*J. Parasitol.*, 93(6), 2007, pp. 1343–1349


# EXPRESSED SEQUENCE TAG ANALYSIS OF GENE REPRESENTATION IN INSECT PARASITIC NEMATODE *HETERORHABDITIS BACTERIOPHORA*

**Xiaodong Bai, Parwinder S. Grewal*, Saskia A. Hogenhout, Byron J. Adams†, Todd A. Ciche‡, Randy Gaugler§, and Paul W. Sternberg‖**

Department of Entomology, The Ohio State University-OARDC, Wooster, Ohio 44691. *e-mail: grewal.4@osu.edu*

ABSTRACT: We compared *Heterorhabditis bacteriophora* GPS11 expressed sequence tags (ESTs) to the ESTs of animal-parasitic, human-parasitic, plant-parasitic, and free-living nematodes. We identified 127 previously nondescribed ESTs of which 119 had homologs in ESTs and 8 had homologs in proteins of free-living nematodes. These ESTs were assigned putative functions in transcription, signal transduction, cell cycle control, metabolism, information processing, and cellular processes, thereby providing better insight into *H. bacteriophora* metabolism, sex determination, and signal transduction. We also identified 36 *H. bacteriophora* ESTs that had significant similarities to ESTs of parasitic nematodes, but not to ESTs or proteins of free-living nematodes species. Among these are the ESTs encoding a centrin, an ankyrin-repeat containing protein, and a nuclear hormone receptor. Our analysis also revealed that parasitic nematode-specific ESTs in this *H. bacteriophora* data set had more homologs in animal-parasitic nematodes than those parasitizing humans or plants.

Insect parasitic (or entomopathogenic) nematodes, belonging to Heterorhabditidae and Steinernematidae, form unique models for the study of parasitism, pathogenicity, and symbiosis in general. These nematodes and their obligate bacterial symbionts form a highly species-specific and mutually beneficial relationship, i.e., heterorhabditids associating with *Photorhabdus* and steinernematids with *Xenorhabdus*, respectively (Boemare, 2002). Although the relationship is obligate, in the laboratory both partners can be cultured separately; the species-specific symbiotic relationship can be restored when combined, making genetic manipulation possible.

Insect parasitic nematodes and their symbiotic bacteria together are important biological control agents of insect pests (Grewal et al., 2005b). The infective juvenile (IJ) or dauer juvenile (DJ) is the only stage that can infect insect hosts. The IJs persist in soil in search of a suitable insect host and invade through natural body openings or the cuticle (Griffin et al., 2005). Upon reaching the hemocoel, the IJs release symbiotic bacteria, which multiply rapidly, killing the host via septicemia, usually within 24–48 hr (Dowds and Peters, 2002). The nematodes then feed on the symbiotic bacteria and complete 1 to 3 generations within the cadaver. When the food source depletes, IJs are formed, which exit the cadaver to search for new insect hosts (Poinar, 1990).

The potential of insect parasitic nematodes as biological control agents can be further improved with respect to IJ longevity, bacterial retention, tolerance to heat, ultraviolet radiation and desiccation, resistance to encapsulation in the hemocoel of some key insect pests, and trait stability (Grewal et al., 2005a). These improvements are now becoming feasible partly due to the determination of the complete genome sequences of 2 closely related nematodes, i.e., *Caenorhabditis elegans* (The *C. elegans* Sequencing Consortium, 1998) and *C. briggsae* (Stein et al., 2003), and of the bacterial symbiont of *Heterorhabditis bacteriophora* (*Photorhabdus luminescens* subsp. *laumondii* TT01) (Duchaud et al., 2003). Furthermore, the increase in the availability of information on the expressed sequence tags (ESTs) of many nematode species together with the initiation of *H. bacteriophora* complete genome sequence project (supported by the National Human Genome Research Institute and the United States Department of Agriculture [http://genome.wustl.edu/genome.cgi?GENOME=Heterorhabditis%20bacteriophora]), enhances possibilities for both comparative and functional genomics.

To pursue genomics in insect parasitic nematodes, we carried out a pilot-scale EST sequencing project for *H. bacteriophora* IJs. Useful information was revealed when the EST sequences were searched against WormBase and SwissProt (Sandhu et al., 2006). Here, we report the findings of a more comprehensive comparative analysis of these ESTs with ESTs of other nematodes.

## MATERIALS AND METHODS

### EST sequences

*Heterorhabditis bacteriophora* GPS11 ESTs in this study were from previously described EST data set (Sandhu et al., 2006). *Heterorhabditis bacteriophora* GPS11 EST sequences were further processed by removing the SMART primer sequences. EST sequences were deleted if their lengths were shorter than 50 nucleotides after removal of SMART primer sequences. The ESTs were clustered by removing redundant sequences and grouping partially aligned sequences, resulting in 988 unique *H. bacteriophora* GPS11 ESTs. ESTs of other nematodes were obtained from the GenBank Entrez nucleotide database by keyword searching. The nematode species whose ESTs were included in this study are summarized in Table I. They included 21 nonprimate animal-parasitic nematode (APN) species, 8 human-parasitic nematode (HPN) species, 18 plant-parasitic nematode (PPN) species, and 7 free-living nematode (FLN) species. These nematodes were classified into clades following Dorris et al. (1999). At the time the present report was written, the EST sequences totaled 747,259 in the nematodes described above.

### Computational analysis

All computational analyses were performed in a Linux workstation. The ESTs from APNs, HPNs, PPNs, and FLNs were combined together to form a database. The geninfo identifier (GI) numbers of the EST sequences for nematodes belonging to each of the 4 categories were separately collected into individual files. To identify the homologs of *H. bacteriophora* ESTs in ESTs of all nematodes in the 4 categories,

Received 28 February 2007; revised 20 April 2007; accepted 15 May 2007.

