Single worm transcriptomics identifies a developmental core network of oscillating genes with deep conservation across nematodes

Shuai Sun, Christian Rödelsperger and Ralf J. Sommer

Genome Res. 2021 31: 1590-1601 originally published online July 22, 2021
Access the most recent version at doi:10.1101/gr.275303.121

Single worm transcriptomics identifies a developmental core network of oscillating genes with deep conservation across nematodes

Shuai Sun, Christian Rödelsperger, and Ralf J. Sommer

Max-Planck Institute for Developmental Biology, Department for Integrative Evolutionary Biology, 72076 Tübingen, Germany

High-resolution spatial and temporal maps of gene expression have facilitated a comprehensive understanding of animal development and evolution. In nematodes, the small body size represented a major challenge for such studies, but recent advancements have helped overcome this limitation. Here, we have implemented single worm transcriptomics (SWT) in the nematode model organism *Pristionchus pacificus* to provide a high-resolution map of the developmental transcriptome. We selected 38 time points from hatching of the J2 larvae to young adults to perform transcriptome analysis over 60 h of post-embryonic development. A mean sequencing depth of 4.5 million read pairs allowed the detection of more than 23,135 (80%) of all genes. Nearly 3000 (10%) genes showed oscillatory expression with discrete expression levels, phases, and amplitudes. Gene age analysis revealed an overrepresentation of ancient gene classes among oscillating genes, and around one-third of them have 1:1 orthologs in *C. elegans*. One important gene family overrepresented among oscillating genes is collagens. Several of these collagen genes are regulated by the developmental switch gene *eud-1*, indicating a potential function in the regulation of mouth-form plasticity, a key developmental process in this facultative predatory nematode. Together, our analysis provides (1) an updated protocol for SWT in nematodes that is applicable to many microscopic species, (2) a 1- to 2-h high-resolution catalog of *P. pacificus* gene expression throughout postembryonic development, and (3) a comparative analysis of oscillatory gene expression between the two model organisms *P. pacificus* and *C. elegans* and associated evolutionary dynamics.

[Supplemental material is available for this article.]

The evolution of novelty and phenotypic divergence depends on two major molecular processes. First, the emergence of new genes through gene duplication, divergence, and de novo formation constantly change the genomic composition of organisms (for review, see Rödelsperger et al. 2019b). Second, changes in the expression of individual genes during development can lead to phenotypic divergence. In general, transcriptomes show highly dynamic spatiotemporal patterns that are orchestrated by multiple regulatory events. However, a detailed understanding of transcriptional changes and their significance for evolution requires sophisticated molecular tools and comparative analyses. In recent years, the ENCODE and modENCODE consortia have generated unprecedented amounts of RNA sequencing data with increasing resolution, which also allowed for the comparison across distant species (Gerstein et al. 2014). Most recently, high-resolution spatial maps of gene expression are made possible through single-cell transcriptional profiling even in animals with small body size such as *Caenorhabditis elegans* (Cao et al. 2017). High temporal resolution throughout embryonic and/or postembryonic development requires the analysis of many parallel samples. In *C. elegans*, the use of highly synchronized cultures has provided the first comprehensive maps of gene expression throughout development (Levin et al. 2012; Kim et al. 2013; Hendriks et al. 2014; Meeuse et al. 2020). However, such methods are restricted to organisms that can be massively cultured and that stay synchronous throughout development. One alternative methodology is single worm transcriptomics (SWT), a method that has originally been introduced in two entomopathogenic nematodes of the genus *Steinernema* (Macchietto et al. 2017) and recently has been revised (Chang et al. 2021). Although such a method could in theory be adopted in many organisms, to the best of our knowledge, it has not yet been widely applied. Here, we implemented SWT in the model nematode *Pristionchus pacificus* with a new updated protocol that uses (1) a different method of worm lysis, (2) even further reduced cDNA input, and (3) reduced PCR amplification cycles to provide a high-resolution map of the transcriptome throughout postembryonic development.

P. pacificus is a well-established model system in evolutionary biology first described in 1996 (Sommer et al. 1996). Based on its rapid growth with a generation time of only 4 d in the laboratory and its hermaphroditic mode of reproduction, forward and reverse genetic techniques including CRISPR-associated engineering have been successfully implemented in this nematode (Witte et al. 2015; Han et al. 2020; Nakayama et al. 2020). Although original work focused on the evolution of developmental processes and comparison to *C. elegans* (Wang and Sommer 2011), more recent studies developed *P. pacificus* as a model system for developmental plasticity and the evolution of novelty and complexity (Sommer 2020). In contrast to *C. elegans*, *P. pacificus* forms two alternative mouth-forms with different teeth-like denticles, enabling different feeding strategies (Bento et al. 2010). So-called "stenostomatous" (St) animals have a single tooth and are strict bacterial feeders, whereas "eurystomatous" (Eu) animals have two teeth and can

Corresponding author: ralf.sommer@tuebingen.mpg.de
Article published online before print. Article, supplemental material, and publication date are at https://www.genome.org/cgi/doi/10.1101/gr.275303.121.

31:1590–1601 Published by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/21; www.genome.org

additionally feed on fungi and other nematodes by predation (Ragsdale et al. 2013). By now, the gene regulatory network controlling mouth-form plasticity has been identified (Sieriebriennikov et al. 2018, 2020), and the long-term evolutionary consequences of mouth-form plasticity for the evolution of novelty were carefully investigated (Susoy et al. 2015, 2016).

The development of teeth-like denticles, mouth-form plasticity, and predation are characters that are shared by several taxa of the Diplogastridae, the family to which *P. pacificus* belongs, but they are absent in *C. elegans* and its relatives (Susoy et al. 2015). Importantly, all diplogastrid nematodes share another developmental feature unknown from *C. elegans*. Specifically, *P. pacificus* hatches in the J2 juvenile stage after the first molt occurs in the egg. This heterochronic feature has severe consequences for the development of these organisms (Fürst von Lieven 2005). Specifically, there are only three free-living preadult larval stages, the J2, J3, and J4 stages, respectively. Thus, the *P. pacificus* J2, J3, and J4 stages correspond to the *C. elegans* L2, L3, and L4 stages, respectively. Consequently, there are only three molts to adulthood after hatching versus four molts in *C. elegans*. Finally, the time of development spent in the egg after fertilization is extended in *P. pacificus* to nearly 24 h (20°C), whereas hatching occurs after 18 h in *C. elegans* (Wood 1988). In consequence, postembryonic development in *P. pacificus* is concluded after 45 h of development and three molting cycles.

P. pacificus and *C. elegans* belong to different nematode families, the Diplogastridae and Rhabditidae, respectively (Kanzaki and Giblin-Davis 2015). Both species reproduce as hermaphrodites, which made them important model organisms with the associated establishment of extensive technology platforms as indicated above. *P. pacificus* and *C. elegans* are thought to have shared a last common ancestor ~100 mya (Prabh et al. 2018; Werner et al. 2018). Their genomes differ in size and the total number of genes, with the 100-Mb genome of *C. elegans* containing 20,040 genes and the 168-Mb genome of *P. pacificus* containing 28,896 genes (Athanasouli et al. 2020). Importantly, 30% of the *P. pacificus* genes have 1:1 orthologs in *C. elegans* (Rödelsperger et al. 2019a). *P. pacificus* is one of around 50 species in the genus *Pristionchus*, all of which are available as living cultures for molecular investigations (Kanzaki et al. 2021). Building on this high phylogenetic resolution, recent transcriptomic studies and deep taxon sampling revealed high evolutionary dynamics of novel gene families in *Pristionchus* evolution (Prabh et al. 2018; Rödelsperger et al. 2018). Together, the availability of various integrative genomics platforms and the large number of culturable species with its high phylogenetic resolution make *Pristionchus* a primary study system for animal and nematode evolution, the evolutionary dynamics of gene families, and new gene origin (for review, see Rödelsperger et al. 2019b). This study aims at establishing SWT for *P. pacificus* to generate a high-resolution temporal map of gene expression throughout its postembryonic development and at focusing on the investigation of oscillatory gene expression patterns.

Results

Detailed characterization of *P. pacificus* postembryonic development highlights major heterochrony in comparison to *C. elegans*

To establish a robust developmental framework for studying the temporal dynamics of gene expression in *P. pacificus* at high resolution, we first followed postembryonic development in highly synchronized cultures, starting with the J2 stage that hatches from the egg ~24 h after fertilization. When fed with *E. coli* OP50 on NGM agar plates at 20°C, *P. pacificus* needs ~46 h to reach adulthood. We first measured the duration of the intermolt period of the three juvenile stages J2, J3, and J4 and the time required for molting (Fig. 1A; Supplemental Fig. S1). There was little difference in the duration of the three molts, which consisted of an ~2-h lethargus–apolysis phase followed by a relatively short ecdysis of 10–15 min. These time points are similar to those measured in *C. elegans* under similar culture conditions (Kim et al. 2013; Hendriks et al. 2014). In contrast, the length of the intermolt differed substantially between stages and species. The *P. pacificus* J2 stage is by far the longest with 16 h, whereas the J3 and J4 stage are 9.6 and 12 h, respectively. In comparison to *C. elegans*, all juvenile stages are longer in *P. pacificus*, reflecting the somewhat longer overall generation time. However, the first free-living juvenile stages, *C. elegans* L1 and *P. pacificus* J2, are substantially longer than all others. In contrast, the *P. pacificus* J1 stage in the egg shell is drastically reduced with the J1 cuticle being formed 16–17 h after fertilization and hatching of the J2 larva after 24 h (Fürst von Lieven 2005). Thus, the *P. pacificus* J1 takes only ~8 h, whereas the L1 stage of *C. elegans* lasts for ~20 h (Kim et al. 2013). This recharacterization of postembryonic development is largely consistent with previous studies (Félix et al. 1999; Fürst von Lieven 2005) and highlights major differences in developmental timing between *C. elegans* and *P. pacificus*.

