Entomopathogenic Nematode Genome Consortium: Sequencing and Annotation of *Heterorhabditis bacteriophora*

Project Summary: This project aims to sequence the genome of the entomopathogenic nematode *Heterorhabditis bacteriophora* strain TTO1 as a representative of 45 other species that are used as biological control agents for insect pests worldwide. This nematode is closely related to the two already sequenced free-living nematodes, *Caenorhabditis elegans* and *C. briggsae,* and is a sister taxon to several vertebrate parasitic nematodes, but possesses unique attributes including an obligate mutualistic association with the insect pathogenic bacterium, *Photorhabdus luminescens* and ability to parasitize model insects including *Drosophila melanogaster*. Furthermore, the recently sequenced genome of *P. luminescens* provides unique insights into genes associated with pathogenicity and mutualism. Aside from the significance to biological control, the *intellectual merit* of the project lies in the comparison between the genomes of *H. bacteriophora* and *C. elegans* and *C. briggsae* which will enhance our understanding of the evolutionary and biological processes in animals involved in mutualism and parasitism. Although, *H. bacteriophora* is similar in morphology and basic biology to *C. elegans,* its genome is less than half the size of *C. elegans* (~40 Mb vs 100 Mb). Thus, the genome sequence of *H. bacteriophora* will provide a new view of the patterns and processes that shape the genomes in response to mutualism and parasitism, and will identify novel genes and regulatory sequences involved in infection, parasitism, pathogenicity, and mutualism. The project will address questions such as; Are the genes involved in basic biology of *H. bacteriophora* (e.g., dauer formation and sex determination) the same as those of *C. elegans*? Did unique genome features arise *de novo*, via horizontal gene transfer, or by modification of existing genes and pathways? Therefore, the specific objectives of this proposal are to: (i) sequence and annotate the genome of *H. bacteriophora* and disseminate the information through the WormBase database, (ii) compare *H. bacteriophora*, *C. elegans, C. briggsae, Brugia malayi* (a human parasitic nematode), and other in-progress Secernentean genomes for synteny, trans-splicing, regulatory DNA and proteins, and unique genes, and (iii) engage students, scientific community, extension personnel, and general public in the genome project through education, outreach, and workshops.

The *broader impacts* of the project include (i) revolutionization of research in over 100 academic and industrial laboratories developing entomopathogenic nematodes and their symbiotic bacteria as biological control agents worldwide, (ii) facilitation of functional genomic research geared towards the enhancement of infective juvenile longevity, stress tolerance, and virulence that will result in improved efficacy and wider use of nematode in biocontrol, leading to reduced reliance on chemical insecticides, (iii) establishment of *H. bacteriophora* as a tractable model for the study of parasitism, mutualism, and pathogenicity, (iv) provision of avenues for investigating the obligate tripartite interactions among *H. bacteriophora, P. luminescens* and insect larvae, including recognition and signaling, leading to immediate and direct applications in agriculture (e.g., targets for pharmacological intervention and genetic manipulation), (v) training of undergraduate, graduate, and post-graduate students including minorities and women in genome annotation, and (vi) enhanced awareness of the value of genome sequencing among extension personnel, farmers, and public. Chances of our success are greatly enhanced as *H. bacteriophora* shares many attributes that enabled *C. elegans* become an established model system and techniques for establishing inbred lines, genetic selection, hybridization, mutagenesis, and transgene expression have already been established for both *H. bacteriophora* and its symbiotic bacterium. We are also well prepared because our team members have experience with genomic projects including *C. elegans* and *C. briggsae*.