* To whom correspondence should be addressed.
† Department of Microbiology and Molecular Biology, Brigham Young University, Provo, Utah 84602.
‡ Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan 48824.
§ Department of Entomology, Rutgers University, New Brunswick, New Jersey 08901.
‖ Howard Hughes Medical Institute and Division of Biology, California Institute of Technology, Pasadena, California 91125.

TABLE I. Summary of GenBank ESTs included in this study.

| Species | Clade* | No. of ESTs |
|---|---|---|
| APNs | | |
| *Ancylostoma caninum* | V | 9,618 |
| *Ancylostoma ceylanicum* | V | 10,651 |
| *Angiostrongylus cantonensis* | V | 1,279 |
| *Ascaris suum* | III | 40,704 |
| *Brugia pahangi* | III | 28 |
| *Dictyocaulus viviparus* | V | 2 |
| *Dirofilaria immitis* | III | 4,005 |
| *Haemonchus contortus* | V | 21,967 |
| *Litomosoides sigmodontis* | III | 2,699 |
| *Nippostrongylus brasiliensis* | V | 3,781 |
| *Oesophagostomum dentatum* | V | 299 |
| *Ostertagia ostertagi* | V | 7,006 |
| *Parastrongyloides trichosuri* | IVa | 7,963 |
| *Parelaphostrongylus tenuis* | V | 99 |
| *Strongyloides ratti* | IVa | 14,761 |
| *Teladorsagia circumcincta* | V | 4,313 |
| *Toxocara canis* | III | 4,889 |
| *Trichinella spiralis* | I | 10,767 |
| *Trichostrongylus vitrinus* | V | 368 |
| *Trichuris muris* | I | 2,714 |
| *Trichuris vulpis* | I | 3,063 |
| HPNs | | |
| *Ascaris lumbricoides* | III | 1,822 |
| *Brugia malayi* | III | 26,215 |
| *Loa loa* | III | 27 |
| *Necator americanus* | V | 5,032 |
| *Onchocerca ochengi* | III | 60 |
| *Onchocerca volvulus* | III | 14,974 |
| *Strongyloides stercoralis* | IVa | 11,392 |
| *Wuchereria bancrofti* | III | 4,847 |
| PPNs | | |
| *Bursaphelenchus mucronatus* | IVb | 3,193 |
| *Bursaphelenchus xylophilus* | IVb | 13,327 |
| *Globodera mexicana* | IVb | 17 |
| *Globodera pallida* | IVb | 4,378 |
| *Globodera rostochiensis* | IVb | 11,851 |
| *Heterodera avenae* | IVb | 1 |
| *Heterodera glycines* | IVb | 24,438 |
| *Heterodera schachtii* | IVb | 2,818 |
| *Meloidogyne arenaria* | IVb | 5,018 |
| *Meloidogyne chitwoodi* | IVb | 12,218 |
| *Meloidogyne hapla* | IVb | 24,452 |
| *Meloidogyne incognita* | IVb | 19,935 |
| *Meloidogyne javanica* | IVb | 7,587 |
| *Meloidogyne paranaensis* | IVb | 3,710 |
| *Pratylenchus penetrans* | IVb | 1,928 |
| *Pratylenchus vulnus* | IVb | 2,485 |
| *Radopholus similis* | IVb | 1,154 |
| *Xiphinema index* | I | 9,351 |
| FLNs | | |
| *Caenorhabditis briggsae* | V | 2,424 |
| *Caenorhabditis elegans* | V | 346,064 |
| *Caenorhabditis japonica* | V | 218 |
| *Caenorhabditis remanei* | V | 20,292 |
| *Panagrolaimus davidi* | IVa | 1 |
| *Pristionchus pacificus* | V | 14,663 |
| *Zeldia punctata* | IVb | 391 |
| Total | | 747,259 |

* Clade classification adopted from Dorris et al. (1999).

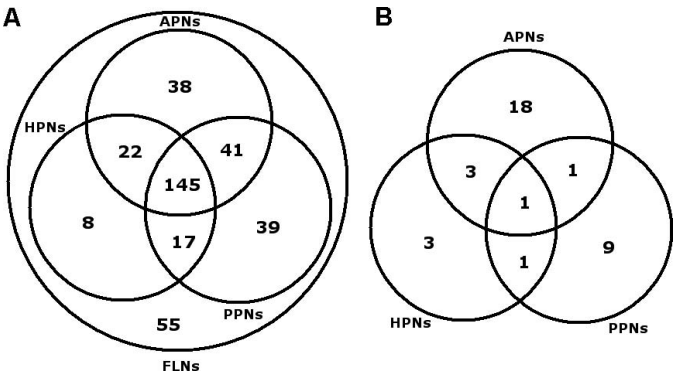


FIGURE 1. Venn diagram depicting the distribution of the number of *Heterorhabditis bacteriophora* GPS11 ESTs common in FLNs (**A**) and parasitic nematode-specific ESTs (**B**) that shared similarity to ESTs of FLNs, APNs, HPNs, and PPNs.

we searched the EST database using tBLASTx algorithm (Altschul et al., 1997) and the −1 option to restrict the search to list of GIs for nematodes in each category. Doing so made the E (expectation) values in the results comparable for individual searches, as the parameters for E value calculation are the same. The BLAST search results were parsed by using in-house written scripts based on bioperl (Stajich et al., 2002), and they were further analyzed manually to identify the matches to *C. elegans* genes in WormBase (Schwarz et al., 2006). To identify *H. bacteriophora* ESTs, the sequences were searched against the GenBank nonredundant (nr) database using BLASTx and BLASTn algorithms. The matches in BLAST searches with E values lower than $10^{-4}$ were considered significant. The ESTs were searched against 5′-untranslated region (UTR) and 3′-UTR databases derived from the nr UTR sequences obtained from UTResource (http://www.ba.itb.cnr.it/UTR/UTRHome.html; Pesole and Liuni, 1999) using BLASTn algorithm.