SWT provides a robust platform for studying the temporal dynamics of gene expression

With this timetable in hand, we decided to use 38 time points for transcriptomic analysis (Fig. 1B; Supplemental Table S1). These 38 time points cover the first 58 h after hatching of the J2 larvae with a 1- to 2-h resolution. Seven of them are after the final molt to provide some information about transcriptional changes within young adults. We conducted SWT using a modified protocol originally generated by Macchietto and coworkers (for details, see Methods) (Macchietto et al. 2017). We performed all experiments in triplicate and ran 114 independent samples on three lanes on an Illumina HiSeq 3000 (Fig. 1C; Supplemental Fig. S2). We obtained a mean sequencing depth of 4.5 million read pairs and only six out of 114 samples had fewer than 2 million read pairs (Supplemental Fig. S2C). We subsampled the raw data of one SWT and one standard RNA-seq (STD) library and profiled the distribution of the number of detected genes among nine levels of sequencing depth. We found that 2 million reads are sufficient to detect 95% of genes that are found with 12 million reads (Supplemental Fig. S2B). We removed the six libraries with fewer than 2 million PE reads in the following analysis. To compare gene expression between SWT and STD samples, we first investigated exon coverage at selected gene loci (Fig. 1D) and found that the gene structure of most genes showed complete agreement between STD and SWT (Supplemental Fig. S2D; Supplemental Table S13). Furthermore, we observed similar density distributions of gene expression levels between SWT and newly generated STD libraries (Fig. 1E). Finally, we found a strong correlation of gene expression when comparing the data set of the last lethargus of the fourth molt with a Pearson's correlation of $R^2 = 0.86$ (Fig. 1F). Note that we found similar correlations between the STD replicates ($R^2 = 0.97$) and between the SWT replicates ($R^2 = 0.92$) (Supplemental Fig. S2A). Thus, our modified protocol with (1) a different type of worm lysis, (2) lower cDNA

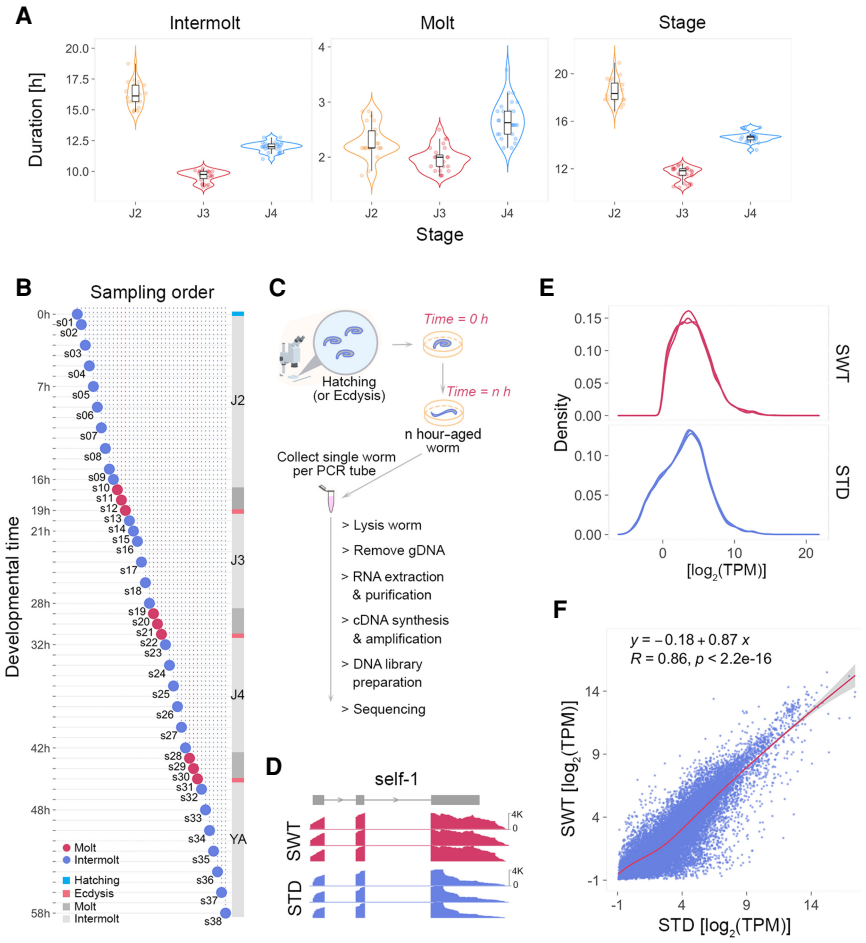

Figure 1. Single worm transcriptomic (SWT) sequencing in *P. pacificus*. (*A*) The violin plots display the duration of intermolt, molt, and the whole stage for all three juvenile stages (J2, J3, and J4). The head structure of the wild type at hatching, lethargus, and ecdysis is shown in Supplemental Figure S1. (*B*) The schematic shows the distribution of time points for SWT samples, which cover 58 h of postembryonic development from hatching to young adults. The three time points of each molt indicated in red represent the substages of molting: lethargus, apolysis, and ecdysis, respectively (for visualization, see Supplemental Fig. S1). Animals were collected hourly or every 2 h in the molt or intermolt stage. (*C*) Schematic summary of the workflow for SWT. (*D*) RNA-seq read coverage is visualized at the *self-1* locus using the Integrative Genomics Viewer (IGV) (Robinson et al. 2011) as one example for comparing SWT and standard RNA-seq (STD). (*E*) Density of gene expression of SWT and STD samples, six replicates each. (*F*) The scatterplot shows the correlation of the mean TPM of individual genes between SWT and STD data at the lethargus of the fourth molt. Three biological replicates for each method.

input, (3) reduced amplification cycles, and (4) larger sample pooling allows more high-throughput investigations and reduces associated costs without a strong decrease in data accuracy. Together, these data reveal that SWT in *P. pacificus* provides a robust platform for studying the temporal dynamics of gene expression throughout postembryonic development.

SWT establishes the most highly resolved developmental expression atlas of *P. pacificus*

Across all 38 time points, we detected 23,135 reliably expressed genes (REGs) with an average of aligned paired-end (PE) reads higher than or equal to three in at least one time point (Supplemental Table S2). More than 92% of these genes show expression evidence in at least two replicates of a single timepoint (Supplemental Fig. S2E), and ~95% of these genes are expressed in at least two time points (Supplemental Fig. S2F). This represents 80.1% of the currently annotated 28,896 genes of the *P. pacificus* genome (Athanasouli et al. 2020). Cross-correlation plots of gene expression profiles revealed two major developmental clusters representing early larval stages (6–38 h) and late larvae, including young adults (42–58 h) (Fig. 2A). This is consistent with a previous analysis of the developmental transcriptome of *P. pacificus*, in which principal component analysis (PCA) and hierarchical clustering also separated between early larvae and late larvae/adults. The differences between early and late transcriptomes can be largely explained by the onset of sexual maturation, as gene clusters with high expression at late stages were significantly enriched for reproductive genes such as major sperm proteins and vitellogenins (Baskaran et al. 2015). However, because only five developmental stages were studied previously, our current data set represents the most highly resolved developmental transcriptome data of *P. pacificus*. This high-resolution catalog of gene expression in *P. pacificus* can assist in many future investigations and the comparative analysis of *P. pacificus* with other organisms. As the general features of the developmental transcriptome of *P. pacificus* were characterized previously, here, we focus on oscillatory gene expression, which can only be studied using the current data set with high temporal resolution.

Thousands of genes oscillate during postembryonic development

The cross-correlation analysis between different time points also revealed signals for periodic expression, suggesting that a fraction of genes show oscillating gene expression (Fig. 2A).

Indeed, when we performed PCA, we found that three of the four top principal components that explained 22% of the variation were associated with oscillatory patterns (Fig. 2B). PC1 reflects developmental genes with pronounced expression differences between early and late larval stages. In contrast, PC2–PC4 display an oscillatory pattern, which is in large parts synchronized with the molting cycle (Fig. 2B). In total, 2964 genes show oscillating expression patterns, whereas 20,171 are nonoscillating (Fig. 2C; Supplemental Fig. S3A). Most of the oscillating genes show up to three expression peaks (Fig. 2C), which is consistent with the fact that *P. pacificus* undergoes three larval molts after hatching from the egg. Further analysis also showed hundreds of peaks in the molting period, suggesting that the oscillating genes involved