I. Introduction

Nematodes, or roundworms, are a diverse group of organisms that occur worldwide in virtually every environment. In the animal kingdom, nematodes are second only to insects in the number of species described, but are the most abundant animals on earth (Bongers and Ferris, 1999). One cubic foot of soil may contain millions of individual nematodes belonging to several phylogenetically distinct taxonomic groups (Barker *et al*., 1994). While only a few nematode species are parasitic on humans, domesticated animals, and plants, the majority of species are beneficial to agriculture and the environment. One group, the entomopathogenic (or insect-parasitic) nematodes have emerged as excellent biological control agents of important insect pests (Gaugler and Kaya, 1990; Gaugler, 2002). Entomopathogenic nematodes, Heterorhabditidae and Steinernematidae, are symbiotically associated with insect pathogenic bacteria in the family Enterobacteriaceae: heterorhabditids are associated with *Photorhabdus* bacteria and steinernematids with *Xenorhabdus* (Boemare, 2002). The infective juvenile (IJ) or dauer (enduring) juvenile persists in the soil in search of a susceptible insect host. Following entry through the cuticle or natural body openings, the infective juveniles release the symbiotic bacteria into the insect hemocoel which rapidly kill the host, usually within 24-48 h (Poinar, 1990; Ciche and Ensign, 2003). Nematodes feed on symbiotic bacteria, complete 1-3 generations in the host cadaver, and as food resources are depleted new infective juveniles are produced which disperse in search of new hosts. In the laboratory, each partner can be cultured separately, but when combined they present a high degree of specificity.

The symbiontic bacteria occupy two different ecological niches or states in the life cycle, and thus interact with the nematode at two levels (Forst and Clarke, 2002). The first is a phoretic state where the bacteria are retained in, and interact with, the gut mucosa of the non-feeding infective juvenile. *Photorhabdus* bacteria are mainly located in the anterior part of the intestine in *Heterorhabditis* (Boemare *et al*., 1996; Ciche and Ensign, 2003), while *Xenorhabus* are retained in a specialized vesicle (Bird and Akhurst, 1983). The second state is a vegetative one, when the bacteria produce an impressive arsenal of virulence factors ensuring rapid insect mortality (ffrench Constant *et al*., 2003; Dowds and Peters, 2002). Bioconversion of the insect cadaver by exoenzymes produced by the bacteria allows the bacteria to multiply and nematodes to reproduce. During this phase the bacteria also produce antibiotics to prevent invasion of the insect cadaver by competing soil microbes (Webster *et al*., 2002), enabling the nematodes and bacteria to re-associate specifically before leaving the used cadaver.

Progress in research and development during the past three decades has made the nematodes available for large-scale application in citrus groves, strawberry plantations, cranberry bogs, ornamentals (in nurseries and greenhouses), and turf (Grewal and Georgis, 1998). However, to realize their full potential, particularly in large acreage field crops, improvements are needed in the infective juvenile longevity, bacterial retention, tolerance to environmental extremes (e.g. heat, ultra violet radiation, and desiccation), resistance to encapsulation in the hemocoel encountered in some key insect pests (e.g. white grubs and wireworms), trait stability, and persistence and colonization in the field following application. In this regard, recent developments have opened new possibilities for improving the biological control potential of entomopathogenic nematodes and their symbiotic bacteria. First, the genome of *Photorhabdus luminescens* subsp *luminescens*, the bacterial symbiont of *Heterorhabditis bacteriophora* TTO1, has been fully sequenced. The 5.7 Mb bacterial genome contains 4,839 predicted genes many of which encode proteins putatively involved in host-bacterial interactions, including Type I-IV and

TPS secretion systems, adhesion proteins, toxins, hemolysins, proteases, lipases, and antibiotics (Duchaud *et al*., 2003). Comparisons with genomes of related bacteria suggest that *P. luminescens* acquired virulence factors through extensive horizontal transfer and provide clues about the evolution of insect pathogens. This genome sequence information has also established a solid foundation for developing new approaches to improve bacterial virulence to insect pests. Second, the availability of the full genome sequence of two closely related free-living nematodes, *C. elegans* and *C. briggsae*, together with a full complement of genetic and molecular tools (*C. elegans* Genome Sequencing Consortium, 1998; Stein *et al*., 2003) offer new means for improving the biological control potential of *H. bacteriophora*. In this regard, gene orthology between *C. elegans* and *H. bacteriophora* has been confirmed (see Preliminary data) and the expression of *C. elegans* gene sequences in *H. bacteriophora* has been demonstrated (Hashmi *et al*., 1996). However, further increase in our knowledge of the underlying mechanisms controlling important biological processes in *H. bacteriophora* is urgently needed to accelerate the pace of research and development to improve the biological control potential of nematodes. Therefore, we propose to sequence the *H. bacteriophora* genome and functionally describe (annotate) genes on the basis of orthology. Specific objectives of this proposal are:

1. **Sequence and annotate the genome at 3-fold coverage of *H. bacteriophora* and disseminate the information through the *wormbase* database.**
2. **Compare *H. bacteriophora, C. elegans, C. briggsae, Brugia malayi* (a human parasitic nematode), and other in-progress Secernentean genomes for syntheny, trans-splicing, regulatory DNA and proteins, and unique genes.**
3. **Engage students, scientific community, extension personnel, and general public through education, outreach, and workshops.**

The *H. bacteriophora* genome sequence will be used immediately in a number of ways:

1. Gene prediction. The compact size of nematode genomes (≤100 Mb) and the available information on intron-exon structures of the *C. elegans* (e.g, Stein et al., 2003) allows efficient gene prediction. Most initial gene predictions of *C. elegans* turned out to be correct, and we expect that the intron-exon structures of *C. elegans* and *H. bacteriophora* will be similar. Further, co-PI Sternberg is leading the *C. elegans Wormbase* database and is experienced with gene prediction and other annotation tools. Also we have collected sequences from cDNA libraries (see preliminary data) which will help to identify intron-exon structures.
2. Identification of genes potentially involved in infection, parasitism, pathogenicity, and symbiosis through genome comparisons. This set of genes will be selected as a result of comparative analysis of the genome sequences of *C. elegans*, *C. briggsae*, *B. malayi*, and ESTs of plant and animal-parasitic nematodes. Based on preliminary data (see below), we expect that this comparative analysis will result in a database of *Heterorhabditis*-specific genes. We will also use available functional data from *C. elegans* and other nematodes (e.g. expression data; mutant phenotypes) to predict what gene sets are involved in *Heterorhabditis* biology.
3. Facilitation of functional genomics. The expression and function of genes can be tested using standard and high-throughput methods, and equipment for high-throughput experiments, such as a colony picker, are available here on the Ohio State University campus (http://www.oardc.ohio-state.edu/mcic). In addition, Stuart Kim's Stanford-style microarrays

for *C. elegans* were generated from primer pairs based on large predicted ORFs, and oligonucleotide-glass slide microarrays were generated from 70-mer oligos, and Affymetrix type micorarrays from shorter oligos. Importantly, J. Ahringer and colleagues used predicted *C. elegans* genes to generate their RNAi feeding library. Similarly, we will explore the possibility to use this RNAi feeding library or other RNAi technologies for functional analysis of *H. bacteriophora* genes. The co-PIs Ciche, Sternberg, and Grewal, and collaborators, Burnell and Fodor are already working on an approach to combine the *H. bacteriophora* symbiont, *P. luminescens* subsp *luminescens,* and *E. coli* for delivering RNAi constructs into *H. bacteriophora.* Thus, the genome sequence will allow RNAi, the expression of proteins, and generation of microarrays for coding sequences.

4. Transcriptional cis-regulation. The genome sequence will be used to identify candidate regulatory sequences of genes of interest (e.g., parasitism, symbiosis, sex determination, longevity and environmental tolerance). Simple tests of the function of these sequences can be done in *C. elegans* or in *H. bacteriophora*. These experiments will become increasingly popular as the bioinformatics tools for identifying cis-regulatory regions are becoming more sophisticated and user-friendly. These experiments will address when and where the *Heterorhabditis*-specific genes are expressed, as well as alterations in gene expression associated with symbiosis and infection (e.g., in gut-expressed genes). *C. elegans* researchers now routinely examine the function of sequences from orthologous *C. briggsae* genes as part of transcriptional regulatory analysis.
5. Positional cloning. The genome sequence will enable positional cloning in two ways. First, it will provide the ability to obtain single nucleotide polymorphisms (SNPs) for high resolution genetic mapping and correlation of physical and genetic maps. Over 20 divergent *H. bacteriophora* strains exist (Grewal *et al*., 2002b), and these can be used for the identification of inter-strain SNPs. The ability to do SNP mapping is important to order contigs (e.g., in *C. briggsae*, SNP mapping of all contig ends is being used to generate an accurate map (R. Waterston, personal communication). Second, it will provide a physical map. The inclusion of fosmid clones and fosmid-end sequences as are part of the assembly strategy will provide a valuable reagent for transgenic rescue, reporter and site-directed mutagenesis studies. As classical genetic techniques and transgene expression have already been established in *H. bacteriophora*, we believe that a genome sequence of *H. bacteriophora* will further increase the utility of this model in agriculture and enhance our understanding of interesting biological phenomena including mutualism and pathogenicity.