## RESULTS

The *H. bacteriophora* ESTs searched against the EST sequences from APNs, HPNs, PPNs, and FLNs using tBLASTx algorithms are summarized in Table II. Among the 988 *H. bacteriophora* ESTs, 569 (57.6%) had no significant similarities to other nematode ESTs, whereas 365 (36.9%) had significant similarities to ESTs of FLNs and 54 (5.5%) to ESTs of only parasitic nematodes. The search against UTR databases revealed that 7 ESTs had significant matches, and the rest had no significant matches to known UTRs.

Among the 365 *H. bacteriophora* ESTs having significant similarities to ESTs of FLNs, 55 had significant similarities to ESTs of only FLNs, and the other 310 had significant similarities to ESTs of both free-living and parasitic nematodes. Further categorization of *H. bacteriophora* ESTs based on homologies to APNs, HPNs, or PPNs (Fig. 1A) showed that the numbers of *H. bacteriophora* ESTs similar to ESTs of APNs and PPNs are about equal, and both slightly outnumber those similar to ESTs of HPNs.

Current analysis identified 119 additional *H. bacteriophora* ESTs having significant similarities to ESTs from FLNs (Table III). Among these, 28 had significant similarities to ESTs of only FLNs, whereas 91 had significant similarities to ESTs of both free-living and parasitic nematodes. Based on the results from the BLASTx search against the GenBank protein nr database, 48 of these 119 ESTs were assigned tentative functions (Table III; Supplementary Table I available at [http://www2.oardc.ohio-state.edu/nematodes/genomics/publications/

TABLE II. Summary of sequence similarity search results of *Heterorhabditis bacteriophora* GPS11 ESTs.

| Category | No. of ESTs (% of total) |
|---|---|
| ESTs not similar to any nematode ESTs | 569 (57.6) |
| ESTs similar to ESTs of FLNs | 365 (36.9) |
| ESTs similar to ESTs of FLNs | 54 (5.5) |
| ESTs similar to ESTs of both FLNs and parasitic nematodes | 310 (31.4) |
| ESTs not similar to ESTs of FLNs | 55 (5.6) |
| ESTs not similar to ESTs of FLNs but similar to proteins of FLNs | 15 (1.5) |
| ESTs similar to ESTs of parasitic nematodes but not to ESTs or proteins of FLNs | 39 (3.9) |
| Total | 988 |

JParasitol07.html]), including metabolism, information storage and processing, cellular processes and signaling, and general function prediction. The remaining 71 ESTs were not assigned specific functions but further categorized as (1) matching to uncharacterized, unnamed, or hypothetical proteins from *C. elegans* (46 ESTs); (2) matching to proteins from non-nematode organisms (12 ESTs); (3) vector contamination (1 EST); and (4) without significant BLASTx matches (12 ESTs) (Table III).

Among the 54 *H. bacteriophora* ESTs having no significant similarities to any ESTs of FLNs, the deduced protein sequences of 15 ESTs had significant similarities to proteins from *C. elegans* or *C. briggsae*. Among these 15 ESTs, 7 have been described previously (Sandhu et al., 2006), including HbGPS11.10A10, HbGPS11.10B02, HbGPS11.14B05, HbGPS11.14F12, HbGPS11.14L16, HbGPS11.6L01, and HbGPS11.6P07. The other 8 are summarized in Supplementary Table I available at (http://www2.oardc.ohio-state.edu/nematodes/genomics/publications/JParasitol07.html), together with ESTs having greatest similarities to ESTs from FLNs. The remaining 39 *H. bacteriophora* ESTs had no significant similarities to either ESTs or proteins from FLNs, while having significant similarities to ESTs in at least 1 category of parasitic nematodes derived mostly from infective stages, such as third stage (L3) larvae of APN *Ancylostoma caninum* and microfilaria of HPN *Brugia malayi* (Table IV).

Among these 39 ESTs, HbGPS11.6O16, HbGPS11.6G19, and HbGPS11.6M13 seemed to be of bacterial origin. The remaining 36 *H. bacteriophora* ESTs had significant similarity to ESTs of only parasitic nematodes, but not to those of the FLNs, therefore, designated as parasitic-nematode specific ESTs (Table IV). Six of these ESTs had significant similarities to proteins in GenBank nr databases. Among the 36 parasitic-nematode–specific ESTs, 18 (50%) had similarities only to ESTs of APNs, 9 (25%) had similarities only to PPNs, and 3 (8.3%) had simi-

TABLE III. Summary of novel *Heterorhabditis bacteriophora* GPS11 ESTs.