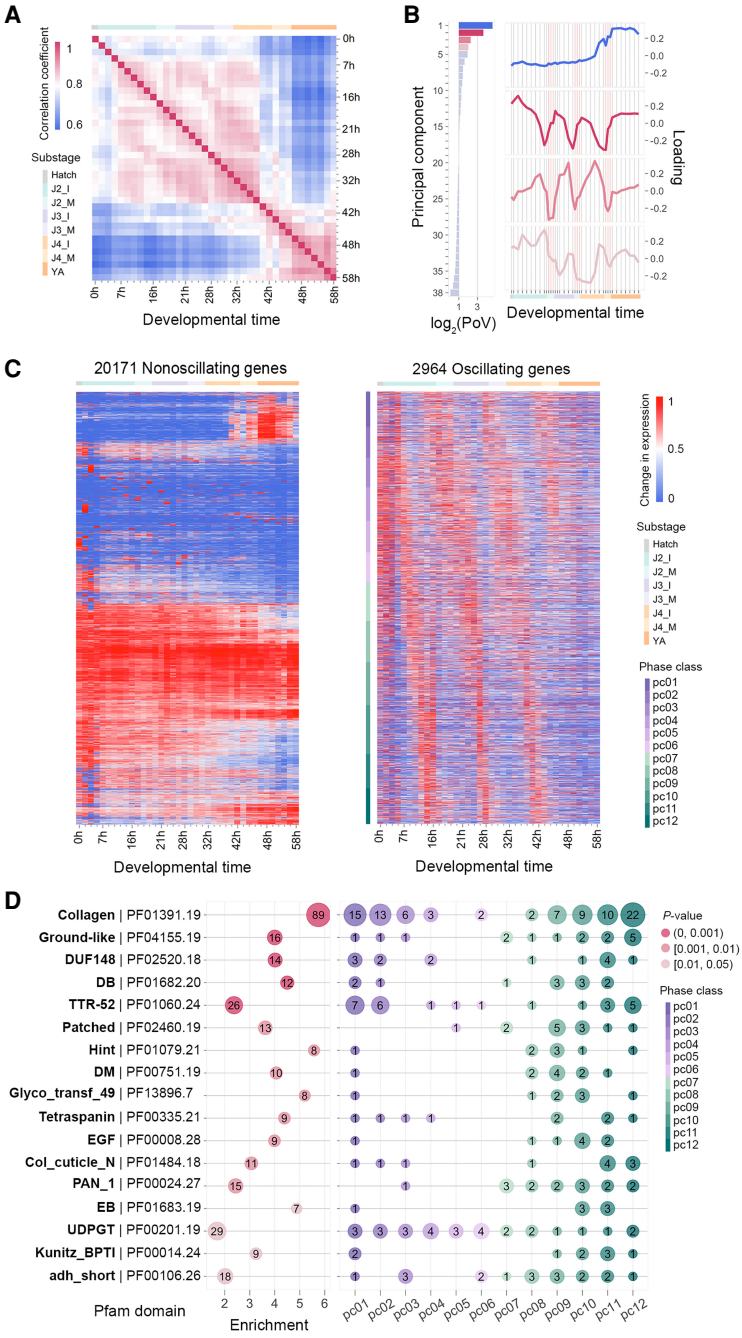

Figure 2. Thousands of oscillating genes are synchronized with the molting cycle in *P. pacificus*. (*A*) The heatmap displays the cross-correlation of $\log_2$-transformed expression patterns of all reliably expressed genes (reg; $n = 23{,}135$) obtained from the 38 time points. The time course was broken down into seven substages: J2 intermolt (J2_I), J2 molt (J2_M), J3 intermolt (J3_I), J3 molt (J3_M), J4 intermolt (J4_I), J4 molt (J4_M), and young adults (YA). (*B*) Proportion of variance (PoV) of gene expression by 38 principal components (PCs; *left* graph) and changes (loadings) of expression for each of the four top PCs (*right* graph). (*C*) The heatmaps visualize changes in gene expression throughout development for nonoscillating genes ($n = 20{,}171$; *left*) and oscillating genes ($n = 2964$; *right*). The mean TPM was calculated from three biological replicates and was normalized by its maximum of each expression trace. The order of rows in the heatmap with the oscillating genes was sorted by phase class, which indicated a phase difference of 30, corresponding to a peak shift by ~1 h. (*D*) Oscillating genes were tested for overrepresentation of protein domains. The *left* part shows the 17 protein domains that are most significantly enriched among oscillating genes, the *x*-axis presents the enrichment score, and the *y*-axis shows the name of protein domains. The size of circles corresponds to the number of oscillatory genes with a given protein domain, and the three different colors indicate significance levels as measured by Fisher's exact test. The *right* graph profiles the overrepresented protein domains according to the 12 phase classes. The number of oscillatory genes in each phase class is marked.

in these expression peaks have specific functions during molting (see Methods) (Supplemental Fig. S3E,F). However, the developmental oscillations are not totally synchronized but instead occur in multiple phases (Supplemental Fig. S3B).

Overrepresentation analysis of protein domains showed the strongest enrichment for collagens, which confirms the observed correlation to molting (Fig. 2D). We found an overrepresentation of collagens in 10 phase classes, suggesting that specific collagens are expressed at different time points of development and that not all collagens are exclusively expressed before molting. Indeed, visualization of collagen expression levels across development confirms the existence of "collagen classes" that are expressed at specific times in the intermolt cycle (Supplemental Fig. S3D). Specifically, 89 of 151 collagen genes of *P. pacificus* displayed oscillation patterns (Fig. 2D; Supplemental Fig. S3D), and oscillating collagens showed a significantly higher expression level than the nonoscillating collagens (Supplemental Fig. S3C). Also, 38 of these oscillating collagens showed expression peaks associated with molting events ($P < 10^{-4}$, Fisher's exact test) (Supplemental Fig. S3F). Taken together, our analysis indicates that thousands of genes in *P. pacificus* show oscillatory gene expression during postembryonic development with a strong enrichment of several protein domains including collagens.

Developmental oscillations are mediated by ancient gene classes

To test if developmental oscillations represent a highly conserved process or are driven by novel genes, we performed an analysis of gene ages of all oscillating genes. Specifically, we defined gene age classes based on the presence of orthologs (see Methods) in distantly related species using comparative genomic data of six *Pristionchus* species, two close relatives of the genera *Parapristionchus* and *Micoletzkya* and four more distantly related species, including *C. elegans* (Fig. 3A). The associated ladder-like phylogeny results in 13 branches (Fig. 3A, p01–p13) that separate different age classes and, thus, allows one to date the emergence of new genes. For example, p01–p03 contain genes only known from *P. pacificus* and its closest relatives *Pristionchus exspectatus* and *Pristionchus arcanus*. In contrast, genes in p12 and p13 are highly

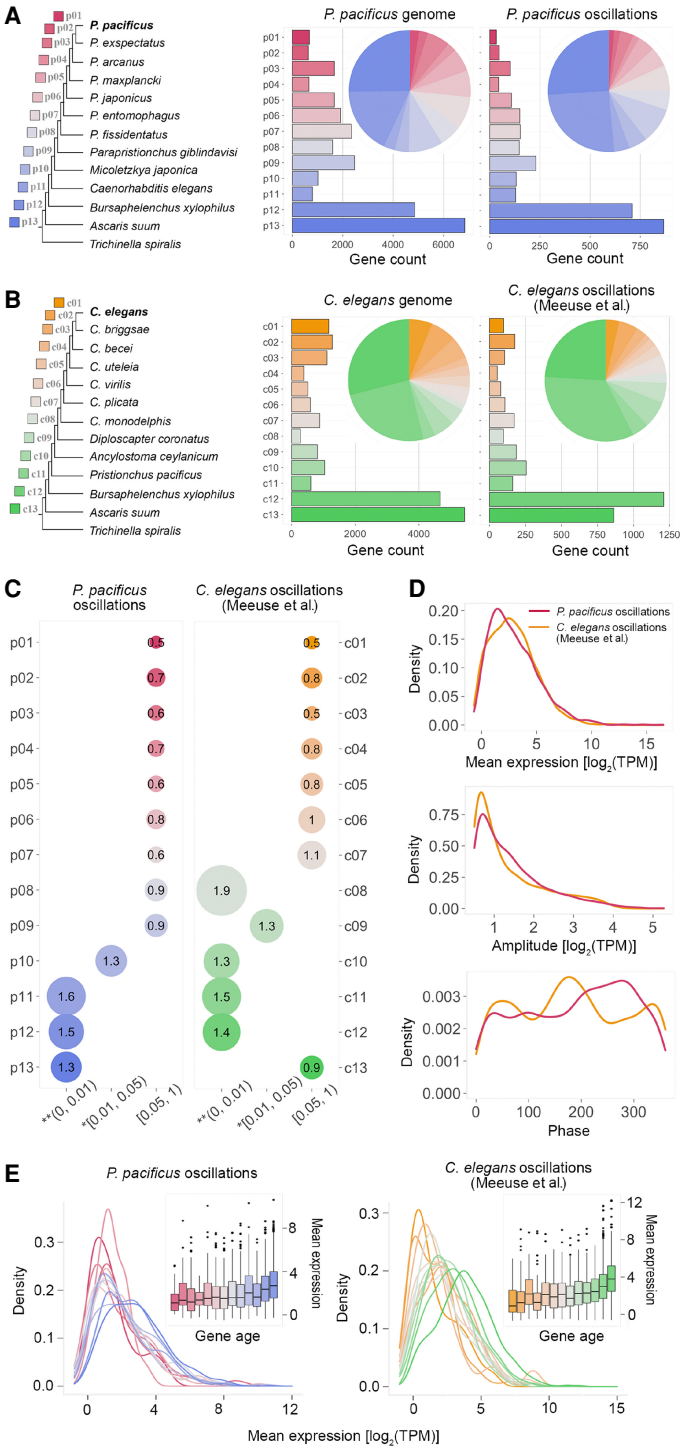

Figure 3. Ancient gene classes are enriched for oscillatory genes. (*A*) *P. pacificus* genes were assigned to age classes based on the most distant ortholog in a given orthology cluster. The age classes are labeled in a red (young)–blue (old) gradient. The distribution of gene number among the 13 age classes for the *P. pacificus* genome ($n = 28,896$) is shown in the *central* graph, and the distribution for the *P. pacificus* oscillating genes ($n = 2964$) is shown in the *right* part. (*B*) A similar analysis for the *C. elegans* oscillating genes set of Meeuse and coworkers (2020). (*C*) The graphs indicate the overrepresentation of gene age classes among the *P. pacificus* oscillatory genes (*left*) and the *C. elegans* oscillating genes sets of Meeuse et al. (2020). The size of the circles indicates the enrichment score; the *x*-axis shows the significance level measured by Fisher's exact test. (*D*) The plots show the distribution of average expression (*top*), amplitude (*middle*), and phase (*bottom*) of the oscillating genes in *P. pacificus* (red) and *C. elegans* (yellow). (*E*) The plots show the distribution of the mean expression of oscillating genes across 13 gene age classes in both species.

conserved and are even shared with distantly related nematodes of clade IV (*Bursaphelenchus xylophilus*) and clade iii (*Ascaris suum*). Gene age analysis indicates that most oscillating genes in *P. pacificus* are highly conserved in evolution (Fig. 3A). Further analysis reveals that oscillating genes in age classes p10, p11, and p12 are indeed strongly overrepresented, whereas younger gene classes contain fewer oscillating genes than expected by chance (Fig. 3C). Next, we reanalyzed the developmental transcriptomes of *C. elegans* and repeated the gene age analysis using seven *Caenorhabditis* species and a related outgroup, including *P. pacificus*, using the data of Kim et al. (2013) and Meeuse et al. (2020; Fig. 3B; Supplemental Fig. S3A). Again, we found that the majority of oscillating genes in *C. elegans* belong to older age classes, in this case in c08–c12 (Fig. 3C; Supplemental Fig. S3B). Unlike the divergence in the phase distribution between *P. pacificus* and *C. elegans*, the distribution of the average expression and amplitude are largely conserved in both species (Fig. 3D; Supplemental Fig. S3C). Genes in different age classes show different mean expression levels in both organisms. Specifically, genes in older age classes show a higher expression level in all three analyzed data sets (Fig. 3E; Supplemental Fig. S3D). Together, these results indicate that the majority of oscillating genes belong to old age classes and are conserved over larger evolutionary distance.