We have assembled an international interdisciplinary team combining expertise in entomopathogenic nematology, *C. elegans* genetics, evolutionary biology, molecular biology, genomics, and bioinformatics. The principal investigators listed on this proposal are supported by a consortium of scientists (see Table 1) working on diverse aspects of entomopathogenic nematodes and their symbiotic bacteria, *C. elegans*, and other nematodes. Sequence and bioinformatics expertise and resources including the Linux database are also available at the Ohio State University's Molecular and Imaging Center in Wooster. The chances of our success are further enhanced by the fact that majority of the sequencing, bioinformatics, and database effort will be carried out by Integrated Genomics Inc., which has successfully lead several genome projects, and has completed or is involved in over a dozen other genome projects of microorganisms and invertebrates and manages an outstanding genome database, ERGO. The data will also be uploaded into WormBase that is public, contains *C. elegans* and *C. briggsae* genome sequence data, and allows comprehensive annotation and genome comparisons.

II. Relevance of the nematode species and scientific merit of the project

Insect pests are a major concern for agriculture and human health. Worldwide insect pests cause 13-16% crop loss which is estimated to be about $244 billion per year (Pimental, 1997). Increasing concerns about the risks of large-scale application of chemical pesticides in environmental degradation and human health have generated strong impetus to develop alternatives, including biological insecticides. The entomopathogenic nematodes, Steinernematidae and Heterorhabditidae, possess many desirable attributes for biological control agents including broad host range, safety to non-target organisms and the environment, exemption from registration in many countries, ease of mass-production and application, ability to search for pests, rapid host mortality, potential to recycle in the environment, and compatibility with many agricultural chemicals (Gaugler and Kaya, 1990; Gaugler, 2002). These nematodes are attractive alternatives to chemical insecticides and are ideally suited for integrated pest management (IPM) and organic systems particularly for soil pests (Grewal and Georgis, 1998). Entomopathogenic nematodes cause rapid mortality (24-48 h) to many agriculturally and medically significant insects while non-target organisms are unaffected (Georgis *et al*., 1991). Furthermore, the application of nematodes is compatible with existing agricultural methods (i.e. spray and irrigation systems).

Efficient *in-vitro* mass-production techniques involving solid substrates (Bedding, 1984) and liquid media (Friedman, 1990; Ehlers, 2002) and formulations with ambient storage stability (Grewal, 2000a, b, 2002) have been developed. Progress and production and application technology has led to large-scale application in citrus groves, strawberry plantations, cranberry bogs, ornamentals (in nurseries and greenhouses), and turf (Grewal and Georgis, 1998). In the USA, over 30,000 acres of citrus are treated annually with the nematodes for the control of root-feeding *Diaprepes* weevil (Clay McCoy, pers. comm). Entomopathogenic nematodes have replaced aldrin as the most effective control measure for the black vine weevil in greenhouse and nursery industries in Europe (Richardson, 1990; Klingler, 1998) and are the only control available for the Hunting Billbug on golfcourses in Japan and against black vine weevil and cranberry girdler in cranberry bogs in North America (Grewal and Georgis, 1998). *Heterorhabditis bacteriophora* and *H. zealandica* are used to control white grubs throughout the world and *H. megidis*, *H. marelata,* and *H. bacteriophora* are used for the control of black vine weevil in Europe and North America. However, to increase the applicability of the nematodes, particularly in large acreage field crops, problems of short infective juvenile longevity, sensitivity to environmental extremes (e.g. heat, ultra violet radiation, and desiccation), low virulence against certain pests due to their encapsulation in the insect hemocoel (e.g. white grubs), rapid trait deterioration during mass culture and handling, and poor persistence and colonization in the field following application need to be addressed (Grewal and Georgis, 1998).