| Category | No. of novel *H. bacteriophora* GPS11 ESTs | |
|---|---|---|
| | with matches to ESTs of FLNs | without matches to ESTs of FLNs |
| 1. Metabolism | 14 | 1 |
| 1.1 Energy production and conversion | 4 | 0 |
| 1.2 Secondary metabolites biosynthesis, transport, and catabolism | 3 | 0 |
| 1.3 Amino acid transport and metabolism | 1 | 0 |
| 1.4 Nucleotide transport and metabolism | 2 | 0 |
| 1.5 Carbohydrate transport and metabolism | 3 | 1 |
| 1.6 Inorganic ion transport and metabolism | 1 | 0 |
| 2. Information storage and processing | 11 | 1 |
| 2.1 Transcription | 3 | 1 |
| 2.2 Translation, ribosomal structure, and biogenesis | 2 | 0 |
| 2.3 RNA processing and modification | 6 | 0 |
| 3. Cellular processes and signaling | 11 | 3 |
| 3.1 Posttranslational modification, protein turnover, chaperons | 2 | 0 |
| 3.2 Signal transduction mechanisms | 8 | 2 |
| 3.3 Cell cycle control, cell division, chromosome partitioning | 1 | 1 |
| 4. General function prediction | 12 | 0 |
| 5. No function assignment | 71 | 3 |
| 5.1 Matching to uncharacterized, unnamed or hypothetical proteins from *C. elegans* | 46 | 3 |
| 5.2 Matching to proteins from non-nematode organisms | 12 | 0 |
| 5.3 Vector contamination | 1 | 0 |
| 5.4 Without significant BLASTx matches | 12 | 0 |

TABLE IV. *Heterorhabditis bacteriophora* GPS11 ESTs that only matched to ESTs of parasitic nematodes.

| | Top tBLASTx match against ESTs of APNs | | | | Top tBLASTx match against ESTs of HPNs | | | |
|---|---|---|---|---|---|---|---|---|
| ESTs | Acc. no. | Organism* | Stage | E value | Acc. no. | Organism* | Stage | E value |
| *H. bacteriophora* ESTs potentially involved in parasitism | | | | | | | | |
| HbGPS11.6B13 | — | — | — | — | — | — | — | — |
| HbGPS11.6C10 | CB015842 | *Ha. contortus* | Adult | 4.00E-18 | — | — | — | — |
| HbGPS11.6C13 | CA957026 | *Ha. contortus* | — | 2.00E-05 | — | — | — | — |
| HbGPS11.6D11 | BQ693696 | *Ta. spiralis* | Adult | 5.00E-06 | — | — | — | — |
| HbGPS11.6D13 | CB037049 | *Te. circumcincta* | Adult | 1.00E-13 | — | — | — | — |
| HbGPS11.6E01 | AW627116 | *An. caninum* | L3 | 1.00E-07 | — | — | — | — |
| HbGPS11.6E23 | BQ482254 | *D. immitis* | Adult | 1.00E-06 | — | — | — | — |
| HbGPS11.6H04 | — | — | — | — | BU088675 | *Ne. americanus* | L3 | 4.00E-10 |
| HbGPS11.6H05 | BF228414 | *As. suum* | Adult | 1.00E-05 | — | — | — | — |
| HbGPS11.6H09 | AW181703 | *An. caninum* | L3 | 6.00E-08 | BU087062 | *Ne. americanus* | L3 | 4.00E-08 |
| HbGPS11.6I03 | — | — | — | — | — | — | — | — |
| HbGPS11.6I21 | BM279529 | *Ni. brasiliensis* | Adult | 5.00E-21 | AW191554 | *Br. malayi* | Microfilaria | 4.00E-11 |
| HbGPS11.6J06 | CA957511 | *Ha. contortus* | — | 8.00E-09 | — | — | — | — |
| HbGPS11.6K11 | — | — | — | — | — | — | — | — |
| HbGPS11.6L04 | BF423206 | *Ha. contortus* | Adult | 2.00E-10 | BU089164 | *Ne. americanus* | L3 | 2.00E-06 |
| HbGPS11.6L11 | DN190749 | *Ag. cantonensis* | — | 4.00E-06 | — | — | — | — |
| HbGPS11.6L17 | BF250079 | *An. caninum* | L3 | 1.00E-07 | — | — | — | — |
| HbGPS11.6L23 | BF250165 | *An. caninum* | L3 | 1.00E-11 | — | — | — | — |
| HbGPS11.6M09 | CB036921 | *Te. circumcincta* | Adult | 6.00E-05 | — | — | — | — |
| HbGPS11.10B20 | — | — | — | — | BE580982 | *S. stercoralis* | — | 4.00E-27 |
| HbGPS11.10E05 | — | — | — | — | — | — | — | — |
| HbGPS11.10E09 | — | — | — | — | — | — | — | — |
| HbGPS11.10G04 | — | — | — | — | — | — | — | — |
| HbGPS11.10L03 | BQ667089 | *An. caninum* | L3 | 7.00E-06 | — | — | — | — |
| HbGPS11.10N13 | BI744312 | *An. caninum* | L3 | 4.00E-13 | — | — | — | — |
| HbGPS11.10N17 | BM138832 | *Ha. contortus* | Adult | 1.00E-07 | — | — | — | — |
| HbGPS11.14C02 | — | — | — | — | BU666114 | *Ne. americanus* | L3 | 2.00E-12 |
| HbGPS11.14D15 | — | — | — | — | — | — | — | — |
| HbGPS11.14G12 | — | — | — | — | BU087348 | *Ne. americanus* | L3 | 7.00E-06 |
| HbGPS11.14H18 | — | — | — | — | — | — | — | — |
| HbGPS11.14L05 | CB190383 | *An. ceylanicum* | Adult | 3.00E-05 | — | — | — | — |
| HbGPS11.14L17 | CB099921 | *Ha. contortus* | — | 2.00E-07 | — | — | — | — |
| HbGPS11.14N09 | BF250079 | *An. caninum* | L3 | 7.00E-14 | — | — | — | — |
| HbGPS11.14O08 | BF250513 | *An. caninum* | L3 | 9.00E-07 | — | — | — | — |
| HbGPS11.14O12 | BI397360 | *S. ratti* | L1 | 7.00E-07 | — | — | — | — |
| HbGPS11.14P05 | BI743949 | *P. trichosuri* | | 2.00E-05 | BG224956 | *S. stercoralis* | — | 2.00E-05 |
| HbGPS11.14P20 | — | — | — | — | — | — | — | — |
| *H. bacteriophora* ESTs of bacterial origin | | | | | | | | |
| HbGPS11.6G19 | — | — | — | — | — | — | — | — |
| HbGPS11.6M13 | BM130215 | *An. caninum* | L3 | 3.00E-05 | BE202334 | *On. volvulus* | Microfilaria | 7.00E-05 |
| HbGPS11.6O16 | BI781487 | *As. suum* | — | 1.00E-06 | BE581808 | *S. stercoralis* | — | 1.00E-05 |