A core network of oscillating genes in both species

The observation that oscillatory genes are highly enriched in ancient gene classes does not necessarily indicate that the gene expression patterns are also conserved. To explicitly test this, we identified orthologous clusters between *P. pacificus* and *C. elegans*. Of the total 28,896 *P. pacificus* genes, only 6594 (22%) were found as 1:1 orthologs of *C. elegans* genes (Fig. 4A; Supplemental Fig. S6A), which reflects the high level of sequence divergence since the separation of both lineages ~100 mya (Prabh et al. 2018; Werner et al. 2018). Of the 2964 oscillating genes, 1002 (33.8%) have one-to-one orthologs (Fig. 4A; Supplemental Fig. S6B). In *C. elegans*, 1285 (34.4%) of the 3739 oscillating genes have one-to-one orthologs (Fig. 4A; Supplemental Fig. S6B). Comparing the 1:1 orthologous genes with oscillatory gene expression patterns in both species

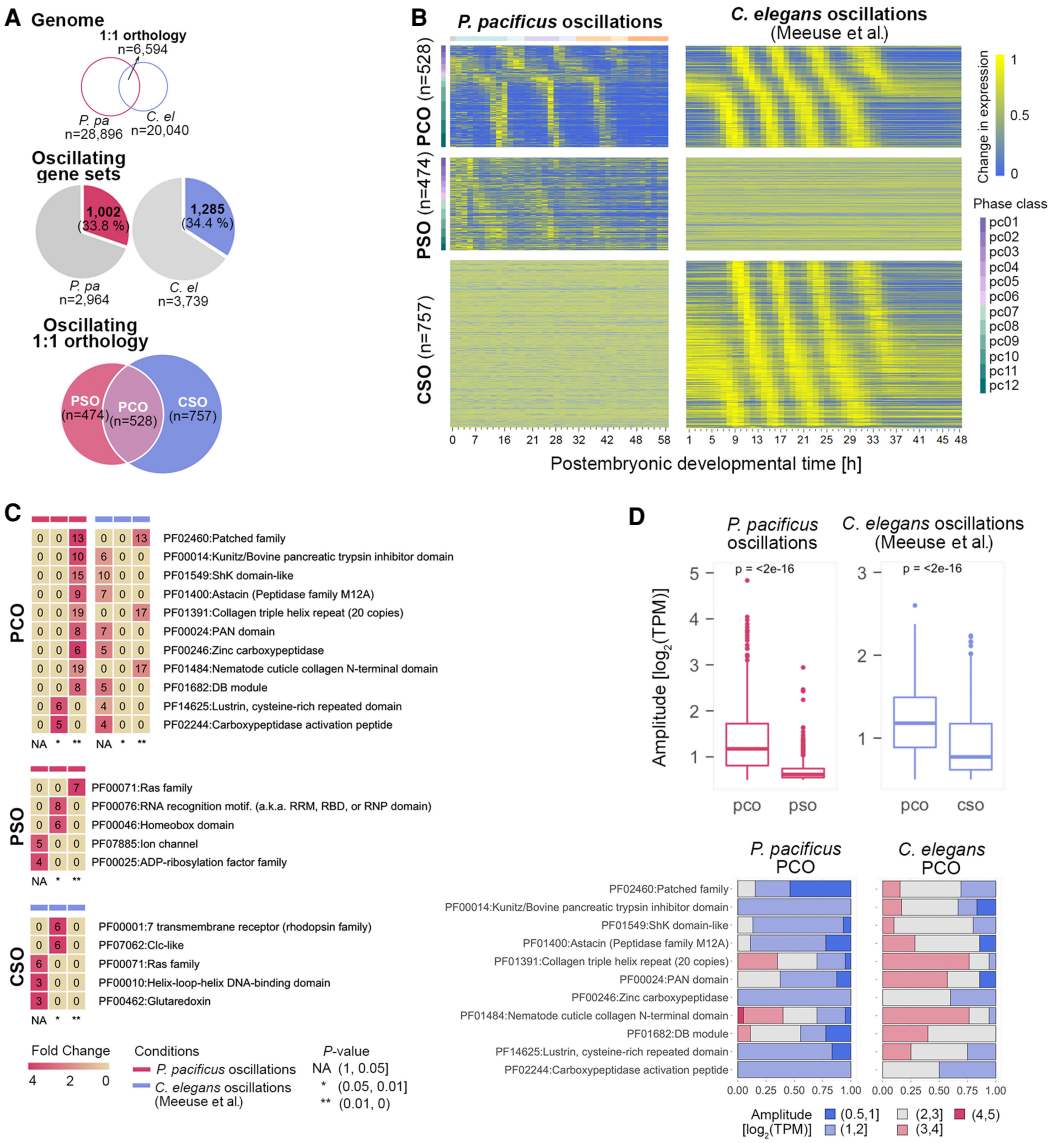

Figure 4. Orthologous oscillatory genes between *P. pacificus* and *C. elegans*. (*A*) A total of 6594 one-to-one orthologous gene pairs were identified between *P. pacificus* and *C. elegans*. In *P. pacificus*, the oscillating gene set contains 1002 one-to-one orthologous genes; in *C. elegans*, it includes 1285 one-to-one orthologous genes. Of these genes, 528 one-to-one orthologous genes are oscillating in both species (defined as *P. pacificus* conserved oscillations; PCO), 474 orthologous genes show a *P. pacificus*–specific oscillating pattern (PSO), and 757 orthologous genes were classified as *C. elegans*–specific oscillations (CSO). (*B*) The heatmaps show the developmental expression of the three gene classes (PCO, PSO, and CSO) in *P. pacificus* and *C. elegans*. (*C*) The matrices show the levels of significance for the tests of protein domain overrepresentation among the three oscillating gene classes. The color scale indicates the fold change; the labeling of matrix cells shows the number of oscillating genes containing a specific protein domain. (*D*) The box plots display the variation in the amplitude of oscillation between PCO and PSO in *P. pacificus* (red) and between PCO and CSO in *C. elegans* (blue). The bar plots show the distribution of genes with five different amplitude classes among top 11 protein domains highly enriched in PCO.

revealed that 528 (52.7%) of them showed conserved oscillation (Fig. 4A,B; Supplemental Fig. S5A–C). Even for oscillating *P. pacificus* genes with multiple orthologs in *C. elegans*, comparative analysis shows that in around half of the cases, at least one of the orthologs shows shared oscillation in *C. elegans* (Supplemental Fig. S6D,E). This level of conserved gene expression is in strong contrast to a recent analysis of spatial transcriptomes, where only a few dozens of one-to-one orthologs showed shared regional expression between both species (Rödelsperger et al. 2020).

The gene set of one-to-one orthologs with shared oscillatory expression likely represents a highly conserved core network mediating developmental oscillations in nematodes. We defined this gene set as the class of *P. pacificus*–*C. elegans* conserved oscillation (PCO) (Fig. 4A,B; Supplemental Table S7) and tested for overrepresentation of protein domains. This identified Patched receptors as one of the most highly enriched gene families (Fig. 4C). Patched receptors have been shown to have undergone gene family expansions early in the nematode phylum, and RNAi-mediated knockdown experiments in *C. elegans* revealed that they affect multiple aspects of development, including molting (Zugasti et al. 2005). Comparing the amplitude among PCOs, *P. pacificus*–specific oscillation (PSO) and *C. elegans*–specific oscillation (CSO) gene sets

revealed that highly conserved oscillatory patterns (PCO) displayed greater amplitude than both PSO and CSO (Fig. 4D; Supplemental Fig. S5C). The two protein domains collagen and nematode cuticle collagen showed a high percentage of genes with greater amplitudes, reflecting the essential role of collagens in the molting cycle (Fig. 4D). Taken together, our comparative analysis of oscillatory expression patterns identifies gene families with functional relevance in nematode development.

Regional expression analysis reveals a set of collagens with potential function in the cuticle remodeling of the head and tail region during molting

To determine where oscillatory gene expression occurred, we used the spatial transcriptomes of *P. pacificus* to study the spatial expression of oscillating genes (Rödelsperger et al. 2020). In total, we identified 380 genes with oscillating and regional expression (Supplemental Table S11), and these genes cover all 11 regions in *P. pacificus* (Fig. 5A). We found the oscillations presenting in the head region (P1, P2, and P3) and tail region (P11) displayed gene expression with higher mean expression and greater amplitude (Fig. 5A). Overrepresentation analysis of protein domains of these regional oscillatory genes showed the strongest enrichment for collagens. The 11 collagens specifically distributed in the head region (P2 and P3) and tail region (P11) and preferentially displayed in the molting-related phase classes (PC01, PC04, and PC12) (Fig. 5B,C), suggesting that these regional oscillating collagens presumably have an important function in the cuticle remodeling of head (including mouth-form and pharyngeal) and tail during the molting stages.

The novel gene *eud-1* was integrated as an upstream regulator of multiple oscillatory core genes

The major morphological differences between the two phenotypically plastic mouth-forms in *P. pacificus*, the St and the Eu form, suggest a role of multiple developmental genes in controlling these alternative ontogenic paths. The gene *eud-1*, which is the master regulator of mouth-form plasticity in *P. pacificus* (Ragsdale et al. 2013), is a novel gene that is derived from recent duplications in the diplogastrid lineage (Casasa et al. 2021). *eud-1* showed a graded expression pattern during larval development and is highly expressed at the second molt (Supplemental Fig. S7A). Hypothesizing that *eud-1* may regulate other developmental and possibly also oscillatory genes, we performed SWT on a *eud-1*-mutant line to identify differentially expressed genes at two selected timepoints (20 h and 28 h) (Supplemental Fig. S7B). In total, 205 genes were found to be significantly differentially expressed between the wild-type and *eud-1*-mutant conditions (Fig. 6A; Supplemental Table S12). When we compared this gene set with oscillatory genes, we found that 126 (61.5%) differentially expressed genes are indeed oscillatory genes ($P < 2.2 \times 10^{-16}$, Fisher's exact test) (Fig. 6B; Supplemental Fig. S7C). Moreover, 40 members of the oscillatory core network genes are among the set of differentially expressed genes (Fig. 5B). Specifically, 38 collagen genes that are regulated by *eud-1* showed oscillations in 11 different phases, and 11 of these 38 collagens are highly enriched in phase 2, which corresponds to the molting stage (Fig. 6C). These findings may indicate that the *eud-1*-dependent oscillatory collagen genes might represent mouth-form-specific collagens of *P. pacificus*. More generally, these findings imply that the taxonomically restricted *eud-1* gene had been integrated into an ancient regulatory network and controls the expression of highly conserved oscillatory genes.

