been uniquely lost in the lineage exclusive to *L. texanum* and *D. magnus*, we used BlastP to identify the set of *P. pacificus* genes that had identifiable homologs in *A. sudhausi* or *K. luziae* but none in *L. texanum* and *D. magnus*, thereby rooting the loss of those genes to the monomorphic clade.

Inference of Gene-Family Evolution

To determine whether genes belonged to rapidly evolving families, we examined gene-family history against a completely resolved species tree, inferred as described above. Prior to this analysis, we fitted the tree to an ultrametric length using r8s (Sanderson 2003), with the age of the tree calibrated using a molecular-clock estimate of divergence between *P. pacificus* and *C. elegans* (250 Ma; Dieterich et al. 2008). To decrease potential error due to deep divergences, our analysis used a tree pruned with Newick Utilities (Junier and Zdobnov 2010) to be restricted to Diplogastridae. After using OrthoMCL (Li et al. 2003) to find orthologous clusters across all species, we estimated gene family expansion and contraction across Diplogastridae using CAFE 5.0, which allows statistical tests of gene-family evolution while also accounting for variable error in genome assemblies and annotations (Han et al. 2013). To account for assembly errors, CAFE was run using three replicates and a Poisson distribution with applied estimated error.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We thank the Indiana University Center for Genomics and Bioinformatics for sequencing, and particularly Dr Doug Rusch and Dr Ram Podicheti (CGB) for analysis consultation and genome assembly (*Bunonema* sp., *Koerneria luziae*). This

work was funded by the National Science Foundation (Grant No. IOS 1911688 to E.J.R.).

Author Contributions

S.C. and E.J.R.: study conception and design; J.F.B.: strain inbreeding and sequence preparation; G.D.K.: genome assembly and species-tree inference; S.C.: analyses and interpretation of data; S.C. and E.J.R.: writing of the article.

Data Availability

Nuclear genome sequences have been deposited in NCBI GenBank as BioProject PRJNA655932 under accession numbers GITX00000000, GITY00000000, GITZ00000000, GIUA00000000, and GIUD00000000 for *Diplogasteroides magnus*, *Levipalatum texanum*, *Bunonema* sp. RGD898, *Koerneria luziae*, and *Allodiplogaster sudhausi*, respectively. Raw sequencing reads are available at the NCBI Sequence Read Archive (PRJNA655932). Mitochondrial genome sequences are separately available from GenBank (MT920391–MT920395).

References

Abouheif E, Wray GA. 2002. Evolution of the gene network underlying wing polyphenism in ants. *Science* 297(5579):249–252.

Antebi A, Culotti JG, Hedgecock EM. 1998. *daf-12* regulates developmental age and the dauer alternative in *Caenorhabditis elegans*. *Development* 125(7):1191–1205.

Antipov D, Hartwick N, Shen M, Raiko M, Lapidus A, Pevzner PA. 2016. plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinformatics* 32(22):3380–3387.

Aubin-Horth N, Landry CR, Letcher BH, Hofmann HA. 2005. Alternative life histories shape brain gene expression profiles in males of the same population. *Proc Biol Sci*. 272(1573):1655–1662.

Aubin-Horth N, Renn SCP. 2009. Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Mol Ecol*. 18(18):3763–3780.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 19(5):455–477.

Baskaran P, Rödelsperger C, Prabh N, Serobyan V, Markov GV, Hirsekorn A, Dieterich C. 2015. Ancient gene duplications have shaped developmental stage-specific expression in *Pristionchus pacificus*. *BMC Evol Biol*. 15(1):185.

Beldade P, Mateus ARA, Keller RA. 2011. Evolution and molecular mechanisms of adaptive developmental plasticity. *Mol Ecol*. 20(7):1347–1363.

Bento G, Ogawa A, Sommer RJ. 2010. Co-option of the hormone-signalling module dafachronic acid-DAF-12 in nematode evolution. *Nature* 466(7305):494–497.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evol*. 69(2):313–319.

Bhardwaj S, Jolander LSH, Wenk MR, Oliver JC, Frederik Nijhout H, Monteiro A. 2020. Origin of the mechanism of phenotypic plasticity in satyrid butterfly eyespots. *eLife* 9:e49544.

Biddle JF, Ragsdale EJ. 2020. Regulators of an ancient polyphenism evolved through episodic protein divergence and parallel gene radiations. *Proc Biol Sci*. 287(1921):20192595.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.

