

UC Davis

UC Davis Previously Published Works

Title

Genomic and Transcriptomic Insights into the Evolution and Parasitic Strategy of the Woody-Plant Nematode *Pratylenchus vulnus*

Permalink

<https://escholarship.org/uc/item/29m7s9q1>

Journal

Molecular Plant-Microbe Interactions, 39(2)

ISSN

0894-0282

Authors

Dai, Dadong
Zhang, Yali
Latina, Romnick
[et al.](#)

Publication Date

2026

DOI

10.1094/mpmi-10-25-0133-fi

Peer reviewed



Published in final edited form as:

Mol Plant Microbe Interact. 2026 March ; 39(2): 265–277. doi:10.1094/MPMI-10-25-0133-FI.

Genomic and Transcriptomic Insights into the Evolution and Parasitic Strategy of the Woody-Plant Nematode *Pratylenchus vulnus*

Dadong Dai^{1,†}, Yali Zhang^{1,†}, Romnick Latina^{1,2}, Xuyun Wang¹, Valerie M. Williamson³, Simon C. Groen^{4,5,6}, Syed Shamsullah^{1,7}, Charles A. Leslie⁸, Bardo Castro¹, Shahid Siddique^{1,*}

¹Department of Entomology and Nematology, University of California, Davis, CA, 95616

²Institute of Weed Science, Entomology and Plant Pathology, University of the Philippines Los Baños

³Department of Plant Pathology, University of California, Davis, CA, 95616, USA

⁴Department of Nematology, University of California, Riverside, CA 92521, USA

⁵Department of Botany and Plant Sciences, University of California, Riverside, CA 92521, USA

⁶Center for Plant Cell Biology, Institute for Integrative Genome Biology, University of California Riverside, Riverside, CA 92521, USA

⁷Department of Plant Pathology, University of Agriculture, Faisalabad, 38040, Pakistan

⁸Department of Plant Sciences, University of California, Davis, CA 95616, USA

Abstract

The root-lesion nematode *Pratylenchus vulnus* parasitizes a wide range of hosts including woody perennials such as walnut (*Juglans regia*) and grapevine (*Vitis vinifera*), significantly damaging roots and reducing yields. Here, we present a high-quality, chromosome-level genome assembly of *P. vulnus* (61.7 Mb across six chromosomes). Comparative genomic analysis revealed high collinearity in protein-coding genes between *P. vulnus* and the root-knot nematode included in this analysis, *Meloidogyne graminicola*, indicating a close evolutionary relationship between these genera of plant parasites. Large chromosomal regions in *P. vulnus* lack synteny with other nematode genomes, have comparatively low GC content (<30%), and are enriched in genes with unique or lineage-specific functions. Transcriptome analysis highlighted dynamic, stage-specific expression of genes involved in parasitism, development, and metabolism. Additionally, we identified an extensive repertoire of putative effector genes and characterized lineage-specific

*Corresponding author: Professor Shahid Siddique (ssiddique@ucdavis.edu).

†These authors contributed equally to this manuscript.

Author contributions

S.S. conceived and administered the project and acquired the funding; D.D. performed the data analysis; D.D. and Y.Z. generated the figures; D.D., Y.Z., and R.L. collected the stage-specific nematodes; D.D. constructed the stage-specific RNA-seq libraries; X.Y. and B.C. raised the nematodes; D.D. wrote the draft manuscript; S. C. G and S.S. revised the manuscript. All authors contributed to the manuscript.

Competing interests

The authors declare no competing interests.

expansions of cell wall-degrading enzyme families. Overall, these findings provide insight into the genome organization, chromosome evolution, and parasitism-related gene repertoire in a woody plant-parasitizing nematode.

Keywords

Pratylenchus vulnus; Root-lesion nematode; Plant–nematode interaction; Genomics; Effector proteins; Lineage-specific genes; Woody host parasitism

Introduction

Plant-parasitic nematodes (PPNs) are among the most destructive plant pests, causing billions of dollars in crop losses globally each year (Jones et al., 2013). Among PPNs, root-lesion nematodes (RLNs) of the genus *Pratylenchus* are particularly damaging due to their broad host range, affecting both herbaceous and woody plant species (Fosu-Nyarko and Jones, 2016, Jones et al., 2013). The infection cycle of RLNs begins when second-stage juveniles (J2s) hatch from eggs deposited in the soil or in root tissues. Unlike sedentary endoparasitic nematodes, all life stages of RLNs—including third- and fourth-stage juveniles (J3s and J4s), as well as adult males and females—remain motile and can enter and leave host root tissues guided by root exudates and other environmental cues (Friebe et al., 1998, Fosu-Nyarko and Jones, 2016). Once inside, they move inter- and intracellularly through the root cortex, feeding on live cells with a stylet and causing characteristic necrotic lesions. The entire life cycle can be completed within a few weeks under favorable environmental conditions, allowing for multiple generations per growing season and contributing to their persistence and damage, especially in perennial cropping systems (Chihani-Hammas et al., 2018, Karaca et al., 2020, Fosu-Nyarko and Jones, 2016).

Among RLNs, *Pratylenchus vulnus* poses a significant threat to perennial crops such as walnut (*Juglans regia*), grapevine (*Vitis vinifera*), and various stone fruits (Karaca et al., 2020, Chihani-Hammas et al., 2018). In California, where walnuts are a major agricultural commodity, *P. vulnus* poses an existential threat to the industry (Pollack, 1998, McKenry, 1989). Although comprehensive statewide surveys are lacking, yield losses attributed to RLNs in walnut orchards are estimated to range from 10% to 20%, potentially translating to \$140 million in annual losses (Personal Communication, California Walnut Board and Commission). In severe cases, yield reductions can reach 50% or more. Traditionally, management in walnut orchards has relied on soil fumigants such as 1,3-Dichloropropene. However, growing regulatory restrictions and environmental concerns have made these chemical control methods less viable, highlighting the urgent need for alternative, sustainable strategies.

Despite the agricultural importance of *P. vulnus*, genomic resources for this species—and RLNs more generally—remain extremely limited (Kikuchi et al., 2017). By comparison, substantial progress has been made in the genomic analysis of sedentary endoparasites such as root-knot nematodes (RKNs, *Meloidogyne* spp.) (Dai et al., 2023) and cyst nematodes (CNs, *Heterodera* and *Globodera* spp.) (Siddique et al., 2022), which primarily parasitize annual herbaceous crops. Interestingly, molecular phylogenetic analyses have shown that

although both RKNs and CNs are sedentary PPNs, RKNs are more closely related to RLNs, whereas their evolutionary relationship with CNs is more distant (Qing et al., 2024). Genomic studies of RKNs have revealed complex ploidy differences, evidence of horizontal gene transfer (HGT), and effector diversification underpinning their parasitic strategies (Dai et al., 2023). However, the lack of comparable data for RLNs restricts our understanding of parasitism and the evolution of host adaptation in this group (Kikuchi et al., 2017).

Here, we present a chromosome-level genome assembly of *Pratylenchus vulnus*, providing a foundational genomic resource for RLNs. We integrated genome-wide analyses of gene content, synteny with other nematode species, and chromosome structure with comparative transcriptomic profiling to investigate host-induced gene expression dynamics and effector deployment.

Results

Chromosome-level assembly and annotation of the *P. vulnus* genome

To better understand the genomic features associated with the ability of migratory nematodes to parasitize woody plants with lignified xylem, we sequenced and assembled the genome of *P. vulnus* strain *Yolo1*. First, we generated 15 Gb of short-read Illumina sequencing data and analyzed it using Smudgeplot, confirming that *P. vulnus* is diploid (Fig. S1A). Using GenomeScope2, we estimated a genome size of approximately 57 Mb (Fig. S1B). For genome assembly, we utilized 20 Gb of long-read Nanopore data and 8 Gb of PacBio HiFi data, yielding a 61.7-Mb draft genome assembled into 63 contigs (Table 1). Subsequently, three rounds of polishing with Illumina short reads improved sequence accuracy. To assess potential contamination from bacterial or other non-target DNA sequences in our assembly, we performed BlobPlot analysis (Challis et al., 2020) of the assembled contigs. The results showed that 89.5% of the total length was classified as coming from Nematoda, no hits were found for 6.9% (no-hit), and only three contigs, accounting for 3.6% of the total sequence length, were initially identified as having non-nematode origins (Fig. S2). We further examined the GC content of contigs with no-hit and non-nematode origin. Their GC content ranges (34.3–39.2% and 27.4–43.9%, respectively) largely overlapped with that of contigs classified as having nematode origins (25.7–54.7%), indicating no obvious difference in GC content distribution among these groups. Upon closer manual inspection of the BLAST alignment details for contigs identified as non-nematode in origin, we found that only short, discontinuous regions of a few hundred base pairs aligned to potential contaminant taxa. Therefore, we consider these contigs to be more appropriately categorized as “no-hit”.

To achieve a chromosome-level assembly, we performed chromatin conformation capture (Hi-C) sequencing, generating 78 Gb of data. Using chromatin interaction information, we anchored 88.89% of the contigs to six pseudochromosomes (Fig. 1A), consistent with previous karyotype studies on this species (Roman and Triantaphyllou, 1969). To assess the completeness of our assembled genome, we searched for tandem repeat sequences and identified from 67 to over 4,000 copies of the typical telomeric repeat sequence TTAGGC at the ends of each chromosome (Fig. 1B). G-quadruplex motifs were unevenly distributed across the chromosomes (Fig. 1B), showing enrichment in some regions and scarcity in

others, suggesting potential structural or regulatory roles in specific chromosomal domains (Du et al., 2009).