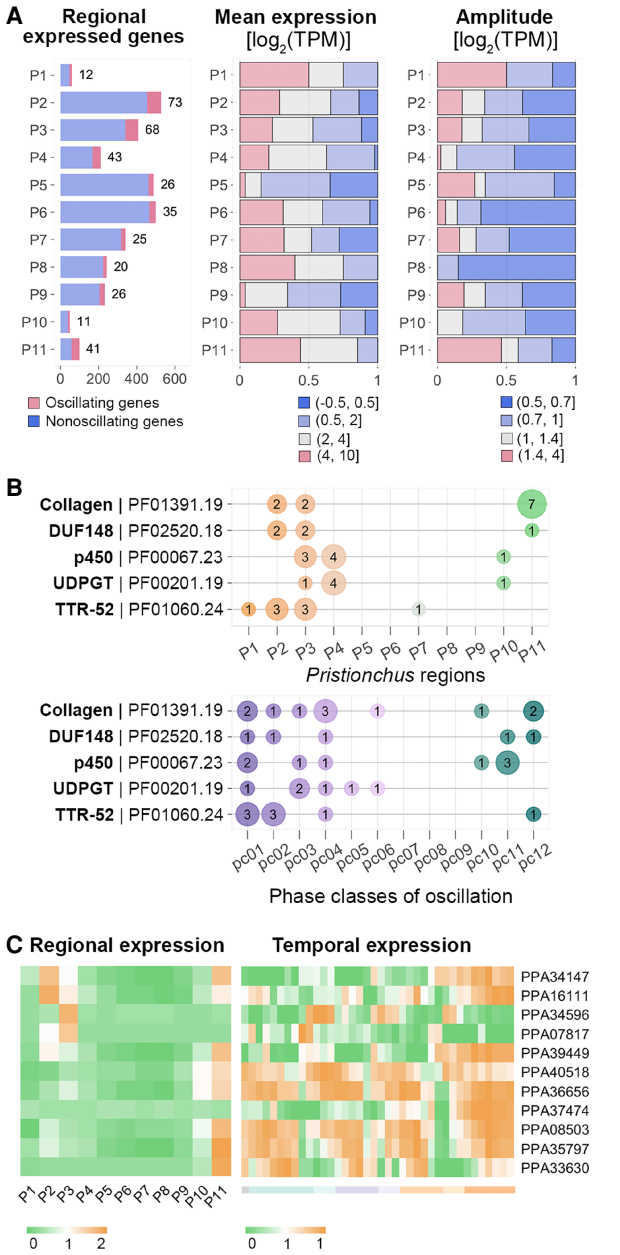

Figure 5. Oscillations are found in multiple *P. pacificus* body regions. (*A*) The bar plots show the numbers of oscillating genes across the 11 *P. pacificus* regions (*left*) as defined by spatial transcriptomics, the mean expression level of oscillatory genes with regional expression (*middle*), and four different amplitude classes (*right*). (*B*) The plots show regional expression and phases of the top five overrepresented protein domains. Circles indicate the number of oscillatory genes with expression in a specific region (*upper* part) and their assignment into 12 phase classes (*lower* part). (*C*) The *left* heatmap presents the median *z*-score-normalized expression values of the 12 oscillatory regional collagens across the 11 regions. The heatmap on the *right* shows their changes in gene expression throughout the entire postembryonic development.

Discussion

In this study, we have implemented SWT in the nematode model organism *P. pacificus*, providing a high-resolution developmental transcriptome with 38 time points from hatching of the J2 larvae

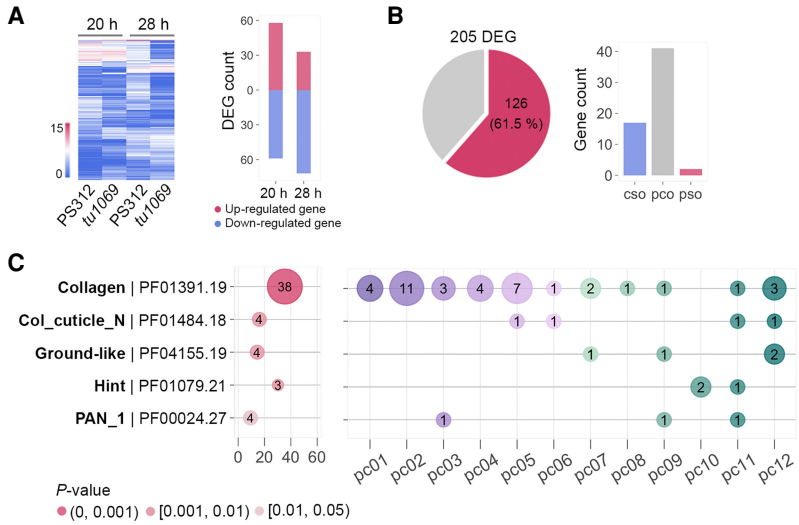

Figure 6. Integration of *eud-1* as upstream regulator of oscillatory core genes. (*A*) The heatmap shows the expression levels of 205 differentially expressed genes (DEGs) between the *eud-1(tu1069)*-mutant and wild-type (PS312) animals at 20 h and 28 h. The bar plot indicates the number of up-/down-regulated genes. (*B*) The pie chart displays the overlap between DEGs and oscillating genes. The bar plot shows the number of the oscillating DEGs among the three oscillating gene classes. (*C*) Protein domain enrichment analysis reveals that the collagen gene family is highly enriched in the 205 DEGs, the *x*-axis presents the enrichment score, and the *y*-axis shows the name of protein domains. The size of the circles corresponds to the number of oscillatory genes in a given protein domain, and the three different colors indicate different significance levels as measured by the Fisher's exact test. The *right* graph profiles the overrepresented protein domains according to the 12 phase classes. The number of oscillatory genes in each phase class is marked.

to young adults. This study combines methods development, comparative analysis, and the evolution of oscillating genes in nematodes. Nematodes are the largest phylum of multicellular animals with an expected number of 1–10 million species (Lambshead and Boucher 2003). Besides the well-studied model organism *C. elegans*, the nematodes contain many important parasites of plants, livestock, other animals, and humans. However, although many powerful tools including integrative genomic technologies are available in *C. elegans*, similar studies are often hard to perform in other nematodes. The small body size, the often-limited availability of specimen, and the inability to grow pure cultures under laboratory conditions are some of the major factors that make the application of integrative genomic techniques difficult in many species. After the original work of Macchietto and coworkers in *Steinernema* (2017), our study establishes an updated protocol of SWT for an unrelated nematode. This protocol was inspired by our need for high-throughput investigations at lower costs. In particular, our work on phenotypic plasticity requires the parallel investigation of large numbers of samples for environmental perturbation and in experimental evolution approaches. Although our read depth is somewhat lower than in the original protocol, the representation of gene expression is comparable between SWT and STD approaches. Therefore, this protocol will allow scaling up of transcriptomic approaches to large sample sizes at low cost. Also, we are confident that this protocol can be applied to the majority of nematode species, if live material of different post-embryonic stages can be obtained. Such transcriptomic maps may be of importance for developing drug targets against parasites. Additionally, they may potentially help in diagnostics. Besides these technical aspects, this study has two major conclusions.

First, the gene expression atlas of 38 developmental time points provides an important reference point for future studies in *P. pacificus*. Given the easiness of the methodology, selected developmental stages of mutant animals can be compared to the wild-type catalog presented in this study. This will represent a powerful tool for many mechanistic studies in which the selection of appropriate developmental stages will be superior over mixed stages. For example, the comprehensive analysis of mouth-form plasticity and its genetic, epigenetic, and environmental regulation has resulted in a large number of mutants that can now be compared by SWT (for review, see Sommer 2020). The differential expression analysis of the mouth-form developmental switch gene *eud-1* (Fig. 5) can serve as a proof of principle for future studies. Similarly, changing abiotic and biotic environmental conditions, including bacterial diet (Werner et al. 2017; Akduman et al. 2020), and their influence on gene expression can now be studied with high accuracy. Thus, our high-resolution catalog of gene expression can assist in many future investigations as well as in the comparative analysis of *P. pacificus* with other organisms.

Second, and most importantly, this study focused on developmental oscillations in gene expression. Although the finding of oscillating gene expression per se is not at all surprising, the comparison of these patterns between *P. pacificus* and *C. elegans* allows for the first time to add an evolutionary dimension to developmental oscillations in nematodes. In general, genomic and transcriptomic comparisons between *P. pacificus* and *C. elegans* build on a large history (Dieterich et al. 2008; Rödelsperger et al. 2014, 2017, 2018; Prabh et al. 2018). But most importantly, the work by Prabh and coworkers (2018) established comparative genomic analysis of six *Pristionchus* species with more distant relatives, including *C. elegans* (Fig. 3A). The associated ladder-like phylogeny allows the distinction of different age classes and the emergence of new genes. Oscillatory genes are highly enriched in ancient gene classes and allowed us to identify a highly conserved core network of 528 one-to-one orthologous genes. This *P. pacificus–C. elegans* shared oscillation (PCO) network will be useful for future studies when selecting individual genes for the identification of the regulatory aspects of expression oscillations and their conservation.

Perhaps the most unexpected finding is the diversity of collagen expression in *P. pacificus*. For nematodes with their exoskeleton, collagens are one of the most important gene families. The current 151 collagen genes annotated in the genome can be grouped into two different classes. First, 62 of these genes do not show oscillatory gene expression. Second, among the 89 oscillating collagen genes, subsets expressed in 10 distinct phase classes can be identified (Supplemental Fig. S2). Of the 89 oscillating collagen genes, 38 showed expression peaks during the molting periods (Supplemental Fig. S3F), suggesting that the synthesis of phase-specific collagens could make different contributions to cuticle growth during the molting and intermolt periods. One unexpected finding was the observation that *eud-1*-dependent oscillating collagens are expressed in different phases and that

11 of these collagens show a strong enrichment in molting (phase class 2). Future studies might indicate that such collagens have an important role during mouth-form development in *P. pacificus* and might have a specific function in the stoma. In any case, the comparison of wild-type and *eud-1*-mutant animals reveal the power of SWT for future studies in developmental biology.