* The genus names of the organisms were abbreviated as follows: *Ag.*, *Angiostrongylus*; *An.*, *Ancylostoma*; *As.*, *Ascaris*; *Br.*, *Brugia*; *Bu*, *Bursaphelenchus*; *D.*, *Dirofilaria immitis*; *G.*, *Globodera*; *Ha.*, *Haemonchus*; *He.*, *Heterodera*; *M.*, *Meloidogyne*; *Ne.*, *Necator*; *Ni.*, *Nippostrongylus*; *On.*, *Onchocerca*; *Os.*, *Ostertagia*; *P.*, *Parastrongyloides*; *S.*, *Strongyloides*; *Ta.*, *Trichinella*; *Te.*, *Teladorsagia*; *Tr.*, *Trichuris*; *X.*, *Xiphinema.*

larities only to human-parasitic nematodes. The rest had similarities to ESTs from nematodes belonging to at least 2 categories (Fig. 1B).

## DISCUSSION

Previously, we reported *H. bacteriophora* 283 EST sequences having significant similarities to *C. elegans* proteins in WormBase (Sandhu et al., 2006). Present analysis yielded another 119 *H. bacteriophora* ESTs having significant similarities to ESTs from FLNs, including those encoding 4 transcription factors (HbGPS11.6A21, HbGPS11.10H22, HbGPS11.2L15, and HbGPS11.6E22), an F-box–containing protein (HbGPS11.14P23), a homology of human α subunit of sarcoglycan complex (HbGPS11.10I15), and other proteins.

HbGPS11.6A21 had high similarity (E value = 4e-34) to *C. elegans* doublesex-family transcription factor Y67D8A.3 related to *Drosophila doublesex* and *C. elegans mab-3*. Similar *doublesex*- or *mab-3*–related transcription factors have been identified in invertebrates such as the fruit fly *Drosophila melanogaster* (Burtis and Baker, 1989) and the Queensland fruit

TABLE IV. Extended.

| Top tBLASTx match against ESTs of PPNs | | | | Top BLASTx match against GenBank nr database | | |
|---|---|---|---|---|---|---|
| Acc. no. | Organism* | Stage | E value | Acc. no. | Description | E value |
| CV509972 | *X. index* | — | 8.00E-11 | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | XP_741933 | Centrin (*Plasmodium chabaudi chabaudi*) | 7.00E-08 |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| CV509439 | *X. index* | — | 4.00E-09 | CAF89842 | Unnamed protein product (*Tetraodon nigroviridis*) | 6.00E-05 |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| CV511724 | *X. index* | — | 4.00E-07 | XP_795809 | PREDICTED: similar to synaptobrevinlike 1, partial (*Strongylocentrotus purpuratus*) | 7.00E-10 |
| BE578503 | *M. javanica* | — | 1.00E-07 | CAJ82965 | Calcyclin binding protein (*Xenopus tropicalis*) | 1.00E-12 |
| — | — | — | — | — | No matches found | — |
| BM415851 | *G. pallida* | Mixed | 8.00E-05 | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| EE270325 | *G. rostochiensis* | — | 1.00E-07 | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | EAA12444 | ENSANGP00000006608 (*Anopheles gambiae* str. PEST) | 2.00E-11 |
| BM415650 | *G. pallida* | Mixed | 7.00E-05 | — | No matches found | — |
| CB825560 | *He. glycines* | J3 | 3.00E-24 | — | No matches found | — |
| CV511276 | *X. index* | — | 1.00E-05 | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| CF358052 | *M. arenaria* | J2 | 7.00E-08 | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| BI863274 | *M. javanica* | — | 1.00E-13 | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | CAG01948 | Unnamed protein product (*Tetraodon nigroviridis*) | 1.00E-11 |
| — | — | — | — | — | No matches found | — |
| CB379493 | *He. glycines* | — | 3.00E-08 | — | No matches found | — |
| CK349410 | *He. glycines* | — | 3.00E-05 | — | No matches found | — |
| — | — | — | — | — | No matches found | — |
| — | — | — | — | CAE12886 | Isoleucine tRNA synthetase (*Photorhabdus luminescens* subsp. *laumondii* TTO1) | 4.00E-84 |

fly *Bactrocera tryoni* (Shearman and Frommer, 1998), and in vertebrates such as humans (Raymond, Parker et al., 1999), chickens (Raymond, Kettlewell et al., 1999), and mice (Kim et al., 2003). These transcription factors are required for the control of somatic sexual differentiation in invertebrates and for testis development in vertebrates. However, the role of this gene in the sex determination of nematodes has not been investigated. Nonetheless, the presence of the homologs of this gene in *H. bacteriophora*, APNs, HPNs, PPNs, and FLNs implied the importance of this transcription factor for nematodes.