To support our gene predictions and functional annotation, we collected mixed-stage nematodes parasitizing the roots of walnut trees or grapevine and extracted total RNA for library construction and sequencing. We identified 13,281 protein-coding genes and 18,704 transcripts, including 5,423 alternative isoforms. Functional annotation using the Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Pfam databases assigned putative functions to 14,432 of the transcripts (77.2%). Additionally, approximately 9.6% of the genome consisted of repetitive sequences, comprising 39,078 transposable elements (TEs; Supplementary Table S1). TEs were grouped into major categories including DNA transposons (20,171 copies), LINEs (273 copies), LTR retrotransposons (205 copies), and MITEs (7,711 copies). Annotating non-coding RNA-transcribing genomic loci revealed 371 tRNAs, 88 rRNAs, and 66 snRNAs. We also identified 712 transcription factor-encoding genes from 92 families, 1,629 genes encoding putatively secreted proteins (candidate effectors), and 326 horizontally transferred genes.

We assessed genome completeness using Benchmarking Universal Single-Copy Orthologs (BUSCO) with the nematoda_odb10 dataset, yielding completeness scores of 63.3% for the genome assembly and 68.2% for the predicted protein set (Fig. 1C), comparable to other published PPN genomes (Dai et al., 2023). Since BUSCO analysis treats each transcript isoform as an independent protein sequence, alternative transcripts originating from the same locus are counted as duplicated orthologs, thereby inflating the duplication score. Finally, we used Circos plots (Krzywinski et al., 2009) to visualize the genomic landscape, showing gene density, gene expression levels, GC content, and TE density (Fig. 1D). Notably, the left arm of chromosome 1, the right arm of chromosome 2, and specific regions on chromosomes 4 and 5 exhibited lower gene density, lower TE and GC contents, and lower gene expression levels relative to the genome-wide averages (Fig. 1D).

Developmental, host-related, and cold-induced gene expression signatures in *P. vulnus*

To explore transcriptomic variation across the nematode life cycle, we manually collected nematodes at six distinct developmental stages—eggs, J2s, J3s, J4s, males, and females—and performed deep RNA sequencing (RNA-seq) of each stage (Fig. 2A). In addition, to compare gene expression of nematodes from different hosts, we analyzed mixed-stage transcriptomes of nematodes isolated from walnut and grape. Furthermore, given that *P. vulnus* is known to survive extended periods of low temperatures typical of winter soil conditions that commonly occur in Mediterranean-type climates like California's Central Valley, we sought to investigate its early molecular response to cold exposure. To do this, we exposed a mixed-stage population to 4 °C for 6 hours to capture initial transcriptional changes triggered by acute cold treatment (Fig. 2A). While this short-term exposure provides insight into immediate gene expression shifts, it may not fully reflect the physiological and molecular mechanisms that enable *P. vulnus* to persist under prolonged cold conditions (Towson and Lear, 1982, Acosta and Malek, 1979) (Fig. 2A).

Principal component analysis (PCA) of the RNA-seq data confirmed high reproducibility among biological replicates and revealed transcriptional patterns associated with both

developmental stages and the environmental factors of host plant and temperature (Fig. 2B, 2C). For nematodes across different developmental stages, samples were clearly separated along principal component 1 (PC1), which explains 32.19% of the variance and broadly recapitulates progression through the nematode life cycle. PC2 (21.77%) distinguished egg samples from those of nematodes at post-hatching stages, likely reflecting gene expression patterns specific to early embryogenesis (Fig. 2B). PCA of RNA-seq data from nematodes experiencing different environmental conditions showed that PC1 (explaining 35.44% of the variance) strongly separated samples of cold-treated nematodes from ones of nematodes collected from grapevine and walnut hosts (Fig. 2C). Although this separation indicates a robust transcriptional response to cold exposure, some differences may reflect the fact that the cold-treated nematodes had originally been reared on grapevines. Indeed, PC2 clearly differentiated walnut-associated nematodes from grapevine-associated ones with and without cold exposure (Fig. 2C), indicating that the host environment also significantly influences nematode gene expression.

To identify genes that significantly respond to developmental transitions or environmental conditions, we conducted a differential gene expression analysis. The largest transcriptional shifts during development—reflected by the greatest number of differentially expressed genes (DEGs; $P_{adj} < 0.05$; $\log_2$ fold-change > 1)—occurred at two key transitions: from egg to J2, and from J4 to adult (Fig. 3A). Notably, the egg-to-J2 transition involved substantial upregulation of genes encoding pectate lyases—enzymes involved in plant cell wall degradation—highlighting the activation of parasitism-related processes for host invasion and migration within the host (Fig. 3B). Additionally, genes encoding G protein-coupled receptors (GPCRs) were significantly enriched in J2s compared to eggs (Fig. 3B and Fig. S3). GPCRs constitute a large family of membrane receptors that mediate nematode responses to diverse external and internal signals, including small molecules and peptides (Cardoso et al., 2012, Beets et al., 2023, Holden-Dye and Walker, 2011). The pronounced stage-specific regulation of GPCR genes thus highlights their likely involvement in coordinating developmental processes and environmental responses during the life cycle of *P. vulnus*.

Interestingly, the highest number of DEGs overall was observed after acute cold treatment with 2,306 upregulated and 1,766 downregulated genes (Fig. 3A). Cold-responsive DEGs were primarily enriched for genes encoding transcription factors, ribosomal proteins, and regulators of cellular osmotic pressure (Fig. S4). When we compared the transcriptomes of *P. vulnus* parasitizing grapevine roots versus walnut roots, we detected 2,793 DEGs. Functional enrichment analysis of genes more highly expressed during walnut parasitism revealed significant enrichment in functions related to transcriptional regulation, DNA binding, morphogenesis, and tissue development, particularly epithelial tube (intestine) formation and organ morphogenesis (Fig. S5). Walnuts have a much tougher root than grapes. Therefore, this extensive transcriptional remodeling may facilitate development of stronger musculature, reinforced feeding structures, or enhanced absorptive tissues to exploit the physically demanding walnut root environment.

To characterize stage-specific gene expression dynamics in greater detail, we calculated transcripts per million (TPM) values for each gene and performed soft clustering analysis

across all developmental stages. Each developmental stage exhibited a unique set (cluster) of highly expressed genes (Figs. S6 and S7). Functional enrichment analyses of these clusters revealed stage-specific biological functions. For instance, clusters with male-biased expression were enriched for terms related to male gamete generation, reproduction, and kinase activity (Fig. 3C). Gene clusters with female-biased expression were enriched for reproduction-related functions, lipid metabolism, and small-molecule metabolism (Fig. 3D). Interestingly, genes that were highly expressed in females also exhibited enrichment for the GO terms male gamete generation and spermatogenesis, suggesting that some sex-associated genes have biological functions beyond sex determination alone.

Finally, clustering analysis of RNA-seq data from nematodes that had or had not been exposed to acute cold identified a subset of genes that are specifically and strongly expressed upon cold treatment (cold-induced cluster; Fig. S8A). Functional enrichment analysis of these cold-responsive genes revealed an enrichment of transcription factors, particularly members of the basic leucine zipper (bZIP) family (Fig. S8B), a protein family previously associated with responses to cold in plants (Han et al., 2020, Liu et al., 2019, Liu et al., 2023, Kreps et al., 2002). Genes encoding Na⁺/H⁺ exchangers—known regulators of osmotic balance—were also significantly enriched among cold-responsive genes in *Drosophila* (Enriquez and Colinet, 2019). Although these results highlight the nimble transcriptional responses of *P. vulnus* under short-term cold exposure, additional long-term studies are needed to capture the full scope of its cold-acclimation strategies. Collectively, these transcriptomic data highlight the dynamic and context-dependent nature of gene expression throughout *P. vulnus* development and parasitism, providing valuable insights into its development and responses to specific hosts and environmental conditions.

Expression landscape and genomic distribution of effector and HGT-derived genes

PPNs use effector proteins to facilitate infection and suppress defenses in host plants (Kud et al., 2019). To better understand the genomic features associated with adaptation to a parasitic lifestyle, we examined the distribution and expression patterns of putatively secreted proteins as most characterized effectors of PPNS are secreted (Rehman et al., 2016). We identified 1,629 candidate effector genes, which were relatively evenly distributed across all chromosomes (Fig. 4A). To characterize the developmental regulation of their expression, we calculated the TPM values of each gene across nematode development and performed clustering analysis. Although *P. vulnus* is a migratory PPN capable of infecting plants continuously from the J2 stage onward, the candidate effector genes exhibited distinct and surprisingly stage-specific expression profiles (Fig. 4B and Fig. S9). Such stage-specific expression patterns likely reflect diverse functional demands at each developmental stage, such as root penetration, immune evasion, and nutrient acquisition. In addition, all nematodes, both free living and parasitic, produce a repertoire of secreted proteins that function in their development, neurotransmission, and other functions. Thus, while secreted proteins are likely enriched in effectors, some of the candidate effectors may, in fact, have other functions. One clear example is a cluster dominated by genes expressed almost exclusively in eggs, which are most likely secretory proteins with roles in development rather than parasitism.

PPNs carry an array of genes involved in carbohydrate metabolism that show signatures of having been acquired by HGT, relatives of which are known to have important functions in parasitism in other PPNs (Mitreva et al., 2009). In contrast to the dispersed genomic arrangement of genes predicted to encode secreted proteins, HGT-derived genes exhibited distinctly clustered genomic distributions (Fig. 4A). This clustered arrangement aligns with previous observations in RKNs (Dai et al., 2023), suggesting that these genes may have expanded in copy number following acquisition. Similar to the genes encoding putatively secreted proteins, HGT-derived genes displayed clear stage-specific expression patterns, with distinct subsets highly expressed at particular developmental stages (Fig. 4C).

Functional enrichment analysis of HGT-derived genes identified a significant overrepresentation of ones encoding glycosyl hydrolases and pectate lyases (Fig. S10); such enzymes are known to be important effectors in PPNs, essential for plant cell wall degradation and sugar metabolism during infection (Rai et al., 2015, Dayi, 2024). All identified pectate lyase genes in *P. vulnus* appear to be derived from HGT events. Together, these findings reveal distinct genomic distributions and precisely coordinated expression patterns of candidate secreted protein and HGT-derived genes, offering valuable insights into how *P. vulnus* may have adapted to a parasitic lifestyle.