Bose N, Ogawa A, von Reuss SH, Yim JJ, Ragsdale EJ, Sommer RJ, Schroeder FC. 2012. Complex small-molecule architectures regulate phenotypic plasticity in a nematode. *Angew Chem Int Ed Engl*. 51(50):12438–12443.

Bradshaw AD. 1965. Evolutionary significance of phenotypic plasticity in plants. *Adv Genet*. 13(C):115–155.

Brakefield PM, Gates J, Keys D, Kesbeke F, Wijngaarden PJ, Montelro A, French V, Carroll SB. 1996. Development plasticity and evolution of butterfly eyespot patterns. *Nature* 384(6606):236–242.

Brakefield PM, Reitsma N. 1991. Phenotypic plasticity, seasonal climate and the population biology of *Bicyclus* butterflies (Satyridae) in Malawi. *Ecol Entomol*. 16(3):291–303.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 34(5):525–527.

Brisson JA, Davis GK. 2016. The right tools for the job: regulating polyphenic morph development in insects. *Curr Opin Insect Sci*. 13:1–6.

Bui LT, Ivers NA, Ragsdale EJ. 2018. A sulfotransferase dosage-dependently regualtes mouthpart dimorphism in the nematode *Pristionchus pacificus*. *Nat Commun*. 9(1):4119.

Bui LT, Ragsdale EJ. 2019. Multiple plasticity regulators reveal targets specifying an induced predatory form in nematodes. *Mol Biol Evol*. 36(11):2387–2399.

Bushmanova E, Antipov D, Lapidus A, Prjibelski AD. 2019. rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *Gigascience* 8(9):giz100.

Casasa S, Moczek AP. 2018. Insulin signalling's role in mediating tissue-specific nutritional plasticity and robustness in the horn-polyphenic beetle *Onthophagus taurus*. *Proc Biol Sci*. 285(1893):20181631.

Casasa S, Zattara EE, Moczek AP. 2020. Nutrition-responsive gene expression and the developmental evolution of insect polyphenism. *Nat Ecol Evol*. 4(7):970–978.

Chen S, Krinsky BH, Long MY. 2013. New genes as drivers of phenotypic evolution. *Nat Rev Genet*. 14(9):645–660.

Davidson E. 2006. The regulatory genome. Cambridge (MA): Academic Press.

Dieterich C, Clifton SW, Schuster LN, Chinwalla A, Delehaunty K, Dinkelacker I, Fulton L, Fulton R, Godfrey J, Minx P, et al. 2008. The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nat Genet*. 40(10):1193–1198.

Dlugosz M, Deorowicz S. 2017. RECKONER: read error corrector based on KMC. *Bioinformatics* 33(7):1086–1089.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1):15–21.

Drezen E, Rizk G, Chikhi R, Deltel C, Lemaitre C, Peterlongo P, Lavenier D. 2014. GATB: genome assembly & analysis tool box. *Bioinformatics* 30(20):2959–2961.

Ehrenreich IM, Pfennig DW. 2016. Genetic assimilation: a review of its potential proximate causes and evolutionary consequences. *Ann Bot*. 117(5):769–779.

Emlen DJ, Hunt J, Simmons LW. 2005. Evolution of sexual dimorphism and male dimorphism in the expression of beetle horns: phylogenetic evidence for modularity, evolutionary lability, and constraint. *Am Nat*. 166(4):S42–S68.

Emlen DJ, Nijhout HF. 1999. Hormonal control of male horn length dimorphism in the dung beetle *Onthophagus taurus* (Coleoptera: Scarabaeidae). *J Insect Physiol*. 45(1):45–53.

Emlen DJ, Warren IA, Johns A, Dworkin I, Lavine LC. 2012. A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. *Science* 337(6096):860–864.

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol*. 16(1):157.

Ertöz L, Steinbach M, Kumar V. 2003. Finding clusters of different sizes, shapes, and densities in noisy, high dimensional data. In: Proceedings of the Third Siam International Conference on Data Mining. p. 47–58.

Filteau M, Pavey SA, St-Cyr J, Bernatchez L. 2013. Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. *Mol Biol Evol*. 30(6):1384–1396.