We have identified 3 ESTs of bacterial origin, which may be the result of contamination from symbiotic or other bacteria. HbGPS11.6O16 encoded an isoleucine tRNA synthetase. HbGPS11.6G19 matched a region in the linear chromosome of soil bacterium *Agrobacterium tumefaciens* (GenBank AE009348). HbGPS11.6M13 had similarity to the 16S ribosomal RNA gene from an uncultured soil bacterium (GenBank DQ378206). These genes may be the result of contamination from symbiotic or other bacteria, although oligo(dT) primers have been used in the cDNA library construction (Sandhu et al., 2006). However, we cannot rule out the possibility that at least some of these genes were acquired by the nematodes through horizontal gene transfer (HGT) from bacteria, because all 3 genes had similarities to independently obtained ESTs

from other 5 nematode species. Among these 5, *Onchocerca volvulus* harbors symbiotic *Wolbachia* (Pearlman and Gillette-Ferguson, 2007), whereas the other 4 species have no known symbiotic bacteria. However, these 4 species may encounter bacteria in the soil. In addition, HPN *Strongyloides stercoralis* and APNs *Ancylostoma caninum* and *Ascaris suum* may encounter bacteria in their invertebrate intermediate hosts. Although no HGT is known for APNs or HPNs, it has been suggested that PPNs have acquired bacterial genes such as those encoding cell wall-degrading enzymes via HGT (Scholl et al., 2003; Jones et al., 2005).

We have identified 36 *H. bacteriophora* ESTs having similarity to ESTs of only parasitic nematodes, but not to those of the FLNs, suggesting that they are involved in parasitic nematode-specific functions, e.g., parasitism. Four were assigned putative functions based on sequence similarity to proteins identified in other organisms. For example, HbGPS11.6D11 was similar to centrin protein from *Plasmodium chabaudi chabaudi* (Hall et al., 2005). Centrin is found in many eukaryotes, and it is known as a versatile molecule that has been adapted to perform diverse cellular functions, including signal transduction mechanisms (Hall et al., 2005), cell division (Zamora and Marshall, 2005), and nuclear mRNA export (Fischer et al., 2004). HbGPS11.10B20 had significant similarity (2e-11) to a protein from malaria-transmitting mosquito *Anopheles gambiae* (GenBank EAA12444). This protein contained a domain of ankyrin (ANK) repeats that mediate protein–protein interactions across diverse protein families. HbGPS11.14O12 had significant similarity (1e-11) to an unnamed protein product (GenBank CAG01948) from the green spotted pufferfish, *Tetraodon nigroviridis*, which contained domains of a c4 zinc finger in nuclear hormone receptor and a ligand binding domain of hormone receptors.

Because APNs, HPNs, and PPNs differ in the mechanics of parasitism, we analyzed *H. bacteriophora* ESTs potentially involved in parasitism in the context of phylogeny. Based on established phylogenetic relationships (Blaxter et al., 1998; Dorris et al., 1999), *H. bacteriophora*, as a member of clade V, is closely related to a group of APNs in clade V. However, *H. bacteriophora* is distantly related to APNs and HPNs in clade III. This observation that *H. bacteriophora* shared more similar genes with several APNs is consistent with the phylogenetic relationship of *H. bacteriophora* and these APNs. It should be noted that the decrease in genes shared with HPNs could be the result of underrepresented HPN data set (64,369 ESTs compared with 150,976 for APNs and 147,861 for PPNs). Although insect parasitic *H. bacteriophora* is distantly related to PPNs in clade IVb, both have stages in their life cycles dwelling in the soil, making it possible that they may share certain genetic elements. However, it should be noted that the EST distribution among the nematodes may change when additional EST data become available.

Although parasitism has independently evolved numerous times, resulting in substantial biological diversity within each group of parasitic nematodes (Jasmer et al., 2003), the insect parasitic *H. bacteriophora*, like APNs, HPNs, and PPNs, needs to overcome common obstacles in establishing itself as a successful parasite. The detailed comparison of parasitism among these parasitic nematodes will shed light on the parasitic mechanisms of *H. bacteriophora*. APNs and HPNs generally achieve infection via oral route, skin penetration, or injection by intermediate hosts during feeding (Jasmer et al., 2003), whereas most PPNs directly invade specific parts of plants, including roots, stems, or leaves (Dropkin, 1989). In contrast to other parasitic nematodes, all PPNs have an oral stylet for perforation of plant cell walls and ingestion of cytoplasm from the feeding site (Jasmer et al., 2003). Despite mechanical mechanisms, host invasion and migration seem to be facilitated through the use of enzymes to degrade structural components of tissues, cells, and cell walls in all parasitic nematodes (Jasmer et al., 2003). *Heterorhabditis bacteriophora* invades insect hosts via cuticle or natural openings, such as mouth, anus, and spiracles. Therefore, penetration of nematodes through the cuticle, petritrophic membrane of the gut wall, or spiracles may be facilitated by enzymatic activities as well as mechanical mechanisms. Previously, we have identified several proteases that may play an important role in insect pathogenesis (Sandhu et al., 2006). The ongoing full-scale EST project as part of the *H. bacteriophora* genome sequencing effort should help identify the participating enzymes.

In the host–parasite interaction, parasitic nematodes may modify a host's defense system and interfere with host cell processes and gene expression (Jasmer et al., 2003). Metalloprotease activity in excretory/secretory products of HPN *Necator americanus* may interrupt host signal transduction (Culley et al., 2000), and PPNs may secrete antioxidants to counteract plant host-generated reactive oxygen species (Henkle-Duhrsen and Kampkotter, 2001). Specifically, PPNs modify plant cells into giant feeding cells, creating a syncytium, and they also modulate host gene expression and cell cycle regulation (Gheysen and Fenoll, 2002). Disruption of even 1 of these genes via host-mediated RNA interference has been shown to disrupt parasitism (Huang et al., 2003). In the insect parasitic nematode *H. bacteriophora*, host–parasite interactions also involve its symbiotic bacterium *P. luminescens*. To antagonize the defense mechanisms of insect hosts, *H. bacteriophora* produces a specific extracellular protease that degrades cecropin, a broad-spectrum antibacterial protein generated by insects (Jarosz, 1998). Such a mechanism allows the establishment of bacterial infection, which is the key element of *H. bacteriophora* life cycle and insecticidal ability. The functional studies of these parasitic nematode-specific genes of *H. bacteriophora* identified in our EST analysis await the development of genetic tools such as stable transformation or RNA interference in *H. bacteriophora*.