Genome comparison with other nematode species reveals extensive chromosomal rearrangements and unique genomic regions in *P. vulnus*

To better understand genomic changes that occurred as plant parasitism evolved in the lineage leading up to *P. vulnus*, we conducted a comparative genomic analysis that included several closely related PPNs with sequenced genomes. A recent BUSCO50-based phylogeny identified the migratory PPN genus *Pratylenchus* as most closely related to the sedentary PPN genus *Meloidogyne* (Qing et al., 2024). Thus, we first evaluated the chromosomal collinearity between *P. vulnus* and *M. graminicola*, a well-characterized diploid RKN with a similar genome size, but 18 chromosomes (Fig. 5A). This comparison revealed blocks of synteny, but also several examples where individual chromosomes in *P. vulnus* exhibited collinearity to multiple chromosomes in *M. graminicola*. Notably, chromosome 3 of *P. vulnus* contained collinear segments corresponding to eight separate chromosomes in *M. graminicola* (Fig. 5A). Conversely, several chromosomes in *M. graminicola* each had segments represented on only single chromosomes in *P. vulnus* (Fig. 5A).

Intriguingly, the left arms of chromosomes 1 and 4 as well as the right arm of chromosome 2 in *P. vulnus* lacked corresponding sequences in the *M. graminicola* genome (Fig. 5A), suggesting that these genomic regions were acquired in *P. vulnus* after splitting from a common ancestor with *M. graminicola* or were lost in *M. graminicola*. Interestingly, these regions correspond roughly to those noted to have low GC content and low gene density (Fig. 1D). Further comparative analyses with other migratory but more distantly related PPNs—*Ditylenchus destructor*, *Aphelenchoides besseyi*, and *Bursaphelenchus xylophilus*—as well as free-living nematodes relatively closely related to the latter two species—*Panagrellus* sp. LJ2284 (Fig. S11) and *Panagrolaimus* sp. JU765—similarly failed to identify homologous genomic segments, highlighting the uniqueness of these regions (Fig. S12A–D).

We extended our comparative analysis to include the published genome of *P. scribneri* (Arora et al., 2023), a congener of *P. vulnus*, to further investigate the origin of the non-collinear chromosomal regions. Although the *P. scribneri* genome is somewhat fragmented, we identified partial homology between its sequences and the left arms of chromosomes 1 and 4 as well as the right arm of chromosome 2 in *P. vulnus*—regions that otherwise showed poor synteny with the genomes of *M. graminicola* and other nematodes (Fig. S13). This suggests that these segments may be partially conserved within the genus *Pratylenchus* and may trace back to an ancestral lineage common to this genus.

Nigon elements, which are conserved genomic blocks present across Rhabditida nematodes, serve as useful markers for analyzing chromosome evolution due to their characteristic distribution patterns among nematode chromosomes (Gonzalez de la Rosa et al., 2021). Analysis of Nigon element distribution in *Panagrolaimus* sp. JU765 and *Panagrellus* sp. LJ2284 showed that each autosome primarily contained a single Nigon element, while their X chromosomes originated from ancestral chromosome block fusions involving NigonX, NigonC, and NigonD elements (Fig. 5B and Fig. S14). Similarly, the six chromosomes of *B. xylophilus* predominantly contained single Nigon elements, except for its X chromosome, which arose through the fusion of ancestral NigonX and NigonC chromosomal segments (Gonzalez de la Rosa et al., 2021). Furthermore, collinearity analysis demonstrated that the X chromosome of *Panagrolaimus* sp. JU765 shared collinear segments with chromosomes 3 (NigonX, NigonC) and 4 (NigonD) of *B. xylophilus* (Fig. 5C), reinforcing the utility of Nigon elements for inferring the evolutionary histories of chromosomes. In comparison, the distribution of Nigon elements in *P. vulnus* chromosomes showed a notably fragmented and complex pattern (Fig. 5D), indicating extensive rearrangements, fusions, or fragmentation events compared to ancestral chromosome structures. Integrating Nigon element (NigonX) distribution data with interspecies protein-based synteny analyses (Fig. 5D and Fig. S12), we propose that chromosome 2 is likely the X chromosome of *P. vulnus*.

Remarkably, the left arms of chromosomes 1 and 4, as well as the right arm of chromosome 2 in *P. vulnus* were nearly devoid of Nigon elements. Given the lack of protein-coding gene collinearity between these regions and other examined nematode genomes (Fig. 5A and Fig. S12), we speculate that these segments may have originated from nematodes lacking typical Nigon elements or even other phyla. Collectively, our results demonstrate that the genome of *P. vulnus* has undergone significant chromosomal rearrangements and genomic innovations compared to its closest relatives. These historical events may have shaped its ecological specialization on woody plants by incorporating genomic elements from diverse ancestral sources, underscoring a complex evolutionary trajectory that is distinct from that of other, related nematodes.

Comparative analysis reveals lineage-specific diversification of cell wall–degrading enzyme genes in *P. vulnus* and related nematodes

P. vulnus is one of the most damaging PPNs of woody host plants such as walnut and peach (*Prunus persica*), which have highly lignified and structurally complex cell walls. This specialized ecological niche raises the question of whether *P. vulnus* possesses a distinct repertoire of cell wall–degrading enzymes (CWDEs) that enable it to colonize

such robust tissues. To address this, we performed a comparative analysis of CWDE gene families across eight nematode species spanning a spectrum of lifestyles, from free-living to highly specialized plant parasites (Qing et al., 2024). These included the free-living species *Oscheius tipulae*, *Panagrolaimus* sp. JU765, and *Panagrellus* sp. LJ2284; the migratory parasites *B. xylophilus*, *A. besseyi*, and *D. destructor*; and *P. vulnus*; and the sedentary endoparasite *M. graminicola*. We focused on CWDE families including glycoside hydrolases (GH), pectate lyases (PL) and carbohydrate-binding modules (CBM13) (Rai et al., 2015, Dayi, 2024, Gahoi and Gautam, 2017). Among the analyzed nematodes, *B. xylophilus* had the largest CWDE gene complement (152 genes), followed by the free-living species *Panagrolaimus* sp. JU765 (101 genes) and *Panagrellus* sp. LJ2284 (78 genes) (Fig. 6A and Table S2). By contrast, the sedentary parasite *M. graminicola* possessed the smallest CWDE repertoire, comprising only 28 genes. *P. vulnus* exhibited an intermediate-sized CWDE gene set (66 genes), placing it between free-living nematodes and sedentary plant parasites (Fig. 6A and Table S2). Interestingly, although both *P. vulnus* and *B. xylophilus* parasitize woody hosts, their CWDE gene compositions differed markedly (Table S2).

To explore this further, we conducted phylogenetic analyses of key CWDE families enriched in *P. vulnus*, including GH5, GH27, GH38, GH43, GH47, CBM13, and PL. These analyses indicated that several *P. vulnus* CWDE genes formed clades that were phylogenetically distinct from those in other nematodes, suggesting lineage-specific expansions or independent origins (Fig. 6B and Fig. S15, 16). For instance, the CBM13 family, associated with xylan degradation (Chen et al., 2020), is abundant across nematodes, particularly in *B. xylophilus* (48 genes) and free-living species such as *Panagrellus* sp. LJ2284 (18 genes). Phylogenetic analysis grouped CBM13 proteins into seven distinct clades (Fig. 6B). While some CBM13 genes in *P. vulnus* grouped closely with homologs from other nematodes, three CBM13 genes formed a distinct lineage separate from that of other nematodes, indicating a possible independent evolutionary origin or diversification. Similar patterns were observed for other CWDE families, including GH5 (7 genes), GH38 (4 genes), GH43 (3 genes), and PL (5 genes), which were more abundant in *P. vulnus* relative to *M. graminicola* (Table S2 and Fig. S15, 16).

Further analysis revealed that many of the phylogenetically distinct CWDE genes in *P. vulnus* are located in non-syntenic regions on chromosomes 1, 2, 4, and 5 (Fig. S17), suggesting possible origins through lineage-specific innovation or horizontal gene transfer. To investigate their functional relevance, we examined developmental stage-specific expression patterns for various CWDE gene families in *P. vulnus*. While no consistent pattern emerged at the level of gene families, several *P. vulnus*-specific CWDE genes showed elevated expression during parasitic stages (J2–J4) and in adult females (Fig. 6C). This includes the CBM13 gene *P. vulnus_03856*, which exhibited the highest expression among all CWDEs in females (Fig. 6D). Although the biases in expression patterns across developmental stages were gene-specific rather than family-wide, the transcriptional activity of several phylogenetically distinct genes during parasitic stages underscores their potential relevance to host invasion and colonization. Together, these findings suggest that *P. vulnus* has evolved a unique set of CWDEs through lineage-specific diversification and expression, likely enabling adaptation to the physical and chemical defenses of woody plant tissues.

Discussion

In this study, we present a chromosome-scale genome assembly for *P. vulnus*, consisting of 6 chromosomes. This corresponds to the previously reported chromosome number for this strain. One of the most striking findings was the presence of extensive genomic regions that lack synteny with other sequenced nematode genomes, including well-characterized free-living and parasitic species. These non-collinear regions were structurally distinct and compositionally divergent, exhibiting GC content as low as 25–30%, in contrast to the 32–55% GC content observed in collinear, conserved chromosomal segments. The presence of these low-GC, non-syntenic regions suggests a possible origin from evolutionarily distant nematode lineages, potentially from other clades or orders for which genomic resources are still unavailable. This hypothesis is supported by the distribution of parasitism-related genes in these regions. Several genes encoding cell wall-degrading enzymes (CWDEs), such as those from the CBM13, GH5, GH27, GH38, GH43, and PL families, are in these non-syntenic blocks and show distant phylogenetic relationships to homologs in other nematodes. Likewise, a subset of predicted effector genes—key to host manipulation—are found within these regions, several of which exhibit stage-specific expression across juvenile and adult phases. Alternatively, these regions may represent ancient HGT events or remnants of genome fusion, possibly derived from now-extinct or unsampled species. Their divergence in base composition and lack of conserved gene order support the hypothesis that they evolved under distinct selective pressures or genomic contexts, separate from the core genome.