Methods

Maintenance of worm cultures

All studies reported in this manuscript have been conducted by using the wild-type strain *P. pacificus* PS312, originally isolated in Pasadena, California, in 1988 (Sommer et al. 1996). In *P. pacificus*, we use the term juvenile and designate stages as J2, J3, and J4, as typical for nematodes. Note that in *C. elegans*, juveniles are named L1–L4 for historical reasons. Stock cultures of two strains used in this study, PS312 and *eud-1*(*tu1069*), were reared at 20°C on nematode growth medium (NGM) (Sieriebriennikov et al. 2020).

Nematode synchronization and collection

Bleached egg populations were seeded onto plates to obtain synchronized worm cultures (Werner et al. 2017). Synchronized eggs-J1 were observed under the Discovery V20 microscopy and at hatching (Supplemental Fig. S1A) were isolated to a new 6-cm agar plate with 50 µL OP50 as starting time point. For later juvenile stages, the ecdysis (Supplemental Fig. S1D,G,J) within each molt was chosen as starting time point for animal collection. At the desired time point, single animals were picked into one PCR tube containing 3 µL nuclear-free water and transferred into liquid nitrogen immediately (Fig. 1C). Such samples could be stored at −80°C for up to a month. In total, worms were collected at 38 time points with three biological replicates, covering the entire postembryonic development and young adulthood (Fig. 1B).

SWT library preparation

We started sample preparation by 3× freeze-thawing with liquid nitrogen. Afterward, 3 µL of lysis buffer (low input cDNA synthesis & amplification module; NEB E6421S) was added, and samples were incubated for 40 min at 65°C, followed by 1 min at 85°C. To remove genomic DNA, we used the TURBO DNA-free kit (Invitrogen AM1907), which allows efficient digestion of DNA within 20 min. Also, the genome digest is performed in the same PCR tube without cleaning up the sample through a column. Importantly, we replaced the 10 × TURBO DNase buffer with the 5 × SuperScript II first-strand buffer (Thermo Fisher Scientific, Invitrogen 18064014), owing to the former containing higher concentration of Mg^{2+}, which might inhibit the efficiency of reverse transcription. In the next step, we applied RNA-SPRI beads (Agencourt RNAClean XP, protocol 001298v001) to extract mRNA and clean up the samples. A Single Cell/Low Input cDNA Synthesis & Amplification Module (NEB E6421S) was used to generate full-length cDNA and cDNA amplification with 15 cycles of PCR amplification. The Nextera DNA Flex Library Prep Kit (Illumina) was used to prepare the DNA library (Fig. 1C). To avoid potential amplification biases by different PCR cycles, we used 12 PCR cycles for cDNA amplification and eight PCR cycles for DNA (library) amplification in all samples. Together, this protocol contains four major modifications relative to the work of Macchietto et al. (2017) and Chang et al. (2021): (1) We lysed worms through freeze-thawing; (2) we removed genomic DNA and performed this reaction in the same PCR tube as the cDNA synthesis reaction; (3) for the library preparation, we reduced the input of cDNA (to ~1 ng); and (4) we only amplified the final library through eight cycles of PCR reaction. Raw reads of the 114 SWT samples were submitted to the European Nucleotide Archive (see Data access).

STD library preparation

To prepare three STD libraries for comparison to SWT (Fig. 1F), we used the lethargus stage of the fourth molt. We observed individual worms at lethargus (Supplemental Fig. S1H) under the discovery microscopy and picked appropriate animals to one PCR tube containing 3 µL nuclease-free water. After 3 × freeze-thawing steps with liquid nitrogen, more than 70 PCR tubes containing single animals at the lethargus stage were pooled into one 1.5-mL tube as one replicate. Three independent biological replicates were prepared in different experiments. To extract RNA, worms suspended in TRIzol were frozen and thawed three times in liquid nitrogen, debris was pelletized for 10–15 min at 14,000 rpm at 4°C, and 200 µL of chloroform was added to the supernatant. After vigorous vortexing and incubation at room temperature for 5 min, tubes were microfuged for 15 min at 14,000 rpm at 4°C. The aqueous phase was combined with an equal volume of 100% ethanol, RNA was purified using RNA Clean & Concentrator kit (Zymo Research), and its integrity was verified using RNA Nano chips on the Bioanalyzer 2100 instrument (Agilent). We built libraries using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs). Raw sequencing of the three newly generated STD samples were submitted to the European Nucleotide Archive (see Data access).

Transcriptome data for comparative analysis across protocols and species

This study contains 38 substages covering the first 58 h after hatching of the J2 larvae with a 1- to 2-h resolution. Three substages in each molt, and seven substages after the final molt provide information about transcriptional changes during the molting stage and within young adult animals (Supplemental Table S1). To estimate the similarities of our SWT and STD libraries, we used three SWT samples of the lethargus during the fourth molt and the three STD samples as described above. To examine how much sequencing depth is required to detect most of the expressed genes for SWT and STD libraries, we selected one library containing more than 12 million read pairs from SWT and STD data set; subsampled the raw reads to nine subsamples with different sequencing depth covering 0.5 million, 1 million, 2 million, 3 million, 4 million, 6 million, 8 million, 10 million, and 12 million; and profiled the distribution of number of detected genes among the nine classes of sequencing depth (Supplemental Fig. 2B). For comparison with *C. elegans*, we used the data sets of Kim et al. (2013) with a time course containing 26 samples at 20°C (sample DH2 for L1–L3, sample DH5 for L4) covering 48 h of postembryonic development (Supplemental Table S1). Raw reads were downloaded from the European Nucleotide Archive (ENA; https://www.ebi.ac.uk/ena/browser/home) accession PRJNA212741. Additionally, we compared our data to the one of Meeuse et al. (2020) with a time course including 48 samples at 25°C (TP1–TP13(TC1, L1–L2), TP14–TP48(TC2, L1–YA)) (Supplemental Table S1). Raw reads were obtained from the NCBI Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra) accession SRP195783.

RNA-seq read alignment and estimation of relative abundance

All libraries were sequenced at 150-bp PE on an Illumina HiSeq 3000. Raw reads were trimmed by the program cutadapt (Martin 2011) version 2.10 with parameters "-q 30,25 --gc-content = 50, --minimum-length 25:25." Read pairs were aligned by the HISAT2

(version 2.1.0) (Kim et al. 2015) software to the *P. pacificus* genome assembly (version El Paco) using the gene annotations (version El Paco gene annotation 3) as additional information. HISAT2 was run with the additional parameter "--rna-strandness RF." Mapped reads were summarized to the genomic features by the featureCounts (version 1.6.4) with parameters "-a -F GTF -g Parent -d 50 -D 500." Ninety-five percent expressed genes with 12 million PE reads were detected with at least 2 million PE reads; we removed the six libraries with fewer than 2 million PE reads in the further analysis, including ENA samples ERS5595844, ERS5595845, ERS5595849, ERS5595862, ERS5595942, ERS5595945. The six time points (1 h, 3 h, 5 h, 13 h, 52 h, and 54 h) contained two replicates; another 32 time points included three replicates with enough sequencing depth. Gene expression value, transcripts per million (TPM), was computed through normalization by library size and feature effective length. The mean TPM was calculated from three biological replicates; 23,135 REGs were identified as their mean PE reads are greater than three in at least one of the 38 time points (Supplemental Table S2).

Raw reads of the two *C. elegans* time courses were mapped to the *C. elegans* genome (WS275) using HISAT2 (version 2.1.0) with parameters "--rna-strandness F." Mapped reads were summarized to the longest isoforms by the featureCounts (version 1.6.4) (Liao et al. 2014) with parameters "-a -F GTF -g Parent -M." Gene expression value (TPM) was calculated by the package t-arae/ngscmdr (version 0.1.0.181203) in R (R Core Team 2020). Genes with more than 10 read counts in at least one time point were named as REGs, 17,446 REGs and 17,409 REGs were identified in the study of Kim et al. (2013) and Meeuse et al. (2020), respectively.

Pairwise correlation analyses and principal component analysis

$\log_2$-transformed gene expression levels of 23,135 REG were used to compute the pairwise correlations between transcriptomes at different developmental time points using the R command cor (data, use = "pairwise.complete.obs", method = "Pearson"). To further show the strong temporal relationship between our samples, we performed a PCA using the function princomp (Sigg and Buhmann 2008) in R with default parameters. The loadings corresponding to the PC2, PC3, and PC4 appeared to be periodical patterns.

Classification of oscillatory genes by Meta2D

To classify the oscillatory genes in *P. pacificus*, we applied the MetaCycle (Wu et al. 2016) package (version 1.2.0) in R, 20 time points from 5 h to 44 h covering the entire J3 stage with 2-h intervals were picked to predict oscillation genes. The parameters (cycMethod = c ("ARS", "JTK"), minper = 7, maxperiod = 13, ARSdefaultPer = 10, weightedPerPha = TRUE) were used to perform the meta2d algorithm. We predicted the oscillations for the Kim et al. (2013) data sets with the parameters (cycMethod = c("ARS", "JTK"), minper = 6, maxperiod = 12, ARSdefaultPer = 9, weightedPerPha = TRUE) considering this time course under 20°C. We reanalyzed the oscillating genes in the Meeuse et al. (2020) data sets by using the same parameters. In total, we identified 2964 oscillations in *P. pacificus* (Supplemental Table S3) and 2221 oscillations in the Kim et al. (2013) data set (Supplemental Table S4) by using the two filters, amplitude ≥ 0.5 and FDR < 0.05. 3739 oscillations that classified by Meeuse et al. (2020) were used to perform further comparison.