With the recently published sequence of *P. luminescens* subsp. *luminescens* TT01, a consensus from the international community of researchers was reached to focus on *H. bacteriophora* TT01 as a model and to collaborate in obtaining, characterizing, and utilizing the genome sequence (International meetings in Wooster, Ohio and Eliat, Israel, 2003). This nematode is a superb candidate for genome sequencing because it is: (i) an important biological control agent of insect pests of agriculture, horticulture, veterinary, and medical significance (Grewal and Georgis, 1998), (ii) closely related to the two already sequenced free-living nematodes *C. elegans* and *C. briggsae*, and a member of the Rhabditidae which includes some important vertebrate parasitic nematodes (e.g. the human parasitic *Haemonchus similis*) (Blaxter

et al., 1998), (iii) has several unique biological traits, including a mutualistic association with the insect pathogenic bacteria, *Photorhabdus luminescens,* and the production of a developmental variant, (the dauer or enduring juvenile), which serves as an important survival strategy and the basis for the establishment of symbiosis, parasitism, and pathogenicity (Forst and Clark, 2002), (iv) shares many biological attributes that have enabled *C. elegans* to become an established biological model system (i.e., small size, high fecundity, short generation time, simple development, hermaphroditic and amphimictic animals available, can be frozen and revived to maintain original stock) (Poinar, 1990), (v) emerging as a useful model to study parasitism in nematodes because the invasion process and transmission of bacteria is easily observed in insect larvae, including *Drosophila melanogaster,* and in insect cell-culture supernatants (see below), (vi) serves as a model for the study of symbiosis, pathogencity, host-finding, environmental tolerance and host-parasite associations which are characteristic of other 45 entomopathogenic nematode species of economic significance belonging to *Heterorhabditis* and *Steinernema* (see below), (vii) the recently sequenced genome of the bacterial symbiont, *P. luminescens* TTO1 provides unique insights into genes associated with pathogenicity, immunity and mutualism (Duchaud *et al*., 2003), (viii) despite the intimate tripartite interactions, *H. bacteriophora* can be made axenic and grown on symbiotic bacteria on agar media outside the insect host (Forst and Clark, 2002), (ix) inbred lines, genetic selection, hybridization, mutagenesis, and transgene expression has already been established for both the nematode and symbiotic bacterium (see below), and (x) has a small genome size (~40 Mb, Grenier *et al*., 1997).

The genome size of *H. bacteriophora* is estimated to be 40 Mb which is less than half the size of *C. elegans* and *C. briggsae* (Grenier *et al*., 1997). Unlike *C. elegans, H. bacteriophora* has an amphimictic mode of reproduction generating males and females in addition to the automictic mode of reproduction as in *C. elegans* where self-fertilizing hermaphrodites are formed. This allows both clonal screens and crosses to be utilized in *H. bacteriophora.* Many genetic techniques of *C. elegans* are proven in *H. bacteriophora*. Forward genetics by mutagenesis using ethyl methane sulfonate (EMS) mutagenesis was successfully applied to obtain dumpy mutants (Zioni et al., 1992) and a desiccation tolerant mutant (O'Leary and Burnell, 1997). Moreover, techniques for genetic selection (Gaugler *et al*., 1989a; Grewal *et al*., 1996a, b; 2002; Glazer *et al*., 1991; Segal and Glazer, 2000), hybridization (Shapiro *et al*., 1997), and DNA transformation (Hashmi *et al*., 1995; Vellai *et al*., 1999) have already been developed and proven. Transformation of the *H. bacteriophora* germline with the *C. elegans* heat shock promoter transcriptionally fused to β·galactosidase (Hashmi *et al*., 1995) and *mec-4* (mechanosensitive) promoter transcriptionally fused to GFP (Hashmi *et al*., 1997) suggest that functional analysis of *H. bacteriophora* genes is possible. Data from pilot expressed-sequenced tags (ESTs) and cDNA library sequencing projects of *H. bacteriophora* are now available (see Preliminary data). Reverse genetics by gene silencing using RNAi is currently being developed in the Burnell, Ciche and Sternberg, and Grewal and Fodor laboratories, it should be possible to transform the *sid-2* gene involved in dsRNA transport (Feinberg and Hunter, 2003) into *H. bacteriophora* or to use morpholinos for gene silencing.