In addition to shedding light on genomic novelty within *P. vulnus*, our findings have broader implications for the evolution of PPN karyotypes. Notably, the observed absence of collinearity in certain chromosomal regions parallels chromosomal plasticity observed in other parthenogenetic nematodes, particularly RKNs. These nematodes also display extraordinary variation in chromosome number (Roman and Triantaphyllou, 1969), extensive structural rearrangements, and evidence of ancient hybridization and polyploidy. The fact that several *Pratylenchus* species reproduce by mitotic parthenogenesis, with chromosome numbers ranging from 5 to 32, suggests that *Pratylenchus* may represent a pivotal lineage in understanding the evolutionary origin of the highly derived genomes of *Meloidogyne*. Intriguingly, the lowest chromosome number reported in *Meloidogyne* is 4 (Alvarez-Ortega et al., 2019), similar to that observed in some *Pratylenchus* species. This small chromosome count raises the possibility that extensive chromosomal rearrangements and reductions may have accompanied or facilitated transitions in reproductive mode and host adaptation. The *Meloidogyne* genomes sequenced so far lack telomeres, whereas *P. vulnus* has long runs of canonical telomere repeats.

To test these evolutionary scenarios more robustly, additional high-quality genome assemblies from other *Pratylenchus* and *Meloidogyne* species are needed. In particular, sequencing *Pratylenchus* species with extreme karyotypes (Roman and Triantaphyllou, 1969) or unique ecological niches could help reconstruct ancestral genomic states and delineate the steps underlying chromosome reshaping, gene acquisition, and reproductive transitions in PPNs.

In conclusion, our study highlights the evolutionary complexity and genomic plasticity of *P. vulnus*, a key representative of woody plant–parasitizing nematodes. The discovery of non-collinear, low-GC chromosomal regions points to deep evolutionary events that shaped this genome, potentially involving ancient lineage divergence or acquisition from uncharacterized taxa. Moreover, our findings underscore the importance of expanded taxon sampling and chromosome-resolved assemblies in uncovering the origins of genomic novelty and adaptation in parasitic nematodes.

Methods

Nematode culture and collection

Pratylenchus vulnus nematodes were propagated on sterile grape and walnut roots grown on solidified Driver Kuniyuki Walnut (DKW) basal medium (i.e. no growth regulators) under axenic conditions. After 3 weeks of root growth, sterile nematodes were introduced, and the cultures were maintained at 25°C for 6–8 weeks. For nematode collection, plant roots were finely chopped with a scalpel along with the surrounding medium and transferred to Baermann funnels lined with a single layer of Kimwipes. Water was added until the plant material was fully submerged, and nematodes were allowed to migrate for approximately 4 h. The eluate was then collected from the bottom of the funnel and mixed with 70% sucrose to achieve a final concentration of 35%. The mixture was transferred to a 1.5 mL microcentrifuge tube and centrifuged at 5,000 rpm for 2 min. The pellet, consisting of impurities, was discarded. The supernatant containing the nematodes was then passed through a 3 µm filter membrane to remove the sucrose, yielding purified nematodes. The purified nematodes were used in all experiments. Genomic DNA used for sequencing was extracted from *P. vulnus* infecting grape roots. Transcriptome data from different developmental stages and cold treatment conditions were obtained from nematodes infecting grapevine roots; however, the mixed-stage transcriptome samples were collected from nematodes infecting grape or walnut roots. The cold treatment samples the control (non-cold treated) samples were prepared using the same collection procedure and thus had comparable stage composition. Therefore, observed transcriptomic differences reflect genuine cold-response biology rather than variation in developmental-stage proportions.

Acquisition of nematodes at different stages of development

To collect nematodes at different stages of development, individuals were manually picked under a stereomicroscope. Eggs were the most easily distinguishable stage, and approximately 150 eggs were collected for each biological replicate. Newly hatched juveniles from eggs were designated as second-stage juveniles (J2s), with ~150 individuals collected per replicate. Third-stage (J3) and fourth-stage (J4) juveniles were selected based on body size differences from a mixed-stage nematode population, with ~100 individuals per replicate. Adult females and males were distinguished by their reproductive organs, and ~50 individuals of each sex were collected per replicate.

DNA and RNA extraction and library construction

DNA and RNA were extracted from *P. vulnus* using previously established protocols for RKNs (Dai et al., 2023). Briefly, for DNA extraction, mixed-stage nematodes were

suspended in 200 µL lysis buffer, flash-frozen in liquid nitrogen, and ground into a fine powder using a mortar and pestle. After adding another 300 µL lysis buffer, the lysate was mixed with 500 µL CTAB buffer at 65°C for 30 min, followed by protein removal using phenol:chloroform:isoamyl alcohol extraction. The DNA was precipitated with ethanol. DNA for Illumina sequencing was used directly after extraction, and high-molecular-weight DNA was further purified for Nanopore and PacBio HiFi sequencing using magnetic beads before being submitted to Novogene (Sacramento, CA) for library preparation and sequencing. Hi-C library construction was performed by Phase Genomics using mixed-stage nematodes. Total RNA was extracted from the samples using a RNeasy Mini Kit (QIAGEN, 74106), and RNA-seq libraries were constructed using a VAHTS Universal V10 RNA-seq Library Prep Kit for Illumina (Vazyme, NR606–01). Libraries were sequenced on an Illumina NovaSeq instrument as 150-bp paired-end reads by Novogene.

Assessment of genome size and ploidy

The size, heterozygosity, and ploidy level of the *P. vulnus* genome were estimated based on Illumina short-read sequencing data. *K*-mer-based ploidy analysis was performed using Smudgeplot (Ranallo-Benavidez et al., 2020) to confirm that *P. vulnus* is a diploid organism. Genome size and heterozygosity were estimated using GenomeScope2 (Ranallo-Benavidez et al., 2020) following methods that were previously applied to RKNs (Dai et al., 2023).

Contig assembly

The *P. vulnus* genome was first assembled using long-read Nanopore sequencing data with default parameters in the NECAT assembler (Chen et al., 2021). PacBio HiFi reads were polished using gcpp and subsequently applied to gap filling with PBJelly (English et al., 2012) to improve base accuracy and assembly continuity. Subsequently, three rounds of polishing were performed with Illumina short-read data using Pilon (Walker et al., 2014) to improve sequence accuracy. The error-corrected genome assembly was used for chromosome anchoring.

Chromosome-level genome construction and identification of telomeric elements

The *P. vulnus* contigs were anchored onto chromosomes using Hi-C data, which captured chromatin interaction information. First, the Hi-C reads were aligned to the contig-level assembly using the BWA (v0.7.17-r1188) (Li and Durbin, 2009) and Juicer (v1.5.7) pipelines (Durand et al., 2016b). The 3D-DNA (v180419) pipeline (Dudchenko et al., 2017) was then employed to scaffold adjacent genomic fragments based on chromatin contact signals, generating a draft chromosome-scale assembly. The resulting “.hic” and “.assembly” files were manually curated and corrected using Juicebox (Durand et al., 2016a) to produce the final chromosome-level genome assembly. An in-house Python script was used to identify the tandemly repeated telomeric repeat sequence “TTAGGC”.

For *Panagrellus* sp. LJ2284, the contig assembly was downloaded from NCBI under accession number GCA_024447225.1, along with the corresponding Hi-C dataset (accession number SRR23107840 from the Sequence Read Archive [SRA] at NCBI). A chromosome-level genome was assembled using the strategy described above. During assembly, redundant “unzipped” contigs were identified and removed, likely representing alternative

haplotypes. The final assembly consisted of five chromosomes, with a total genome size of 65.4 Mb. This chromosome-level *Panagrellus* sp. LJ2284 genome assembly was used for genome annotation.

Genome annotation

To annotate protein-coding genes in *P. vulnus*, transcriptome data were integrated from all developmental stages to perform comprehensive gene predictions. First, repetitive sequences in the genome were masked using RepeatMasker (Tarailo-Graovac and Chen, 2009). Subsequently, transcriptome data from eggs, J2s, J3s, J4s, males, females, and mixed samples were sequentially aligned to the genome using HISAT2 (Kim et al., 2019). Low-quality reads (mapping quality [MAPQ] < 30) were removed using SAMtools (Li et al., 2009) to generate the “final.bam” files, which were used for multiple gene predictions. The alignments were used as input for BRAKER3 (Gabriel et al., 2024) to perform *ab initio* gene predictions based on GeneMark-ET/EP and AUGUSTUS, generating the file “braker.gff3”. In parallel, TransDecoder (<https://github.com/TransDecoder/TransDecoder>) and StringTie (Pertea et al., 2015) were used on the same .bam files to predict protein-coding genes, resulting in the file “stringtie.gff3”. Additionally, all RNA-seq datasets were assembled *de novo* using Trinity (Grabherr et al., 2011), and the assembled transcripts (Trinity.fasta) were used as input for PASA (Haas et al., 2008) to identify gene structures, producing the file “pasa_assemblies.gff3”. Finally, the three annotation sources were combined using EVIDENCEModeler (Haas et al., 2008), with weights set to 4 for AUGUSTUS, 4 for TransDecoder, and 5 for PASA. The merged annotation file (EVM.merged.gff3) was processed with the gffrename.py script to standardize gene names and obtain the final gene annotation file. To accurately annotate transcript isoforms and refine gene models, we performed an additional PASA update step. The EVM-derived annotation (EVM.merged.gff3) was supplied to Launch_PASA_pipeline.pl together with the Trinity-assembled transcripts. During this step, PASA realigned full-length transcripts to the genome, updated exon–intron boundaries, added UTRs, and resolved alternative splicing events. This iterative evidence-based refinement allowed PASA to distinguish biologically supported isoforms from incomplete or fragmented transcript structures, ensuring consistent annotation of alternative splicing and preventing artificial inflation of gene numbers. The resulting PASA-updated GFF3 file was used as the final gene and isoform annotation.