## ACKNOWLEDGMENTS

This research is supported by a grant from the Interagency National Science Foundation (NSF)/United States Department of Agriculture (USDA)–Cooperative State Research, Education, and Extension Service (CSREES) Microbial Genome Sequencing Program.

## LITERATURE CITED

ALTSCHUL, S. F., T. L. MADDEN, A. A. SCHAFFER, J. ZHANG, Z. ZHANG, W. MILLER, AND D. J. LIPMAN. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Research **25:** 3389–3402.

BLAXTER, M. L., P. DE LEY, J. R. GAREY, L. X. LIU, P. SCHELDEMAN, A. VIERSTRAETE, J. R. VANFLETEREN, L. Y. MACKEY, M. DORRIS, L. M. FRISSE, J. T. VIDA, AND W. K. THOMAS. 1998. A molecular evolutionary framework for the phylum Nematoda. Nature **392:** 71–75.

BOEMARE, N. 2002. Interactions between the partners of the entomo-

pathogenic bacterium nematode complexes, *Sterinernema-Xenorhabdus* and *Heterorhabditis-Photorhabdus*. Nematology **4:** 601–603.

BURTIS, K. C., AND B. S. BAKER. 1989. *Drosophila* doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. Cell **56:** 997–1010.

CULLEY, F. J., A. BROWN, D. M. CONROY, I. SABROE, D. I. PRITCHARD, AND T. J. WILLIAMS. 2000. Eotaxin is specifically cleaved by hookworm metalloproteases preventing its action in vitro and in vivo. Journal of Immunology **165:** 6447–6453.

DORRIS, M., P. DE LEY, AND M. L. BLAXTER. 1999. Molecular analysis of nematode diversity and the evolution of parasitism. Parasitology Today **15:** 188–193.

DOWDS, B. C. A., AND A. PETERS. 2002. Virulence mechanisms. *In* Entomopathogenic nematology, R. Gaugler (eds.). CABI Publishing, Oxon, U.K., p. 79–98.

DROPKIN, V. H. 1989. Introduction to plant nematology. Wiley, New York, New York, 303 p.

DUCHAUD, E., C. RUSNIOK, L. FRANGEUL, C. BUCHRIESER, A. GIVAUDAN, S. TAOURIT, S. BOCS, C. BOURSAUX-EUDE, M. CHANDLER, J. F. CHARLES, E. DASSA, R. DEROSE, S. DERZELLE, G. FREYSSINET, S. GAUDRIAULT, C. MÉDIGUE, A. LANOIS, K. POWELL, P. SIGUIER, R. VINCENT, V. WINGATE, M. ZOUINE, P. GLASER, N. BOEMARE, A. DANCHIN, AND F. KUNST. 2003. The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. Nature Biotechnology **21:** 1307–1313.

FISCHER, T., S. RODRIGUEZ-NAVARRO, G. PEREIRA, A. RACZ, E. SCHIEBEL, AND E. HURT. 2004. Yeast centrin Cdc31 is linked to the nuclear mRNA export machinery. Nature Cell Biology **6:** 840–848.

GHEYSEN, G., AND C. FENOLL. 2002. Gene expression in nematode feeding sites. Annual Review of Phytopathology **40:** 191–219.

GREWAL, P. S., R. U. EHLERS, AND D. I. SHAPIRO-ILAN. 2005a. Critical issues and research needs for expanding the use of nematodes in biocontrol. *In* Nematodes as biocontrol agents, P. S. Grewal, R. U. Ehlers, and D. I. Shapiro-Ilan (eds.). CABI Publishing, Oxon, U.K., p. 479–489.

———, ———, AND ———. 2005b. Nematodes as biocontrol agents, CABI Publishing, Oxon, U.K., 480 p.

GRIFFIN, C. T., N. E. BOEMARE, AND E. E. LEWIS. 2005. Biology and behavior. *In* Nematodes as biocontrol agents, P. S. Grewal, R. U. Ehlers, and D. I. Shapiro-Ilan (eds.). CABI Publishing, Oxon, U.K., p. 47–64.

HALL, N., M. KARRAS, J. D. RAINE, J. M. CARLTON, T. W. KOOIJ, M. BERRIMAN, L. FLORENS, C. S. JANSSEN, A. PAIN, G. K. CHRISTOPHIDES, K. JAMES, K. RUTHERFORD, B. HARRIS, D. HARRIS, C. CHURCHER, M. A. QUAIL, D. ORMOND, J. DOGGETT, H. E. TRUEMAN, J. MENDOZA, S. L. BIDWELL, M. A. RAJANDREAM, D. J. CARUCCI, J. R. YATES, 3RD, F. C. KAFATOS, C. J. JANSE, B. BARRELL, C. M. TURNER, A. P. WATERS, AND R. E. SINDEN. 2005. A comprehensive survey of the plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. Science **307:** 82–86.

HENKLE-DUHRSEN, K., AND A. KAMPKOTTER. 2001. Antioxidant enzyme families in parasitic nematodes. Molecular and Biochemical Parasitology **114:** 129–142.

HUANG, X., J. WANG, S. ALURU, S. P. YANG, AND L. HILLIER. 2003. PCAP: A whole-genome assembly program. Genome Research **13:** 2164–2170.