III. Broader impact on biological sciences, agriculture and society

Entomopathogenic nematode and bacteria (EPN-EPB) community is comprised of over 600 researchers in more than 100 laboratories worldwide. There have been three international meetings (1998, 1994, and 2003) and many regional meetings in Europe, North and South America, Australia, and Asia. Over 100 researchers from 14 countries attended the 2003

International meeting in Wooster, Ohio (www.oardc.ohio-state.edu/nematodes) in which it was decided to hold the international meeting every two years. Over 1000 papers were published on entomopathogenic nematodes and their symbiotic bacteria during the past 10 years. Several members of the EPN-EPB community have experience with *C. elegans*. With a large contingent of academic, government and industrial labs researching entomopathogenic nematodes both nationally and internationally, the genome sequence of *H. bacteriophora* will undoubtedly have a large impact on society and agriculture. Many students are now trained in entomopathogenic nematode and symbiotic bacterial biology. In addition, workshops are routinely held domestically and abroad to train scientists and agricultural workers on the effective use and production of entomopathogenic nematodes. The genome sequence of *H. bacteriophora* will increase the already broad interest in entomopathogenic nematodes. This is a unique biological system where nematodes can be seen to regurgitate GFP-labeled symbionts in insect blood and insect cadavers which are bioluminescent and red pigmented where 100,000s IJs emerge after 10 days. Most excitingly the genome of *H. bacteriophora* will bridge the gap between the *C. elegans* and biological community and the symbiosis and parasitism research communities. Apart form its significance to biological control and agriculture below we illustrate how *H. bacteriophora* genome sequence will have a huge impact on biological sciences and the society.

1. *H. bacteriophora* serves as a tractable model for the study of symbiosis

Common molecular mechanisms exist as to how bacteria interact with animal or plant cells (Rhame *et al*., 1995, Plotnikova *et al*., 2000) or whether the interaction is mutualistic or pathogenic (Steinert *et al*., 2000). The use of model organisms like *C. elegans* (Abalay and Ausubel, 2002) or *Arabidopsis thaliana* (Plotnikova *et al*., 2000) have proven useful in elucidating the molecular mechanisms of host-pathogen interactions, especially how bacteria interact with the innate immune system of plants and animals. Much of the similarities between bacterial-eukaryotic interactions are due to various components of the Toll signaling pathway present in *D. melanogaster* (reviewed in Hoffman and Reichart, 2002; Hoffman, 2003) and *C. elegans* (Pujol *et al*., 2001). Much less is understood about benign, beneficial or mutualistic interactions, because few symbiotic interactions exist where all partners are of small size, have short generation time, are easily cultured, and have powerful genetic tools like those available for *H. bacteriophora* and *P. luminescens*.

Phylogenetic relationship among species of *Heterorhabditis* and *Photorhabdus* (ffrench-Constant, 2003, in press; Adams and Nguyen, 2002; Fischer-Le Saux *et al*., 1999) are congruent with that of coevolution. Comparative statistical analyses of co-evolution are currently underway (B. Adams, Personal Communication). Typically, each nematode species is associated with a species or subspecies of bacteria (Boemare, 2002). Only rarely are the nematodes associated with symbionts incongruent with their phylogenetic position. Thus, the highly specific nematode-bacterial association appears to be a legacy of tight coevolution, with only a few cases of host-symbiont infidelity. This symbiotic relationship is a mutualism because the nematode 1) requires viable symbiotic bacteria to grow and develop (both in insect and in culture) and 2) retains and transmits only the symbiotic bacteria that are required to kill an insect host. At present, data suggest that both nutrition and signaling are required for nematode growth and reproduction. Two intracellular crystalline inclusion proteins (CipA and CipB) that can account for over 40% of total cell protein and high in essential amino acids are necessary but not sufficient for nematode growth and reproduction (Bintrim and Ensign, 1998). A gene required for the bacteria to support **n**ematode **g**rowth and **r**eproduction was named, *ngrA*, and is

homologous to the phosphopantetheinyl transferases (PPTases) which are required for the assembly of acyl or pepidyl groups to fatty acids, and a large variety of secondary metabolites (i.e. polyketides, non-ribosomal peptides, siderophores) (Ciche *et al*., 2001). However, a catechol siderphore, photobactin, produced by *P. temperata,* is not required for nematode growth and reproduction (Ciche *et al*., 2003).