For genome annotation of *Panagrellus* sp. LJ2284 and *Panagrolaimus* sp. JU765 (GCA_964245515.1), RNA-seq data were downloaded from the SRA database under accession numbers SRR22163891 and SRR5253561, respectively. The two genomes were annotated using the strategy described above, yielding 16,973 and 16,343 protein-coding genes, respectively. These annotated genomes were used for subsequent analysis. The predicted genes were annotated with GO and KEGG terms using eggNOG-mapper (Cantalapiedra et al., 2021), and Pfam domains were identified using the Pfam_scan.pl script (Mistry et al., 2021). GO enrichment analysis was performed using TBtools-II (Chen et al., 2023), while Pfam enrichment analysis was conducted with the Pfam_Enrichment.py script.

Predicting transposable elements, horizontal gene transfer events, transcription factors, and secreted proteins

The EDTA (v2.2.0) pipeline (Ou et al., 2019) was employed to predict TEs in the *P. vulnus* genome, with the parameters “--overwrite 1 --force 1 --sensitive 1 --anno 1 --evaluate 1”. Horizontal gene transfer events were predicted using the AvP pipeline (Koutsovoulos et al., 2022) with the UniRef90 database. Transcription factors were identified using the AnimalTFDB (v4.0) database (Shen et al., 2023). Secreted proteins were predicted by first identifying signal peptides using SignalP5 (Almagro Armenteros et al., 2019), followed by the detection of transmembrane domains using TMHMM (Krogh et al., 2001). Proteins with signal peptides but lacking transmembrane domains were classified as putative secreted proteins. Horizontal gene transfer (HGT) candidates were identified using the AvP pipeline (Koutsovoulos et al., 2022). AvP combines two sources of evidence for HGT detection: (i) calculation of the Alien Index (AI), which compares the similarity of each protein to non-metazoan versus metazoan sequences to detect potential foreign origin; and (ii) phylogeny-based inference, where gene trees (‘phylomes’) are reconstructed to evaluate whether a gene clusters with non-metazoan lineages rather than within metazoan or nematode clades. Genes supported by both a high AI value and a non-metazoan phylogenetic placement were retained as high-confidence HGT candidates.

Analysis of genome collinearity between different nematodes

Due to the distant evolutionary relationships among the nematode species, inter-genomic synteny analysis was conducted based on annotated protein sequences to minimize the impact of synonymous mutations. Homologous proteins between each pair of genomes were identified using MCScan (Wang et al., 2012). First, the gene list in BED format was generated using the `jcvi.formats.gff` script, and all protein sequences from each genome were extracted using `gffread` (Pertea and Pertea, 2020). Cross-species homologous protein sequences were identified using [jcvi.compara.catalog](https://jvarkit.github.io/jcvi-compara/catalog/), and a synteny analysis of the homologs was performed with [jcvi.compara.synteny](https://jvarkit.github.io/jcvi-compara/synteny/) using the parameter `--minspan=10`. Finally, interspecies chromosomal synteny was visualized using [jcvi.graphics.karyotype](https://jvarkit.github.io/jcvi-graphics/karyotype/).

RNA-seq and gene expression profile analyses

For differential gene expression analysis, the transcriptome data were aligned to the genome assembly generated in this study using HISAT2. The gene-level read counts were then quantified using HTSeq-count (v0.13.5) (Anders et al., 2015). Significance testing for DEGs was performed using DESeq2 (Love et al., 2014): genes with an adjusted *p*-value (*P*_{adj}) < 0.05 and an absolute log₂(fold-change) > 1 were considered to be DEGs. To calculate gene expression values as TPM counts for each developmental stage, STAR was used (Dobin et al., 2013) for alignment and RSEM (Li and Dewey, 2011) for quantification. Expression profiling was conducted using Mfuzz (Kumar and Futschik, 2007), and heatmaps were visualized using the pheatmap R package.

Phylogenetic analysis of cell wall-degrading enzymes

DIAMOND (Buchfink et al., 2015) was used to align nematode protein sequences against the “CAZy.dmnd” database to identify CAZy family proteins. Based on previous reports

(Rai et al., 2015, Gahoi and Gautam, 2017), CWDEs from each nematode species were extracted according to their respective CAZy family classifications. These sequences were aligned using MAFFT (Katoh and Standley, 2013), and poorly aligned regions were trimmed using trimAl (Capella-Gutiérrez et al., 2009). Phylogenetic trees were reconstructed using IQ-TREE (Nguyen et al., 2015) and visualized with iTOL (Letunic and Bork, 2024).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the University of California Davis farm server system for their support with data analysis. S.C.G. was funded through NIH-NIGMS grant R35GM151194.

Availability of data and materials

Requests for further information and resources should be directed to the lead contact, Shahid Siddique (ssiddique@ucdavis.edu). This study did not generate any new reagents. All sequencing data have been deposited in the Sequence Read Archive (SRA) at the National Center for Biotechnology Information (NCBI). The Nanopore, HiFi, Illumina, Hi-C, and RNA-seq raw data for *P. vulnus* were deposited under BioProject PRJNA1283215 (31 SRA files); The assembled genome of *P. vulnus* has been deposited to GeneBank with the accession number JBSLOA000000000.1. The genomes of *Panagrellus* sp. LJ2284 and *Panagrolaimus* sp. JU765 were downloaded from NCBI under accession numbers GCA_024447225.1 and GCA_964245525.1, respectively. The RNA-seq data for *Panagrellus* sp. LJ2284 and *Panagrolaimus* sp. JU765 were downloaded from the SRA under accession numbers SRR22163891 and SRR5253561, respectively. The Hi-C data for *Panagrellus* sp. LJ2284 were downloaded from the SRA under accession number SRR23107840.

References

- ACOSTA N & MALEK R 1979. Influence of temperature on population development of eight species of *Pratylenchus* on soybean. *Journal of Nematology*, 11, 229. [PubMed: 19300639]
- ALMAGRO ARMENTEROS JJ, TSIRIGOS KD, SONDERBY CK, PETERSEN TN, WINTHER O, BRUNAK S, VON HEIJNE G & NIELSEN H 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*, 37, 420–423. [PubMed: 30778233]
- ALVAREZ-ORTEGA S, BRITO JA & SUBBOTIN SA 2019. Multigene phylogeny of root-knot nematodes and molecular characterization of *Meloidogyne nataliei* Golden, Rose & Bird, 1981 (Nematoda: Tylenchida). *Sci Rep*, 9, 11788. [PubMed: 31409860]
- ANDERS S, PYL PT & HUBER W 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *bioinformatics*, 31, 166–169. [PubMed: 25260700]
- ARORA D, HERNANDEZ AG, WALDEN KKO, FIELDS CJ & YAN G 2023. First Draft Genome Assembly of Root-Lesion Nematode *Pratylenchus scribneri* Generated Using Long-Read Sequencing. *Int J Mol Sci*, 24.
- BEETS I, ZELS S, VANDEWYER E, DEMEULEMEESTER J, CAERS J, BAYTEMUR E, COURTNEY A, GOLINELLI L, HASAKIOGULLARI I, SCHAFER WR, VERTES PE, MIRABEAU O & SCHOOF L 2023. System-wide mapping of peptide-GPCR interactions in *C. elegans*. *Cell Rep*, 42, 113058. [PubMed: 37656621]

- BUCHFINK B, XIE C & HUSON DH 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*, 12, 59–60. [PubMed: 25402007]
- CANTALAPIEDRA CP, HERNANDEZ-PLAZA A, LETUNIC I, BORK P & HUERTA-CEPAS J 2021. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol Biol Evol*, 38, 5825–5829. [PubMed: 34597405]
- CAPELLA-GUTIÉRREZ S, SILLA-MARTÍNEZ JM & GABALDÓN T 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25, 1972–1973. [PubMed: 19505945]
- CARDOSO JC, FELIX RC, FONSECA VG & POWER DM 2012. Feeding and the rhodopsin family g-protein coupled receptors in nematodes and arthropods. *Front Endocrinol (Lausanne)*, 3, 157. [PubMed: 23264768]
- CHALLIS R, RICHARDS E, RAJAN J, COCHRANE G & BLAXTER M 2020. BlobToolKit–interactive quality assessment of genome assemblies. *G3: Genes, Genomes, Genetics*, 10, 1361–1374. [PubMed: 32071071]
- CHEN C, WU Y, LI J, WANG X, ZENG Z, XU J, LIU Y, FENG J, CHEN H, HE Y & XIA R 2023. TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Mol Plant*, 16, 1733–1742. [PubMed: 37740491]
- CHEN SJ, LAM MQ, THEVARAJOO S, ABD MANAN F, YAHYA A & CHONG CS 2020. Genome analysis of cellulose and hemicellulose degrading *Micromonospora* sp. CP22. *3 Biotech*, 10, 160.
- CHEN Y, NIE F, XIE S-Q, ZHENG Y-F, DAI Q, BRAY T, WANG Y-X, XING J-F, HUANG Z-J & WANG D-P 2021. Efficient assembly of nanopore reads via highly accurate and intact error correction. *Nature Communications*, 12, 60.
- CHIHANI-HAMMAS N, HAJJI-HEDFI L, REGAIEG H, LARAYEDH A, BADISS A, QING Y & NAJET HR 2018. First report of *Pratylenchus vulnus* associated with apple in Tunisia. *J Nematol*, 50, 579–586. [PubMed: 31094160]
- DAI D, XIE C, ZHOU Y, BO D, ZHANG S, MAO S, LIAO Y, CUI S, ZHU Z, WANG X, LI F, PENG D, ZHENG J & SUN M 2023. Unzipped chromosome-level genomes reveal allopolyploid nematode origin pattern as unreduced gamete hybridization. *Nat Commun*, 14, 7156. [PubMed: 37935661]
- DAYI M 2024. Evolution of parasitism genes in the plant parasitic nematodes. *Scientific Reports*, 14, 3733. [PubMed: 38355886]
- 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, 15–21. [PubMed: 23104886]
- DU Z, ZHAO Y & LI N 2009. Genome-wide colonization of gene regulatory elements by G4 DNA motifs. *Nucleic Acids Res*, 37, 6784–98. [PubMed: 19759215]
- DUDCHENKO O, BATRA SS, OMER AD, NYQUIST SK, HOEGER M, DURAND NC, SHAMIM MS, MACHOL I, LANDER ES, AIDEN AP & AIDEN EL 2017. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science*, 356, 92–95. [PubMed: 28336562]
- DURAND NC, ROBINSON JT, SHAMIM MS, MACHOL I, MESIROV JP, LANDER ES & AIDEN EL 2016a. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst*, 3, 99–101. [PubMed: 27467250]
- DURAND NC, SHAMIM MS, MACHOL I, RAO SS, HUNTLEY MH, LANDER ES & AIDEN EL 2016b. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst*, 3, 95–8. [PubMed: 27467249]
- ENGLISH AC, RICHARDS S, HAN Y, WANG M, VEE V, QU J, QIN X, MUZNY DM, REID JG, WORLEY KC & GIBBS RA 2012. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS One*, 7, e47768. [PubMed: 23185243]
- ENRIQUEZ T & COLINET H 2019. Cold acclimation triggers major transcriptional changes in *Drosophila suzukii*. *BMC genomics*, 20, 413. [PubMed: 31117947]
- FOSU-NYARKO J & JONES MG 2016. Advances in Understanding the Molecular Mechanisms of Root Lesion Nematode Host Interactions. *Annu Rev Phytopathol*, 54, 253–78. [PubMed: 27296144]