JAROSZ, J. 1998. Active resistance of entomophagous rhabditid *Heterorhabditis bacteriophora* to insect immunity. Parasitology **117:** 201–208.

JASMER, D. P., A. GOVERSE, AND G. SMANT. 2003. Parasitic nematode interactions with mammals and plants. Annual Review of Phytopathology **41:** 245–270.

JONES, J. T., C. FURLANETTO, AND T. KIKUCHI. 2005. Horizontal gene transfer from bacteria and fungi as a driving force in the evolution of plant parasitism in nematodes. Nematology **7:** 641–646.

KIM, S., J. R. KETTLEWELL, R. C. ANDERSON, V. J. BARDWELL, AND D. ZARKOWER. 2003. Sexually dimorphic expression of multiple doublesex-related genes in the embryonic mouse gonad. Gene Expression Patterns **3:** 77–82.

PEARLMAN, E., AND I. GILLETTE-FERGUSON. 2007. *Onchocerca volvulus*, *Wolbachia* and river blindness. Chemical Immunology and Allergy **92:** 254–265.

PESOLE, G., AND S. LIUNI. 1999. Internet resources for the functional analysis of 5′ and 3′ untranslated regions of eukaryotic mRNAs. Trends in Genetics **15:** 378.

POINAR, G. O. J. 1990. Taxonomy and biology of Steinernematidae and Heterorhabditidae. *In* Entomopathogenic nematodes in biological control, R. Gaugler, and H. K. Kaya (eds.). CRC Press, Boca Raton, Florida, p. 23–61.

RAYMOND, C. S., J. R. KETTLEWELL, B. HIRSCH, V. J. BARDWELL, AND D. ZARKOWER. 1999. Expression of Dmrt1 in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. Developmental Biology **215:** 208–220.

———, E. D. PARKER, J. R. KETTLEWELL, L. G. BROWN, D. C. PAGE, K. KUSZ, J. JARUZELSKA, Y. REINBERG, W. L. FLEJTER, V. J. BARDWELL, B. HIRSCH, AND D. ZARKOWER. 1999. A region of human chromosome 9p required for testis development contains two genes related to known sexual regulators. Human Molecular Genetics **8:** 989–996.

SANDHU, S. K., G. B. JAGDALE, S. A. HOGENHOUT, AND P. S. GREWAL. 2006. Comparative analysis of the expressed genome of the infective juvenile entomopathogenic nematode, *Heterorhabditis bacteriophora*. Molecular and Biochemical Parasitology **145:** 239–244.

SCHOLL, E. H., J. L. THORNE, J. P. MCCARTER, AND D. M. BIRD. 2003. Horizontally transferred genes in plant-parasitic nematodes: A high-throughput genomic approach. Genome Biology **4:** R39.

SCHWARZ, E. M., I. ANTOSHECHKIN, C. BASTIANI, T. BIERI, D. BLASIAR, P. CANARAN, J. CHAN, N. CHEN, W. J. CHEN, P. DAVIS, T. J. FIEDLER, L. GIRARD, T. W. HARRIS, E. E. KENNY, R. KISHORE, D. LAWSON, R. LEE, H. M. MÜLLER, C. NAKAMURA, P. OZERSKY, A. PETCHERSKI, A. ROGERS, W. SPOONER, M. A. TULI, K. VAN AUKEN, D. WANG, R. DURBIN, J. SPIETH, L. D. STEIN, AND P. W. STERNBERG. 2006. WormBase: Better software, richer content. Nucleic Acids Research **34:** D475–D478.

SHEARMAN, D. C., AND M. FROMMER. 1998. The *Bactrocera tryoni* homologue of the *Drosophila melanogaster* sex-determination gene doublesex. Insect Molecular Biology **7:** 355–366.

STAJICH, J. E., D. BLOCK, K. BOULEZ, S. E. BRENNER, S. A. CHERVITZ, C. DAGDIGIAN, G. FUELLEN, J. G. GILBERT, I. KORF, H. LAPP, H. LEHVÄSLAIHO, C. MATSALLA, C. J. MUNGALL, B. I. OSBORNE, M. R. POCOCK, P. SCHATTNER, M. SENGER, L. D. STEIN, E. STUPKA, M. D. WILKINSON, AND E. BIRNEY. 2002. The Bioperl toolkit: Perl modules for the life sciences. Genome Research **12:** 1611–1618.

STEIN, L. D., Z. BAO, D. BLASIAR, T. BLUMENTHAL, M. R. BRENT, N. CHEN, A. CHINWALLA, L. CLARKE, C. CLEE, A. COGHLAN, A. COULSON, P. D'EUSTACHIO, D. H. FITCH, L. A. FULTON, R. E. FULTON, S. GRIFFITHS-JONES, T. W. HARRIS, L. W. HILLIER, R. KAMATH, P. E. KUWABARA, E. R. MARDIS, M. A. MARRA, T. L. MINER, P. MINX, J. C. MULLIKIN, R. W. PLUMB, J. ROGERS, J. E. SCHEIN, M. SOHRMANN, J. SPIETH, J. E. STAJICH, C. WEI, D. WILLEY, R. K. WILSON, R. DURBIN, AND R. H. WATERSTON. 2003. The genome sequence of *Caenorhabditis briggsae*: A platform for comparative genomics. PLoS Biology **1:** E45.

THE *C. ELEGANS* SEQUENCING CONSORTIUM. 1998. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. Science **282:** 2012–2018.

ZAMORA, I., AND W. F. MARSHALL. 2005. A mutation in the centriole-associated protein centrin causes genomic instability via increased chromosome loss in *Chlamydomonas reinhardtii*. BMC Biology **3:** 15.