The infective juveniles selectively retain *P. luminescens* in their gut mucosa (Ciche and Ensign, 2003) in contrast to the other nematode stages that digest *P. luminescens.* The molecular mechanisms that confer specificity in the colonization of the infective juvenile intestine are not known, but it is clear that *P. luminescens* interacts differently with the gut epithelium of the infective juvenile versus the other life stages of the nematode and this interaction is highly selective for the species or subspecies that can successfully colonize and persist in the IJ intestine (Gerritson and Smits, 1993). Thus, the nematode and bacteria interact in a host specific manner.

2. *H. bacteriophora* is a unique model for the study of mutualism and pathogenesis

The pathogenic process depends on characteristics of each of the three partners of the interaction: the insect, nematode and bacteria. Pathogenicity is influenced by insect resistance (including humoral and cellular defenses) and by virulence factors of the bacteria and the nematodes (Dowds and Peters, 2002). Most *Photorhabdus* strains examined to date are highly virulent, the LD_{50} usually being < 100 cells for the Waxmoth, *Galleria mellonella* larvae (Akhurst and Boemare, 1990). After being immersed in insect hemolymph for 30 min, the infective juveniles regurgitate the *P. luminescens* in a staggered and gradual manner for more than 300 min (Ciche and Ensign, 2003, see time-lapsed movie at http://www.stanford.edu/~taciche). Regurgitation is induced by non-proteinaceous and small molecular weight component(s) present in hemolymph or cell cultures of a variety of insects (Ciche and Ensign, 2003). This provides a convenient assay for studying successful pathogen transmission and parasitism induction. After being released, the bacteria are phagocytosed by hemocytes (Dunphy and Thurston, 1990; Bowen *et al*., 2003) where they survive and presumably produce toxins and other virulence factors that kill the insect host, usually within 6 hours (Au *et al*., 2003).

It has been discovered that *Photorhabdus* (Bowen and Ensign, 1998; Bowen *et al*., 1998; Guo *et al*., 1999) and *Xenorhabdus* (Ensign *et al*., 2000, Morgan *et al*., 2001) produce orally active insecticidal protein toxins. The toxins exist in at least 4 toxin complexes of more than 100 kDa weight (Bowen and Ensign., 1998) consisting of several distinct peptides that are expressed from 4 operons (Toxin complex, *tca, tcb, tcc, tcd*) (Bowen *et al*., 1998). Genomic data from other pathogenic bacteria have revealed *tc* homologs in *Serratia entomophila, Yersinia pestis* CO29**,** *Pseudomonas syringe* pv. *tomato, Fibriobacter succinogenes* S85 (rumen associated) and *Treponema denticola* II:11:33520 involved in periodontal disease (ffrench-Constant, 2003). Hence, the *tc* genes might have broad significance that extends to other animal-bacterial interactions. In fact, on 238 kDa, TcdA toxin was successfully expressed in plants leading to plant protection against insects (Lui *et al*., 2003).

The complete genome sequence of *P. luminescens* subsp. *luminescens* TTO1 (Duchaud *et al*., 2003) and draft sequence of *P. luminescens* subsp. *akhustii* W14 have revealed an abundance of other genes putatively involved in host-bacterial interactions, including several putative pathogenicity islands (PAIs) containing Type I-IV secretion pathway homologs (ffrench-Constant *et al*., 2000, 2003). In addition to *tc* PAIs, other PAIs in *P. luminescens* subsp. *akhurstii* W14 are: PAI-II *mcf*-island which encodes *mcf* [*makes caterpillars floppy* (Daborn *et al*., 2002)] toxin; PAI-III, contains a cytonecrosis-like toxin, and PAI-IV contains genes similar

to *E. coli* O157:H7 PAI and the *phlAB* hemolysin. *Photorhabdus luminescens* contains a Ymt homolog that is essential for transmission of *Y. pestis* in the flea (Hinnebush *et al*., 2002). The function of Ymt in *P. luminescens* is unknown, but hypothesized to function in the insect host and be ancestral to the *Y. pestis* Ymt (Duchaud *et al*., 2003). It is also plausible that Ymt is involved in transmission by the nematode host. Several other virulence factors participate in the pathogenicity of *Photorhabdus* and *Xenorhabdus* (Dowds and Peters, 2002; Forst and Clarke, 2002), including motility (Givaudan *et al*., 1995, 1996; Givaudan and Lanois, 2000) and hemolysins (Brillard *et al*., 2001, 2002, 2003). With this initial data it is evident that *Photorhabdus* is similar to other enteric pathogens, like *Y. pestis* and *E. coli* in containing several PAIs. The function of many ORFs present in the PAIs remain to be determined. Therefore, the genome sequence of *H. bacteriophora* will enable detailed studies to elucidate the molecular mechanisms important to the evolution of symbiosis and pathogenesis.