- FRIEBE A, KLEVER W, SIKORA R & SCHNABL H 1998. Allelochemicals in root exudates of maize: effects on root lesion nematode *Pratylenchus zeae*. *Phytochemical signals and plant-microbe interactions*, 71–93.
- GABRIEL L, BRUNA T, HOFF KJ, EBEL M, LOMSADZE A, BORODOVSKY M & STANKE M 2024. BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. *Genome Res*, 34, 769–777. [PubMed: 38866550]
- GAHOI S & GAUTAM B 2017. Genome-wide analysis of Excretory/Secretory proteins in root-knot nematode, *Meloidogyne incognita* provides potential targets for parasite control. *Comput Biol Chem*, 67, 225–233. [PubMed: 28187376]
- GONZALEZ DE LA ROSA PM, THOMSON M, TRIVEDI U, TRACEY A, TANDONNET S & BLAXTER M 2021. A telomere-to-telomere assembly of *Osccheius tipulae* and the evolution of rhabditid nematode chromosomes. *G3 (Bethesda)*, 11.
- GRABHERR MG, HAAS BJ, YASSOUR M, LEVIN JZ, THOMPSON DA, AMIT I, ADICONIS X, FAN L, RAYCHOWDHURY R & ZENG Q 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature biotechnology*, 29, 644.
- HAAS BJ, SALZBERG SL, ZHU W, PERTEA M, ALLEN JE, ORVIS J, WHITE O, BUELL CR & WORTMAN JR 2008. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Genome Biol*, 9, R7. [PubMed: 18190707]
- HAN N, FAN S, ZHANG T, SUN H, ZHU Y, GONG H & GUO J 2020. SIHY5 is a necessary regulator of the cold acclimation response in tomato. *Plant Growth Regulation*, 91, 1–12.
- HOLDEN-DYE L & WALKER R 2011. Neurobiology of plant parasitic nematodes. *Invertebrate Neuroscience*, 11, 9–19. [PubMed: 21538093]
- JONES JT, HAEGEMAN A, DANCHIN EG, GAUR HS, HELDER J, JONES MG, KIKUCHI T, MANZANILLA-LOPEZ R, PALOMARES-RIUS JE, WESEMAEL WM & PERRY RN 2013. Top 10 plant-parasitic nematodes in molecular plant pathology. *Mol Plant Pathol*, 14, 946–61. [PubMed: 23809086]
- KARACA MS, YAVUZASLANOGLU E, IMRIZ G & SONMEZOGLU OA 2020. Molecular characterization of the *Pratylenchus vulnus* populations on cereals in Turkey. *J Nematol*, 52.
- KATO H & STANDLEY DM 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, 30, 772–80. [PubMed: 23329690]
- KIKUCHI T, EVES-VAN DEN AKKER S & JONES JT 2017. Genome Evolution of Plant-Parasitic Nematodes. *Annu Rev Phytopathol*, 55, 333–354. [PubMed: 28590877]
- KIM D, PAGGI JM, PARK C, BENNETT C & SALZBERG SL 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology*, 37, 907–915.
- KOUTSOVOULOS GD, GRANJEON NORIOT S, BAILLY-BECHET M, DANCHIN EGJ & RANCUREL C 2022. AvP: A software package for automatic phylogenetic detection of candidate horizontal gene transfers. *PLoS Comput Biol*, 18, e1010686. [PubMed: 36350852]
- KREPS JA, WU Y, CHANG HS, ZHU T, WANG X & HARPER JF 2002. Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. *Plant Physiol*, 130, 2129–41. [PubMed: 12481097]
- KROGH A, LARSSON B, VON HEIJNE G & SONNHAMMER EL 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*, 305, 567–80. [PubMed: 11152613]
- KRZYWINSKI M, SCHEIN J, BIROL I, CONNORS J, GASCOYNE R, HORSMAN D, JONES SJ & MARRA MA 2009. Circos: an information aesthetic for comparative genomics. *Genome Res*, 19, 1639–45. [PubMed: 19541911]
- KUD J, WANG W, GROSS R, FAN Y, HUANG L, YUAN Y, GRAY A, DUARTE A, KUHL JC, CAPLAN A, GOVERSE A, LIU Y, DANDURAND LM & XIAO F 2019. The potato cyst nematode effector RHA1B is a ubiquitin ligase and uses two distinct mechanisms to suppress plant immune signaling. *PLoS Pathog*, 15, e1007720. [PubMed: 30978251]
- KUMAR L & FUTSCHIK ME 2007. Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics*, 23, 5. [PubMed: 18084642]
- LETUNIC I & BORK P 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res*, 52, W78–W82. [PubMed: 38613393]

- LI B & DEWEY CN 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics*, 12, 1–16. [PubMed: 21199577]
- LI H & DURBIN R 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25, 1754–60. [PubMed: 19451168]
- LI H, HANDSAKER B, WYSOKER A, FENNELL T, RUAN J, HOMER N, MARTH G, ABECASIS G, DURBIN R & SUBGROUP GPDP 2009. The sequence alignment/map format and SAMtools. *bioinformatics*, 25, 2078–2079. [PubMed: 19505943]
- LIU C, SCHLÄPPI MR, MAO B, WANG W, WANG A & CHU C 2019. The bZIP 73 transcription factor controls rice cold tolerance at the reproductive stage. *Plant Biotechnology Journal*, 17, 1834–1849. [PubMed: 30811812]
- LIU X, BULLEY SM, VARKONYI-GASIC E, ZHONG C & LI D 2023. Kiwifruit bZIP transcription factor AcePosF21 elicits ascorbic acid biosynthesis during cold stress. *Plant physiology*, 192, 982–999. [PubMed: 36823691]
- LOVE MI, HUBER W & ANDERS S 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15, 1–21.
- MCKENRY M 1989. Damage and development of several nematode species in a plum orchard.
- MISTRY J, CHUGURANSKY S, WILLIAMS L, QURESHI M, SALAZAR GA, SONNHAMMER ELL, TOSATTO SCE, PALADIN L, RAJ S, RICHARDSON LJ, FINN RD & BATEMAN A 2021. Pfam: The protein families database in 2021. *Nucleic Acids Res*, 49, D412–D419. [PubMed: 33125078]
- MITREVA M, SMANT G & HELDER J 2009. Role of horizontal gene transfer in the evolution of plant parasitism among nematodes. *Horizontal gene transfer: Genomes in flux*, 517–535.
- 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, 268–74. [PubMed: 25371430]
- OU S, SU W, LIAO Y, CHOUGULE K, AGDA JRA, HELLINGA AJ, LUGO CSB, ELLIOTT TA, WARE D, PETERSON T, JIANG N, HIRSCH CN & HUFFORD MB 2019. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol*, 20, 275. [PubMed: 31843001]
- PERTEA G & PERTEA M 2020. GFF Utilities: GffRead and GffCompare. *F1000Res*, 9.
- PERTEA M, PERTEA GM, ANTONESCU CM, CHANG TC, MENDELL JT & SALZBERG SL 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*, 33, 290–5. [PubMed: 25690850]
- POLLACK S 1998. An Analysis of the Feasibility of Providing Federal Multi-Peril Crop Insurance to Nut Tree Growers: Almonds Hazelnuts.
- QING X, ZHANG YM, SUN S, AHMED M, LO WS, BERT W, HOLOVACHOV O & LI H 2024. Phylogenomic Insights into the Evolution and Origin of Nematoda. *Syst Biol*.
- RAI KM, BALASUBRAMANIAN VK, WELKER CM, PANG M, HII MM & MENDU V 2015. Genome wide comprehensive analysis and web resource development on cell wall degrading enzymes from phyto-parasitic nematodes. *BMC Plant Biology*, 15, 1–15. [PubMed: 25592487]
- RANALLO-BENAVIDEZ TR, JARON KS & SCHATZ MC 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun*, 11, 1432. [PubMed: 32188846]
- REHMAN S, GUPTA VK & GOYAL AK 2016. Identification and functional analysis of secreted effectors from phytoparasitic nematodes. *BMC Microbiol*, 16, 48. [PubMed: 27001199]
- ROMAN J & TRIANTAPHYLLOU A 1969. Gametogenesis and reproduction of seven species of *Pratylenchus*. *Journal of Nematology*, 1, 357. [PubMed: 19325697]
- SHEN WK, CHEN SY, GAN ZQ, ZHANG YZ, YUE T, CHEN MM, XUE Y, HU H & GUO AY 2023. AnimalTFDB 4.0: a comprehensive animal transcription factor database updated with variation and expression annotations. *Nucleic Acids Res*, 51, D39–D45. [PubMed: 36268869]
- SIDDIQUE S, RADAKOVIC ZS, HILTL C, PELLEGRIN C, BAUM TJ, BEASLEY H, BENT AF, CHITAMBO O, CHOPRA D, DANCHIN EGJ, GRENIER E, HABASH SS, HASAN MS, HELDER J, HEWEZI T, HOLBEIN J, HOLTERMAN M, JANAKOWSKI S, KOUTSOVOULOS GD, KRANSE OP, LOZANO-TORRES JL, MAIER TR, MASONBRINK RE, MENDY B, RIEMER E, SOBCZAK M, SONAWALA U, STERKEN MG, THORPE P, VAN