Hahn MW, Kern AD. 2005. Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks. *Mol Biol Evol*. 22(4):803–806.
Han MV, Thomas GWC, Lugo-Martinez J, Hahn MW. 2013. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol Biol Evol*. 30(8):1987–1997.
Harvald EB, Sprenger RR, Dall KB, Ejsing CS, Nielsen R, Mandrup S, Murillo AB, Larance M, Gartner A, Lamond AI, et al. 2017. Multi-omics analyses of starvation responses reveal a central role for lipoprotein metabolism in acute starvation survival in *C. elegans*. *Cell Syst*. 5(1):38–52.
Hoff KJ, Lomsadze A, Borodovsky M, Stanke M. 2019. Whole-genome annotation with BRAKER. *Methods Mol Biol*. 962:65–95.
Hu PJ. 2007. Dauer. In: WormBook, ed. The *C. elegans* Research Community. Available from: http://www.wormbook.org/chapters/www_dauer/dauer.html. Accessed June 26, 2020.
Johnson LS, Eddy SR, Portugaly E. 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* 11(1):431.
Jones BM, Robinson GE. 2018. Genetic accommodation and the role of ancestral plasticity in the evolution of insect eusociality. *J Exp Biol*. 221(23):jeb153163.
Junier T, Zdobnov EM. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26(13):1669–1670.
Kaessmann H. 2010. Origins, evolution, and phenotypic impact of new genes. *Genome Res*. 20(10):1313–1326.
Kannan S, Hui J, Mazooji K, Pachter L, Tse D. 2016. Shannon: an information-optimal *de novo* RNA-Seq assembler. BioRxiv, 039230.
Kapheim KM, Jones BM, Pan HL, Cai L, Harpur BA, Clement F, Zayed A, Ioannidis P, Waterhouse RM, Kingwell C, et al. 2020. Developmental plasticity shapes social traits and selection in a facultatively eusocial bee. *Proc Natl Acad Sci U S A*. 117(24):13615–13625.
Kenkel CD, Matz MV. 2017. Gene expression plasticity as a mechanism of coral adaptation to a variable environment. *Nat Ecol Evol*. 1(1):14.
Kieninger MR, Ivers NA, Rödelsperger C, Markov GV, Sommer RJ, Ragsdale EJ. 2016. The nuclear hormone receptor NHR-40 acts downstream of the sulfatase EUD-1 as part of a developmental plasticity switch in *Pristionchus*. *Curr Biol*. 26(16):2174–2179.
Kijimoto T, Moczek AP. 2016. Hedgehog signaling enables nutrition-responsive inhibition of an alternative morph in a polyphenic beetle. *Proc Natl Acad Sci U S A*. 113(21):5982–5987.
Kimura KD, Tissenbaum HA, Liu YX, Ruvkun G. 1997. *daf-2*, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science* 277(5328):942–946.
Kiontke K, Barriere A, Kolotuev I, Podbilewicz B, Sommer R, Fitch DH, Felix MA. 2007. Trends, stasis, and drift in the evolution of nematode vulva development. *Curr Biol*. 17(22):1925–1937.
Kocot KM, Citarella MR, Moroz LL, Halanych KM. 2013. PhyloTreePruner: a phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. *Evol Bioinform Online*. 9:429–435.
Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A, Wren J. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35(21):4453–4455.
Kück P, Longo GC. 2014. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Front Zool*. 11(1):81.
Kumar S, Jones M, Koutsovoulos G, Clarke M, Blaxter M. 2013. Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. *Front Genet*. 4:237.
Lafuente E, Beldade P. 2019. Genomics of developmental plasticity in animals. *Front Genet*. 10:720.
Lande R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J Evol Biol*. 22(7):1435–1446.
Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9(1):559.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
Li L, Stoeckert CJ, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res*. 13(9):2178–2189.
Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBNET J*. 17(1):10–12.
Masel J, Siegal ML. 2009. Robustness: mechanisms and consequences. *Trends Genet*. 25(9):395–403.
Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* 30(17):i541–i548.
Müllner D. 2013. Fastcluster: fast hierarchical, agglomerative clustering routines for R and Python. *J Stat Softw*. 53(9):1–18.
Namdeo S, Moreno E, Rödelsperger C, Baskaran P, Witte H, Sommer RJ. 2018. Two independent sulfation processes regulate mouth-form plasticity in the nematode *Pristionchus pacificus*. *Development* 145(13):dev66272.
Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 32(1):268–274.
Nijhout HF. 2003. Development and evolution of adaptive polyphenisms. *Evol Dev*. 5(1):9–18.
Nijhout HF, McKenna KZ. 2018. The distinct roles of insulin signaling in polyphenic development. *Curr Opin Insect Sci*. 25:58–64.
Ogawa A, Bento G, Bartelmes G, Dieterich C, Sommer RJ. 2011. *Pristionchus pacificus daf-16* is essential for dauer formation but dispensable for mouth form dimorphism. *Development* 138(7):1281–1284.
Ogawa A, Streit A, Antebi A, Sommer RJ. 2009. A conserved endocrine mechanism controls the formation of dauer and infective larvae in nematodes. *Curr Biol*. 19(1):67–71.
Ogg S, Paradis S, Gottlieb S, Patterson GI, Lee L, Tissenbaum HA, Ruvkun G. 1997. The fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. *Nature* 389(6654):994–999.
Parker BJ, Brisson JA. 2019. A laterally transferred viral gene modifies aphid wing plasticity. *Curr Biol*. 29(12):2098–2103.
Portereiko MF, Mango SE. 2001. Early morphogenesis of the *Caenorhabditis elegans* pharynx. *Dev Biol*. 233(2):482–494.
Prabh N, Roeseler W, Witte H, Eberhardt G, Sommer RJ, Rödelsperger C. 2018. Deep taxon sampling reveals the evolutionary dynamics of novel gene families in *Pristionchus* nematodes. *Genome Res*. 28(11):1664–1674.
Projecto-Garcia J, Biddle JF, Ragsdale EJ. 2017. Decoding the architecture and origins of mechanisms for developmental polyphenism. *Curr Opin Genet Dev*. 47:1–8.
Ragsdale EJ, Ivers NA. 2016. Specialization of a polyphenism switch gene following serial duplications in *Pristionchus* nematodes. *Evolution* 70(9):2155–2166.
Ragsdale EJ, Müller MR, Rödelsperger C, Sommer RJ. 2013. A developmental switch coupled to the evolution of plasticity acts through a sulfatase. *Cell* 155(4):922–933.
Renn SCP, Schumer ME. 2013. Genetic accommodation and behavioural evolution: insights from genomic studies. *Anim Behav*. 85(5):1012–1022.
Rödelsperger C, Neher RA, Weller AM, Eberhardt G, Witte H, Mayer WE, Dieterich C, Sommer RJ. 2014. Characterization of genetic diversity in the nematode *Pristionchus pacificus* from population-scale resequencing data. *Genetics* 196(4):1153–1165.
Sanderson MJ. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19(2):301–302.
Scheiner SM. 1993. Genetics and evolution of phenotypic plasticity. *Annu Rev Ecol Syst*. 24(1):35–68.
Schlichting CD, Pigliucci M. 1993. Control of phenotypic plasticity via regulatory genes. *Am Nat*. 142(2):366–370.
Schneider RF, Li Y, Meyer A, Gunter HM. 2014. Regulatory gene networks that shape the development of adaptive phenotypic plasticity in a cichlid fish. *Mol Ecol*. 23(18):4511–4526.