Identification of gene expression peaks

To identify the peaks of temporal gene expression during the time-course, we applied the "findpeaks" function (with nups = 2, ndowns = 2) of the package "pracma" (version 2.2.9; #37) in R. To quantify the number of the expression peaks coinciding with the molting cycle, we defined the molting periods as the intervals spanning 1 h before until 1 h after molting. These time points are for the second molt "TP-2M" (16 h, 17 h, 18 h, 19 h, and 20 h), the third molt "TP-3M" (28 h, 29 h, 30 h, 31 h, 32 h), and the fourth molt "TP-4M" (42 h, 43 h, 44 h, 45 h, 46 h). Similarly, they are for the intermolt periods "TP-2I" (3 h, 5 h, 7 h, 9 h, 11 h, 13 h, 15 h), "TP-3I" (21 h, 22 h, 24 h, 26h) and "TP-4I" (34 h, 36 h, 38 h, 40 h). Time points for young adult (TP-YA) include 48 h, 50 h, 52 h, and 54 h. To study the distribution of oscillation peaks during three molting events, we defined seven different classes based on the peak distribution. C1, C2, or C3 indicates oscillating genes that show only one peak during TP-2M, TP-3M, or TP-4M, respectively. C4, C5, or C6, represents oscillating genes that display two peaks during TP-2M and TP-3M, TP-2M and TP-4M, or TP-3M and TP-4M, respectively. C7 indicates oscillating genes that show three peaks during the three molting periods.

Pfam domain prediction and overrepresentation analysis in *P. pacificus* oscillation genes

To predict Pfam domains of all genes in the *P. pacificus* genome, we applied the HMMER (Mistry et al. 2013) version 3.3 with parameters "hmmscan --tblout Pfam-A.hmm"; 17,664 of all 28,896 genes contain predicted Pfam domains (*e*-value of full seq < 0.01) (Supplemental Table S8). Overrepresented Pfam domains in the set of oscillating genes were determined by calculating the fold enrichment of the number of overlapping genes compared to what to expect by chance given the number of genes in a particular Pfam domains and the number of oscillating genes with only expressed genes considered. To minimize the large enrichments that would otherwise be caused by Pfam domains with small numbers of genes, we added a pseudocount of 12 to the total number of genes for a given Pfam domain before calculating the actual ratio. The *P*-value was generated by the Fisher's exact test.

Gene age analysis

For *C. elegans,* we performed gene age analysis by obtaining annotated protein sets for *C. elegans* (WS275), *Caenorhabditis briggsae* (WS275), *Caenorhabditis becei* (QG2083_v1), *Caenorhabditis uteleia* (JU2585_v1), *Caenorhabditis virilis* (JU1968_v1), *Caenorhabditis plicata* (SB355_v1), *Caenorhabditis monodelphis* (JU1667_v1), *Diploscapter coronatus* (WBPS14), *Ancylostoma ceylanicum* (WS248), *P. pacificus* (El_Paco_v3), *B. xylophilus* (WS248), *A. suum* (WS248), and *Trichinella spiralis* (WS248) from WormBase (https://wormbase.org) and from http://www.pristionchus.org. The set of *C. elegans* proteins (the longest isoform per gene) was taken as query sequences, and gene age classes were defined from the OrthoFinder (version 2.3.12) (Emms and Kelly 2020), DIAMOND (version 0.9.31, with parameters BLASTP -e 0.001) (Buchfink et al. 2015) searches against all other species and clustered orthogroups by the MCL algorithm (Enright et al. 2002) with default parameters; 27,132 (93.9%) *P. pacificus* annotated genes and 18,750 (93.6%) *C. elegans* annotated genes were assigned into orthogroups, respectively (Supplemental Fig. S3E). Further, we used a phylostratigraphy approach (Domazet-Lošo et al. 2007; Prabh et al. 2018; Prabh and Rödelsperger 2019; Rödelsperger et al. 2020), to trace the origin of the assigned genes on the phylogeny in *P. pacificus* and *C. elegans*. Based on the presence of the most distant outgroup species in a given orthologous cluster, we identified 13 gene age classes in both species (Fig. 3A,B; Supplemental Tables S5, S6). Note that singletons or the unsigned genes in orthogroup (6.1% of *P. pacificus genes* and 6.4% of *C. elegans* genes) were removed from this analysis considering they are

different from any gene age class as in previous analysis (Prabh et al. 2018). Gene age analysis for *P. pacificus* was performed using the nine data sets of predicted proteins from the recently sequenced diplogastrid genomes (Prabh et al. 2018) and proteins from *C. elegans*, *B. xylophilus*, *A. suum*, and *T. spiralis*.

One-to-one ortholog prediction between *P. pacificus* and *C. elegans*

For *C. elegans*, 20,040 protein sequences were extracted from the gene annotation (WS275) by taking the longest isoform of each annotated gene. Similarly, 28,896 protein sequences of the last version of *P. pacificus* (El_Paco_v3) were used to identify the one-to-one orthologs. We applied the OrthoFinder (version 2.3.12) program to prediction, DIAMOND (version 0.9.31, with parameters BLASTP -e 0.001) searches against protein data from both species were performed, and the orthogroups were clustered by the MCL algorithm with default parameters. In total, 6594 one-to-one orthologs, 529 one-to-many (*Ppa* vs. *Cel*) orthologs, 756 one-to-many (*Cel* vs. *Ppa*) orthologs, and 187 many-to-many (*Ppa* vs. *Cel*) orthologs were classified from the resulting orthologous clusters.

Gene Ontology (GO) analysis for orthologous oscillatory genes

We applied the DAVID (Dennis et al. 2003) v6.8 web-accessible programs to do the Gene Ontology (GO) analysis by using the UniProt accessions of the *C. elegans* genes, as obtained from the gene annotation file (WS275). For UniProt accession annotation of *P. pacificus* genes, we first applied DIAMOND (version 0.9.31, with parameters BLASTP -e 0.001) searches against protein sequences from both species and obtained the best-reciprocal pairs with lowest cross-species e-value, and then we annotated the UniProt accession of *P. pacificus* gene based on the UniProt accession of *C. elegans* gene that is the best-reciprocal hit. The GO analysis includes biological process (BP), molecular function (MF), and cellular component (CC), and we used gene enrichment and EASE scores from the output of DAVID for visualization (Supplemental Fig. S4D; Supplemental Table S9).

Regional expression analysis of oscillatory genes

The spatial transcriptomic analysis defined 11 regions across the anterior posterior axis, which correspond to anatomical structures such as head, intestine, germline, and tail (Rödelsperger et al. 2020). The corresponding gene sets were defined based on a relative enrichment over the mean expression (z-score > 1). As the spatial transcriptome was analyzed for a reference annotation that was based on a strand-specific transcriptome assembly (Rödelsperger et al. 2016, 2018), we applied the following set of rules to assign the genes of the last version of *P. pacificus* (El_Paco_v3): (1) applied DIAMOND (version 0.9.31, with parameters blastx -e 0.00001) searches against protein sequences of the last version of *P. pacificus* (El_Paco_v3) and obtained the best-hits with the lowest e-value, (2) used BEDTools (version 2.26.0, with parameters intersect -wa -wb -s) to confirm the same genomic location, and (3) selected the longest predicted transcript as the final corresponding locus to the El_Paco version 3. This could automatically assign 3155 of the 3502 regional genes to the corresponding gene identifiers in the latest annotation (El Paco gene annotation, version 3) (Supplemental Table S10). We identified 380 *P. pacificus* genes with oscillatory and regional expression by overlapping the 2964 oscillations in this study and 3155 regional expression genes (Supplemental Table S11).

Differential expression analysis

To identify differentially expressed genes regulated by *eud-1*, we picked two time points, 20 h (the first hour after the second molt) and 28 h (the last hour before the third molt), to perform the gene expression comparison between the *eud-1(tu1609)* mutant and wild type (PS312). These experiments were performed independent of the wild-type SWT for logistic reasons. Differential expression analysis was performed in R (version 4.0.3) using DESeq2 (version 1.18.1) (Love et al. 2014). We applied an adjusted *P*-value cut-off of 0.05 and a fold change cutoff of two to identify the DEG genes. Raw reads of the six SWT samples of the *eud-1 (tu1069)* mutant were submitted to the European Nucleotide Archive (see Data access).

Data access

Raw and processed data sets from this study have been submitted to the European Nucleotide Archive (ENA; https://www.ebi.ac.uk/ena) under accession numbers PRJEB42613, PRJEB42633, and PRJEB42635.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank the members of the Sommer laboratory for thoughtful comments of experiments, results, and interpretation. We thank Dr. Wen-Sui Lo for her assistance with the bioinformatic analysis and Dr. Kohta Yoshida for his help with the SWT protocol. S.S. was funded by a fellowship from the Chinese Sponsorship Council (CSC). This work was funded by the Max-Planck Society.

Author contributions: S.S. and R.J.S. conceived and designed all experiments. S.S. conducted all experimental studies. S.S. and C.R. performed all bioinformatic analysis. All three authors contributed to writing.

References

Akduman N, Lightfoot JW, Röseler W, Witte H, Lo WS, Rödelsperger C, Sommer RJ. 2020. Bacterial vitamin B_{12} production enhances nematode predatory behavior. *ISME Journal* **14:** 1494–1507. doi:10.1038/s41396-020-0626-2

Athanasouli M, Witte H, Weiler C, Loschko T, Eberhardt G, Sommer RJ, Rödelsperger C. 2020. Comparative genomics and community curation further improve gene annotations in the nematode *Pristionchus pacificus*. *BMC Genomics* **21:** 708. doi:10.1186/s12864-020-07100-0

Baskaran P, Rödelsperger C, Prabh N, Serobyan V, Markov G, Hirsekorn A, Dieterich C. 2015. Ancient gene duplications have shaped developmental stage-specific expression in *Pristionchus pacificus*. *BMC Evol Biol* **15:** 185. doi:10.1186/s12862-015-0466-2

Bento G, Ogawa A, Sommer RJ. 2010. Co-option of the hormone-signalling module dafachronic acid-DAF-12 in nematode evolution. *Nature* **466:** 494–497. doi:10.1038/nature09164

Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12:** 59–60. doi:10.1038/nmeth.3176

Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ, et al. 2017. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357:** 661–667. doi:10.1126/science.aam8940

Casasa S, Biddle JF, Koutsovoulos GD, Ragsdale EJ. 2021. Polyphenism of a novel trait integrated rapidly evolving genes into ancestrally plastic networks. *Mol Biol Evol* **38:** 331–343. doi:10.1093/molbev/msaa235

Chang D, Serra L, Lu D, Mortazavi A, Dillman A. 2021. A revised adaptation of the smart-Seq2 protocol for single-nematode RNA-seq. *Methods Mol Biol* **2170:** 79–99. doi:10.1007/978-1-0716-0743-5_6

Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. 2003. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* **4:** R60. doi:10.1186/gb-2003-4-9-r60