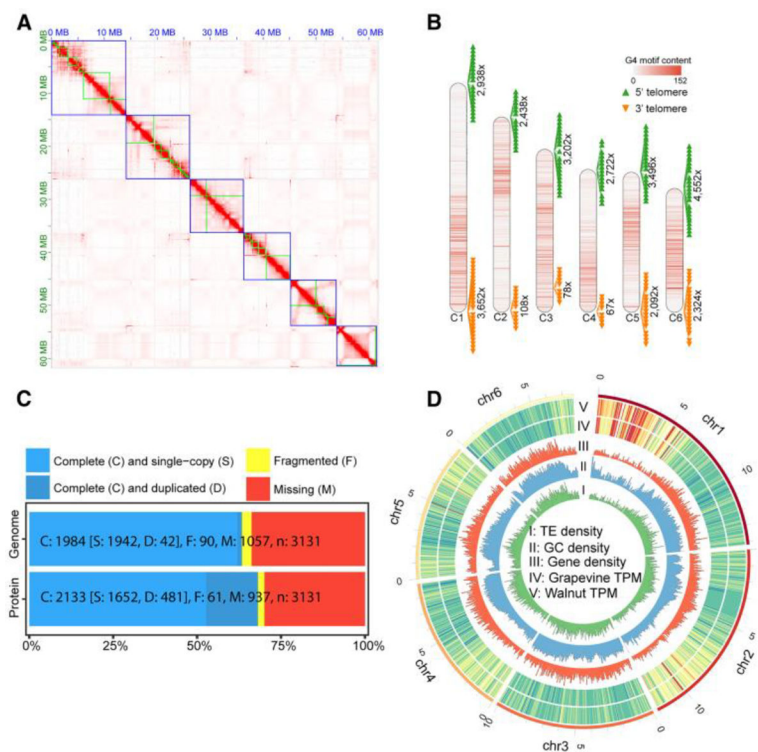
STEENBRUGGE JJM, ZAHID N, GRUNDLER F & EVES-VAN DEN AKKER S 2022. The genome and lifestage-specific transcriptomes of a plant-parasitic nematode and its host reveal susceptibility genes involved in trans-kingdom synthesis of vitamin B5. *Nat Commun*, 13, 6190. [PubMed: 36261416]

TARAILO-GRAOVAC M & CHEN N 2009. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics*, Chapter 4, 4 10 1–4 10 14.

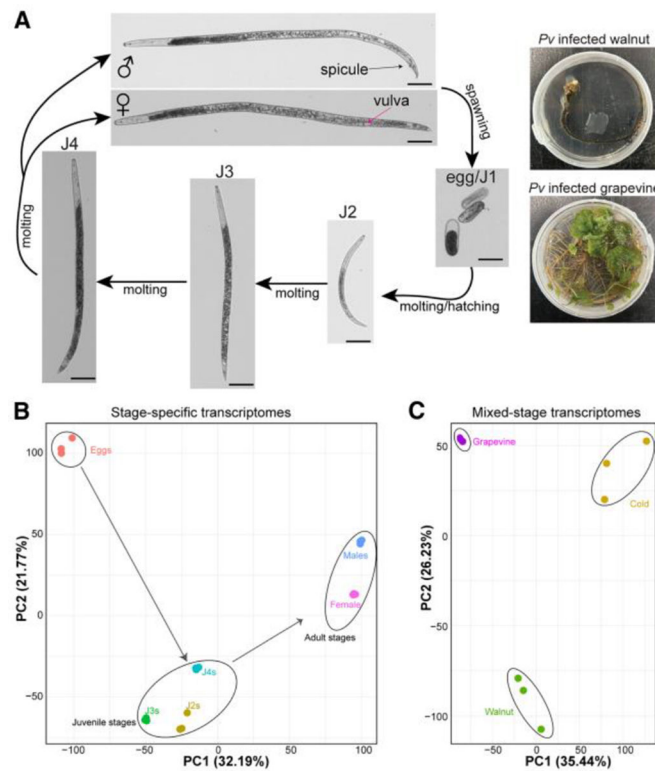
TOWSON A & LEAR B 1982. Effect of temperature on reproduction and motility of *Pratylenchus vulnus*. *Journal of Nematology*, 14, 602. [PubMed: 19295763]

WALKER BJ, ABEEL T, SHEA T, PRIEST M, ABOUELLIEL A, SAKTHIKUMAR S, CUOMO CA, ZENG Q, WORTMAN J, YOUNG SK & EARL AM 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*, 9, e112963. [PubMed: 25409509]

WANG Y, TANG H, DEBARRY JD, TAN X, LI J, WANG X, LEE T-H, JIN H, MARLER B & GUO H 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic acids research*, 40, e49–e49. [PubMed: 22217600]

**Fig. 1:**

Overview of the *Pratylenchus vulnus* genome. **A**, Hi-C contact heatmap showing a chromosome-scale assembly of the *P. vulnus* genome. The green boxes represent contigs, and the blue boxes represent pseudochromosomes. **B**, Chromosomal distribution of G4 DNA motifs and telomeric repeats (TTAGGC). Orange and yellow arrowheads mark the 5' and 3' telomeres, respectively. The heatmap shading indicates G4 motif content. Numerical values (x) indicate the number of occurrences of the telomeric repeat sequence at the corresponding position. **C**, BUSCO analysis showing genome completeness based on gene (top) and protein (bottom) datasets. Categories: complete and single copy (light blue), complete and duplicated (dark blue), fragmented (yellow), and missing (red). **D**, Circos plot displaying genomic features across the six *P. vulnus* chromosomes, displayed in five concentric circles: I, transposable element (TE) density; II, GC content; III, gene density; IV, gene expression in transcripts per million (TPM) on *Vitis vinifera* (grapevine); V, gene expression (TPM) on *Juglans regia* (walnut). In circles IV and V, blue indicates high expression, and red indicates low expression.

**Fig. 2:**

Gene expression dynamics in *Pratylenchus vulnus* (*Pv*). **A**, Representative micrographs of the six developmental stages of *P. vulnus*. Right: images of *P. vulnus*-infected walnut and grapevine roots. Scale bars: 50 μ m. **B**, Principal component analysis (PCA) of RNA sequencing (RNA-seq) data from biological replicates of individual developmental stages. Juvenile stages, adult stages, and eggs form distinct clusters. **C**, PCA of RNA-seq data from mixed-stage populations under three conditions: extracted from grapevine host, extracted from walnut host, and cold shock after extraction from grapevine host. B and C, Each point represents a biological replicate.

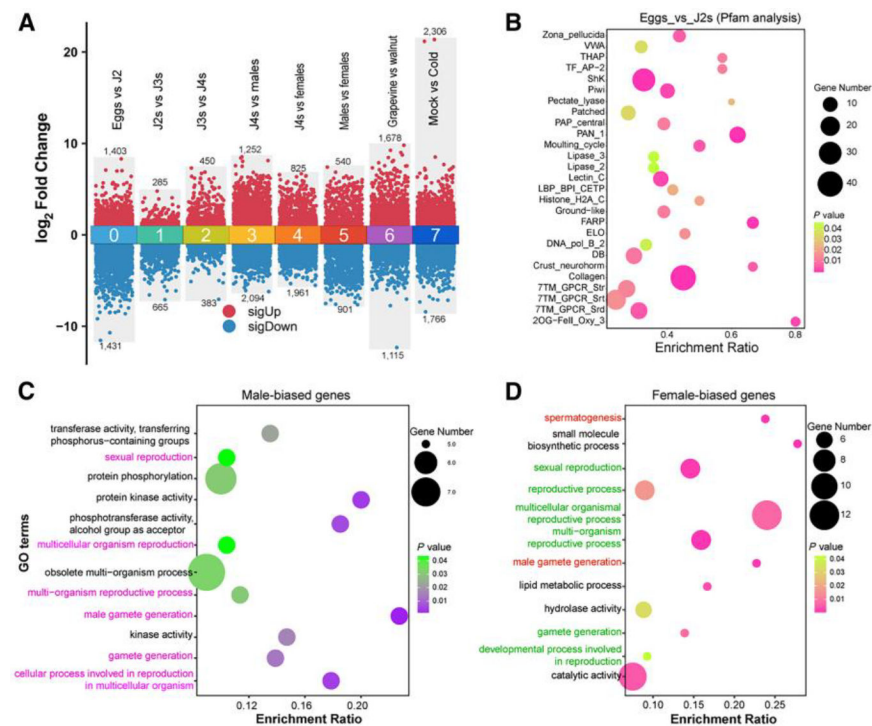


Fig. 3: Differential gene expression and functional enrichment in *Pratylenchus vulnus*. **A**, Differentially expressed genes (DEGs) between successive developmental stages, sexes, host conditions, and responses to acute cold in *P. vulnus*. Red and blue dots represent significantly upregulated (sigUp) and downregulated (sigDown) genes, respectively ($|\log_2 \text{fold change}| \geq 1$, adjusted $P < 0.05$). Numbers above and below each group indicate up- and downregulated gene counts. Gv, grapevine; Wt, walnut; Ct, control. **B**, Pfam domain enrichment analysis of DEGs between the egg and second-stage juvenile (J2) stages. Dot size corresponds to the number of genes per Pfam domain; color intensity indicates significance (*P* value). **C and D**, Gene Ontology (GO) enrichment analyses of male- and female-biased genes, respectively. Dot size corresponds to gene count; color reflects *P* value. Enrichment ratio indicates the proportion of genes in a given category among DEGs compared with that category's proportion in the background gene set.