Schneider RF, Meyer A. 2017. How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations. *Mol Ecol*. 26(1):330–350.

Serobyan V, Ragsdale EJ, Müller MR, Sommer RJ. 2013. Feeding plasticity in the nematode *Pristionchus pacificus* is influenced by sex and social context and is linked to developmental speed. *Evol Dev*. 15(3):161–170.

Serobyan V, Ragsdale EJ, Sommer RJ. 2014. Adaptive value of a predatory mouth-form in a dimorphic nematode. *Proc R Soc B*. 281(1791):20141334.

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(11):2498–2504.

Sieriebriennikov B, Markov GV, Witte H, Sommer RJ. 2017. The role of DAF-21/Hsp90 in mouth-form plasticity in *Pristionchus pacificus*. *Mol Biol Evol*. 34(7):1644–1653.

Sieriebriennikov B, Prabh N, Dardiry M, Witte H, Röseler W, Kieninger MR, Rödelsperger C, Sommer RJ. 2018. A developmental switch generating phenotypic plasticity is part of a conserved multi-gene locus. *Cell Rep*. 23(10):2835–2843.

Sieriebriennikov B, Sun S, Lightfoot JW, Witte H, Moreno E, Rödelsperger C, Sommer RJ. 2020. Conserved nuclear hormone receptors controlling a novel plastic trait target fast-evolving genes expressed in a single cell. *PLoS Genet*. 16(4):e1008687.