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:** 1193–1198. doi:10.1038/ng.227

Domazet-Lošo T, Brajković J, Tautz D. 2007. A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends Genet* **23:** 533–539. doi:10.1016/j.tig.2007.08.014
Emms D, Kelly S. 2020. OrthoFinder manual: accurate inference of orthologues and orthogroups made easy! *Genome Biol* **20:** 238.
Enright AJ, Dongen Sv, Ouzounis CA. 2002. An efficient algorithm for large-scale detection of protein families. *Nucl. Acid Res*. **30:** 1575–1584. doi:10.1093/nar/30.7.1575
Félix MA, Hill RJ, Schwarz H, Sternberg PW, Sudhaus W, Sommer RJ. 1999. *Pristionchus pacificus*, a nematode with only three juvenile stages, displays major heterochronic changes relative to *Caenorhabditis elegans*. *Proc Roy Soc B* **266:** 1617–1621. doi:10.1098/rspb.1999.0823
Fürst von Lieven A. 2005. The embryonic moult in diplogastrids (Nematoda): homology of developmental stages and heterochrony as a prerequisite for morphological diversity. *Zool Anz* **244:** 79–91. doi:10.1016/j.jcz.2005.05.001
Gerstein MB, Rozowsky J, Yan K-K, Wang D, Cheng C, Brown JB, Davis CA, Hillier L, Sisu C, Jessica Li J, et al. 2014. Comparative analysis of the transcriptome across distant species. *Nature* **512:** 445–448. doi:10.1038/nature13424
Han Z, Lo WS, Lightfoot JW, Witte H, Sun S, Sommer RJ. 2020. Improving transgenesis efficiency and CRISPR-associated tools through codon optimization and native intron addition in *Pristionchus* nematodes. *Genetics* **216:** 947–956. doi:10.1534/genetics.120.303785
Hendriks GJ, Gaidatzis D, Aeschimann F, Großhans H. 2014. Extensive oscillatory gene expression during *C. elegans* larval development. *Mol Cell* **53:** 380–392. doi:10.1016/j.molcel.2013.12.013
Kanzaki N, Giblin-Davis RM. 2015. Diplogastrid systematics and phylogeny. In *Pristionchus pacificus, a nematode model for comparative and evolutionary biology* (ed. Sommer RJ), pp. 43–76. Brill, Leiden.
Kanzaki N, Herrmann M, Weiler C, Röseler W, Theska T, Berger J, Rödelsperger C, Sommer RJ. 2021. Nine new *Pristionchus* (nematoda: Diplogastridae) species from China. *Zootaxa* **4943:** zootaxa.4943.1.1. doi:10.11646/zootaxa.4943.1.1
Kim DH, Grün D, van Oudenaarden A. 2013. Dampening of expression oscillations by synchronous regulation of a microRNA and its target. *Nat Genet* **45:** 1337–1344. doi:10.1038/ng.2763
Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12:** 357–360. doi:10.1038/nmeth.3317
Lambshead PJD, Boucher G. 2003. Marine nematode deep-sea biodiversity: hyperdiverse or hype? *J Biogeogr* **30:** 475–485. doi:10.1046/j.1365-2699.2003.00843.x
Levin M, Hashimshony T, Wagner F, Yanai I. 2012. Developmental milestones punctuate gene expression in the *Caenorhabditis* embryo. *Dev Cell* **22:** 1101–1108. doi:10.1016/j.devcel.2012.04.004
Liao Y, Smyth GK, Shi W. 2014. Sequence analysis feature counts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30:** 923–930. doi:10.1093/bioinformatics/btt656
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15:** 550. doi:10.1186/s13059-014-0550-8
Macchietto M, Angdembey D, Heidarpour N, Serra L, Rodriguez B, El-Ali N, Mortazavi A. 2017. Comparative transcriptomics of *Steinernema* and *Caenorhabditis* single embryos reveals orthologous gene expression convergence during late embryogenesis. *Genome Biol Evol* **9:** 2681–2696. doi:10.1093/gbe/evx195
Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17:** 10–12. doi:10.14806/ej.17.1.200
Meeuse MW, Hauser YP, Morales Moya LJ, Hendriks G, Eglinger J, Bogaarts G, Tsiairis C, Großhans H. 2020. Developmental function and state transitions of a gene expression oscillator in *Caenorhabditis elegans*. *Mol Syst Biol* **16:** e9498. doi:10.15252/msb.209975
Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res* **41:** e121. doi:10.1093/nar/gkt263
Nakayama K-I, Ishita Y, Chihara T, Okumura M. 2020. Screening for CRISPR/Cas9-induced mutations using a co-injection marker in the nematode *Pristionchus pacificus*. *Dev Genes Evol* **230:** 257–264. doi:10.1007/s00427-020-00651-y
Prabh N, Rödelsperger C. 2019. *De novo*, divergence, and mixed origin contribute to the emergence of orphan genes in *Pristionchus* nematodes. *G3 (Bethesda)* **9:** 2277–2286. doi:10.1534/g3.119.400326
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:** 1664–1674. doi:10.1101/gr.234971.118
Ragsdale EJ, Müller MR, Rödelsperger C, Sommer RJ. 2013. A developmental switch coupled to the evolution of plasticity acts through a sulfatase. *Cell* **155:** 922–933. doi:10.1016/j.cell.2013.09.054
R Core Team. 2020. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/.
Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29:** 24–26. doi:10.1038/nbt.1754
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:** 1153–1165. doi:10.1534/genetics.113.159855
Rödelsperger C, Meyer JM, Prabh N, Lanz C, Bemm F, Sommer RJ. 2017. Single-molecule sequencing reveals the chromosome-scale genomic architecture of the nematode model organism *Pristionchus pacificus*. *Cell Rep* **21:** 834–844. doi:10.1016/j.celrep.2017.09.077
Rödelsperger C, Röseler W, Prabh N, Yoshida K, Weiler C, Herrmann M, Sommer RJ. 2018. Phylotranscriptomics of *Pristionchus* nematodes reveals parallel gene loss in six hermaphroditic lineages. *Curr Biol* **28:** 3123–3127.e5. doi:10.1016/j.cub.2018.07.041
Rödelsperger C, Athanasouli M, Lenuzzi M, Theska T, Sun S, Dardiry M, Wighard S, Hu W, Sharma DR, Han Z. 2019a. Crowdsourcing and the feasibility of manual gene annotation: a pilot study in the nematode *Pristionchus pacificus*. *Sci Rep* **9:** 18789. doi:10.1038/s41598-019-55359-5
Rödelsperger C, Prabh N, Sommer RJ. 2019b. New gene origin and deep taxon phylogenomics: opportunities and challenges. *Trends Genet* **35:** 914–922. doi:10.1016/j.tig.2019.08.007
Rödelsperger C, Ebbing A, Sharma DR, Okumura M, Sommer RJ, Korswagen HC. 2020. Spatial transcriptomics of nematodes identifies sperm cells as a source of genomic novelty and rapid evolution. *Mol Biol Evol* **38:** 229–243. doi:10.1093/molbev/msaa207
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:** 2835–2843.e4. doi:10.1016/j.celrep.2018.05.008
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:** e1008687. doi:10.1371/journal.pgen.1008687
Sigg CD, Buhmann JM. 2008. Expectation-maximization for sparse and non-negative PCA. In *Proceedings of the 25th International Conference on Machine Learning*, pp. 960–967. Association for Computing Memory, New York.
Sommer RJ. 2020. Phenotypic plasticity: from theory and genetics to current and future challenges. *Genetics* **215:** 1–13. doi:10.1534/genetics.120.303163
Sommer RJ, Carta LK, Kim SY, Sternberg PW. 1996. Morphological, genetic and molecular description of *Pristionchus pacificus* sp. n. (Nematoda: Neodiplogastridae). *Fund Applied Nemat* **19:** 511–521.
Susoy V, Ragsdale EJ, Kanzaki N, Sommer RJ. 2015. Rapid diversification associated with a macroevolutionary pulse of developmental plasticity. *eLife* **4:** e05463. doi:10.7554/eLife.05463
Susoy V, Herrmann M, Kanzaki N, Kruger M, Nguyen CN, Rödelsperger C, Röseler W, Weiler C, Giblin-Davis RM, Ragsdale EJ, et al. 2016. Large-scale diversification without genetic isolation in nematode symbionts of figs. *Sci Adv* **2:** e1501031. doi:10.1126/sciadv.1501031
Wang X, Sommer RJ. 2011. Antagonism of LIN-17/frizzled and LIN-18/RyK in nematode vulva induction reveals evolutionary alterations in core developmental pathways. *PLoS Biol* **9:** e1001110. doi:10.1371/journal.pbio.1001110
Werner MS, Sieriebriennikov B, Loschko T, Namdeo S, Lenuzzi M, Dardiry M, Renahan T, Sharma DR, Sommer RJ. 2017. Environmental influence on *Pristionchus pacificus* mouth form through different culture methods. *Sci Rep* **7:** 7207. doi:10.1038/s41598-017-07455-7
Werner MS, Sieriebriennikov B, Prabh N, Loschko T, Lanz C, Sommer RJ. 2018. Young genes have distinct gene structure, epigenetic profiles, and transcriptional regulation. *Genome Res* **28:** 1675–1687. doi:10.1101/gr.234872.118
Witte H, Moreno E, Rödelsperger C, Kim J, Kim J-S, Streit A, Sommer RJ. 2015. Gene inactivation using the CRISPR/Cas9 system in the nematode *Pristionchus pacificus*. *Dev Genes Evol* **225:** 55–62. doi:10.1007/s00427-014-0486-8
Wood W. 1988. Introduction to *C. elegans* biology. In *The nematode Caenorhabditis elegans* (ed. Wood WB). Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Wu G, Anafi RC, Hughes ME, Kornacker K, Hogenesch JB. 2016. MetaCycle: an integrated R package to evaluate periodicity in large scale data. *Bioinformatics* **32:** 3351–3353. doi:10.1093/bioinformatics/btw405
Zugasti O, Rajan J, Kuwabara PE. 2005. The function and expansion of the patched- and hedgehog-related homologs in *C. elegans*. *Genome Res* **15:** 1402–1410. doi:10.1101/gr.3935405

Received January 25, 2021; accepted in revised form July 14, 2021.