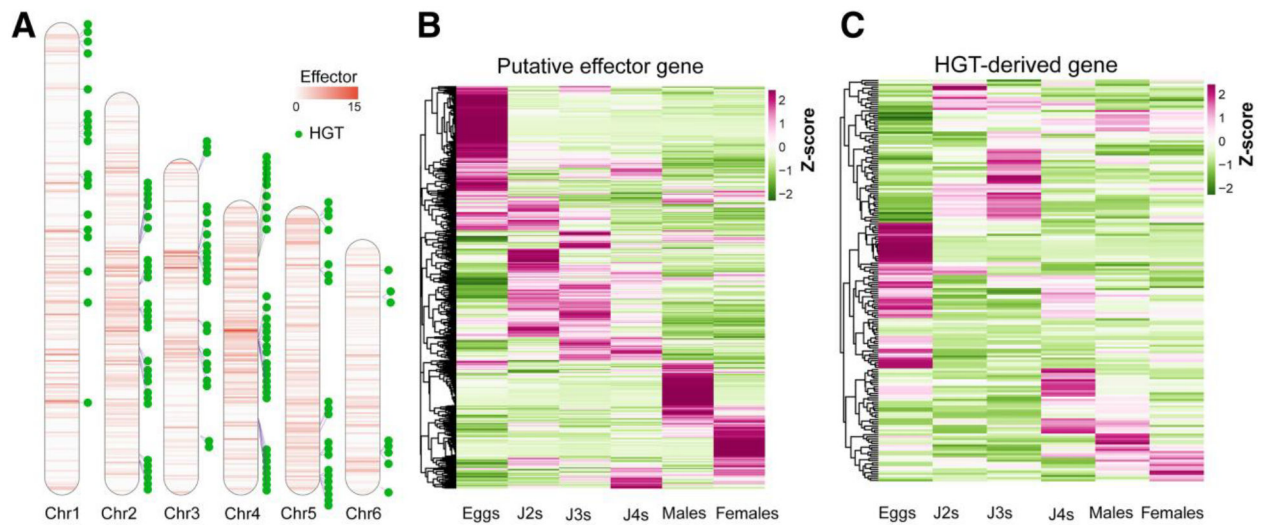
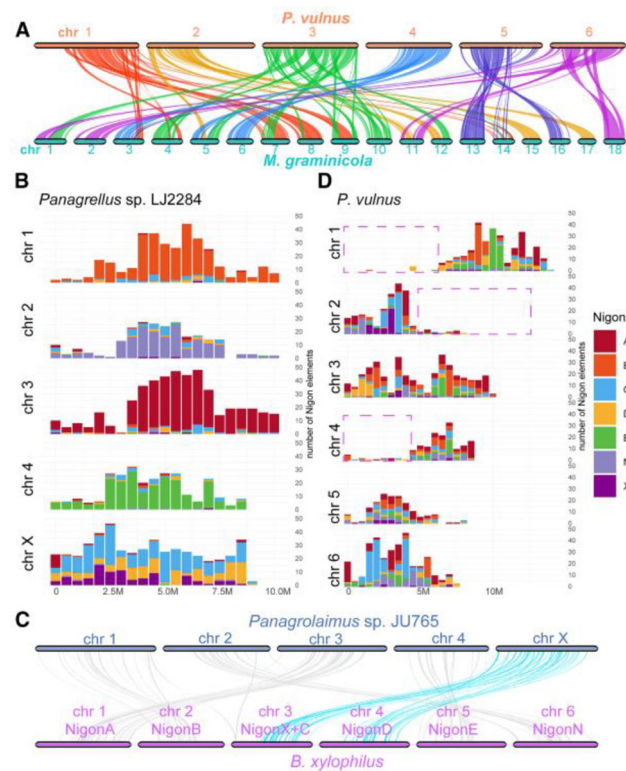
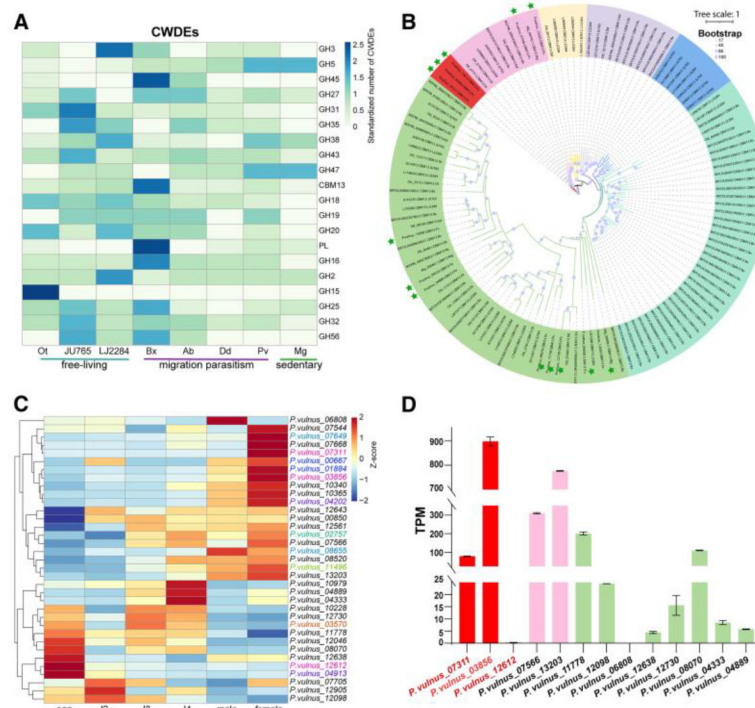


Fig. 4: Genomic distribution and developmental expression of candidate effector and horizontally transferred genes in *Pratylenchus vulnus*. **A**, Chromosomal distribution of putative effector genes (heatmap) and horizontal gene transfer (HGT)-derived genes (green dots) across the six *P. vulnus* chromosomes. Candidate effector gene density is shown as a heatmap (0 to 15 per window), and HGT-derived gene positions are marked as individual points. **B**, Heatmap showing expression (Z score-normalized transcripts per million [TPM]) of putative effector genes across developmental stages. **C**, Heatmap of expression patterns for HGT-derived genes across the same developmental stages.

**Fig. 5:**

Chromosomal collinearity and Nigon element distribution in *Pratylenchus vulnus* and related nematodes. **A**, Chromosomal collinearity between *P. vulnus* and *Meloidogyne graminicola*. Colored ribbons link orthologous genes, with colors corresponding to their *P. vulnus* chromosome of origin. **B**, Distribution of Nigon elements across chromosomes in *Panagrellus* sp. LJ2284. Most autosomes are dominated by a single Nigon element, whereas the X chromosome contains a mosaic of three. **C**, Syntenic relationships between the X chromosome of *Panagrolaimus* sp. JU765 and chromosomes 3 and 4 of *Bursaphelenchus xylophilus*. **D**, Distribution of Nigon elements across *P. vulnus* chromosomes. Each bar represents a genomic window and is color-coded by Nigon element identity. Regions without detectable Nigon assignment are highlighted with dashed purple boxes.

**Fig. 6:**

Characterization of cell wall-degrading enzyme (CWDE) genes across nematode species. **A**, Abundance of CWDE gene families across eight nematode species, normalized by min-max scaling. Species are grouped by ecological lifestyle. Free-living nematodes: *Oscheius tipulae* (Ot), *Panagrolaimus* sp. JU765, *Panagrellus* sp. LJ2284. Migratory plant parasites: *Bursaphelenchus xylophilus* (Bx), *Aphelenchoides besseyi* (Ab), *Ditylenchus dipsaci* (Dd), *Pratylenchus vulnus* (Pv). Sedentary endoparasite: *Meloidogyne graminicola* (Mg). Gene families include glycoside hydrolases (GHs), pectate lyases (PLs), and carbohydrate-binding modules (CBMs). **B**, Phylogenetic tree of CBM13 translated gene (protein) sequences from all eight nematode species. Colored branches represent distinct clades. Asterisks indicate *P. vulnus* genes that form phylogenetically distinct lineages, lacking close homologs in other nematodes. Tree constructed using IQ-TREE with 1,000 ultrafast bootstrap replicates. **C**, Heatmap showing stage-specific expression (Z score of transcripts per million [TPM]) for various *P. vulnus* CWDE gene families. Gene names are colored according to gene family: CBM13 (magenta), PL (purple), GH5 (blue), GH27 (cyan), GH38 (green), GH43 (lime green), GH47 (orange). Colored gene names indicate *P. vulnus*-specific genes that form phylogenetically distinct clades (i.e., without homologs in other nematodes), and genes labeled in black are conserved and cluster with homologs from other species. **D**, RNA sequencing-based expression of CBM13 genes in *P. vulnus* females. Colors match clades from panel B.

Table 1.Statistical summary for the *Pratylenchus vulnus* genome assembly and annotation

Feature	Value
Contig-level assembly	
Initial genome size (NECAT)	61.72 Mb
Polished genome size	61.67 Mb
Number of contigs	63
Contig N50	6.6 Mb
Genes	
Protein-coding genes	13,311
Transcripts (isoforms)	18,704
Transcripts with untranslated regions	24,250
Transcripts with Pfam domains	12,444
Transcripts with Gene Ontology terms	9,016
Chromosome-level assembly (Hi-C)	
Final genome size	61.71 Mb
Number of scaffolds	13
Number of chromosomes	6
Scaffold N50	10.1 Mb
Percentage of sequences anchored to chromosomes	99.88%
Chromosome size	
Chromosome 1	14,202,844 bp
Chromosome 2	12,102,447 bp
Chromosome 3	10,106,057 bp
Chromosome 4	8,859,133 bp
Chromosome 5	8,682,852 bp
Chromosome 6	7,683,060 bp