Smit A, Hubley R, Grenn P. 2015. RepeatMasker Open-4.0. 2013–2015. Available from: http://repeatmasker.org. Accessed June 26, 2020.

Sommer RJ. 2020. Phenotypic plasticity: from theory and genetics to current and future challenges. *Genetics* 215(1):1–13.

Stanke M, Morgenstern B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res*. 33(Web Server):W465–W467.

Sudhaus W. 2011. Phylogenetic systemisation and catalogue of paraphyletic "Rhabditidae" (Secernentea, Nematoda). *J Nematode Morphol Syst*. 14(2):113–178.

Susoy V, Ragsdale EJ, Kanzaki N, Sommer RJ. 2015. Rapid diversification associated with a macroevolutionary pulse of developmental plasticity. *eLife* 4:e05463.

Swafford AJ, Oakley TH. 2019. Light-induced stress as a primary evolutionary driver of eye origins. *Integr Comp Biol*. 59(4):739–750.

Tautz D, Domazet-Lošo T. 2011. The evolutionary origin of orphan genes. *Nat Rev Genet*. 12(10):692–702.

van Bergen E, Osbaldeston D, Kodandaramaiah U, Brattström O, Aduse-Poku K, Brakefield PM. 2017. Conserved patterns of integrated developmental plasticity in a group of polyphenic tropical butterflies. *BMC Evol Biol*. 17(1):59.

van der Maaten L, Hinton G. 2008. Visualizing data using t-SNE. *J Mach Learn Res*. 9:2579–2605.

van Megen H, van den Elsen S, Holterman M, Karssen G, Mooyman P, Bongers T, Holovachov O, Bakker J, Helder J. 2009. A phylogenetic tree of nematodes based on about 1200 full-length small subunit ribosomal DNA sequences. *Nematology* 11(6):927–950.

Vangestel S, Houthoofd W, Bert W, Borgonie G. 2008. The early embryonic development of the satellite organism *Pristionchus pacificus*: differences and similarities with *Caenorhabditis elegans*. *Nematology* 10(3):301–312.

Via S, Gomulkiewicz R, De Jong G, Scheiner SM, Schlichting CD, Van Tienderen PH. 1995. Adaptive phenotypic plasticity: consensus and controversy. *Trends Ecol Evol*. 10(5):212–217.

Voineagu I, Wang XC, Johnston P, Lowe JK, Tian Y, Horvath S, Mill J, Cantor RM, Blencowe BJ, Geschwind DH. 2011. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* 474(7351):380–384.

Wagner A. 2011. The origins of evolutionary innovations: a theory of transformative change in living systems. Oxford: Oxford University Press.

Wagner GP, Erkenbrack EM, Love AC. 2019. Stress-induced evolutionary innovation: a mechanism for the origin of cell types. *BioEssays* 41(4):1800188.

Wagner GP, Pavlicev M, Cheverud JM. 2007. The road to modularity. *Nat Rev Genet*. 8(12):921–931.

Walley JW, Sartor RC, Shen Z, Schmitz RJ, Wu KJ, Urich MA, Nery JR, Smith LG, Schnable JC, Ecker JR, et al. 2016. Integration of omic networks in a developmental atlas of maize. *Science* 353(6301):814–818.

West-Eberhard MJ. 2003. Developmental plasticity and evolution. Oxford: Oxford University Press.

West-Eberhard MJ. 2005. Developmental plasticity and the origin of species differences. *Proc Natl Acad Sci U S A*. 102(Suppl 1):6543–6549.

Xu HJ, Xue J, Lu B, Zhang XC, Zhuo JC, He SF, Ma XF, Jiang YQ, Fan HW, Xu JY, et al. 2015. Two insulin receptors determine alternative wing morphs in planthoppers. *Nature* 519(7544):464–467.

Xue ZG, Huang K, Cai CC, Cai LB, Jiang CY, Feng Y, Liu ZS, Zeng Q, Cheng LM, Sun YE, et al. 2013. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* 500(7464):593–597.

Yang CH, Pospisilik JA. 2019. Polyphenism – a window into gene-environment interactions and phenotypic plasticity. *Front Genet*. 10:132.

Zhang W, Landback P, Gschwend AR, Shen B, Long M. 2015. New genes drive the evolution of gene interaction networks in the human and mouse genomes. *Genome Biol*. 16(1):202.