3. *H. bacteriophora* is a tractable model for the study of parasitism in nematodes

Heterorhabditis bacteriophora is an ideal model for the study of nematode parasitism within an evolutionary framework. Emerging data suggest an association between invertebrate and vertebrate parasitism, with invertebrate-pathogenic and –parasitic clades lying basal to major vertebrate-parasitic ones. Although, the phylogenetic relationships among the Rhabditidae have been shown to be paraphylectic (Blaxter *et al*., 1998), *H. bacteriophora* is a member of a monophylectic clade that contains several vertebrate parasitic nematodes, including the human parasite *Haemonchus similis* and is the sister taxon to the clade containing the *Caenorhabditis* group (David Fitch, preliminary results, http://www.nyu.edu/projects/fitch/Prelimtree.gif).

Studies on the host finding behavior of entomopathogenic nematodes have been more extensive than any other group of nematodes (see Lewis, 2002 for review). Host finding behavior of entomopathogenic nematodes is a continuum with ambush and cruise foraging modes being the two extremes. *H. bacteriophora* is a cruise forager as infective juveniles actively search for insect hosts (Lewis *et al*. 1992; Grewal *et al*. 1994a; Campbell and Gaugler, 1997). Host recognition behavior has also been measured by recording changes in several behaviorsof infective juveniles in response to host-related materials (Grewal *et al*., 1993a, b; 1997). Chemotaxis studies have revealed that *H. bacteriophora* ifective juveniles are unresponsive to a large number of compounds which *C. elegans* finds highly attractive (O'Halloran and Burnell, 2003). These results suggest that, associated with the adoption of a parasitic mode of life by *Heterorhabditis,* there was an adaptive change in chemotactic behavior of the infective stage resulting in a decreased sensitivity to volatile by-products of bacterial metabolism and an increased sensitivity towards long chain alcohols and other insect specific volatiles and possibly also to herbivore induced plant volatiles. Following entry into the insect hemocoel, the infective juvenile faces a strenuous immune response mounted by the host and the first few that enter are encapsulated, melanized, and killed (Wang *et al*., 1995; Grewal *et al*., 2002b). Some entomopathogenic nematode species produce anti cecropin proteins as a defense against the host immune response (Wang and Gaugler, 1999).

Like *H. bacteriophora,* most parasitic nematodes infect from the dauer stage (infective juvenile), often after sensing host cues. Recently, it was demonstrated that dauer recovery occurs through a common muscarinic pathway in *Haemonchus contortus* and *C. elegans* (Tissenbaum *et al*., 2000) and preliminary results suggest that this pathway is also utilized in *H. bacteriophora* (Ciche and Sternberg, unpublished results). Thus, the complex process of parasitism can be studied in the more tractable *H. bacteriophora-P. luminescens* system by

monitoring bacterial regurgitation in insect cell-culture supernatants or hemolymph, obviating many of the obstacles associated with similar approaches in other parasite systems. *H. bacteriophora* shares numerous biological similarities with *C. elegans* and proven tools of mutagenesis and transgene expression can be powerful avenues of exploration of the origin and maintenance of genetic and molecular pathways involved in parasitism. Recently, *P. luminescens* was labeled with the green-fluorescent protein allowing the bacteria to be observed in living nematodes while incubated in insect hemolymph (Ciche and Ensign, 2003). A heat and protease resistant, <3.5 kDa component present in insect hemolymph and cell-culture supernatants induced the regurgitation process. This provides a convenient assay to monitor the induction of parasitism in nematodes and sensing and signal transduction of the arthropod cues.

4. *Heterorhabditis bacteriophora* genome sequence will serve as a bridge between *C. elegans* and more distantly related nematode parasites

The *C. elegans* project based on a complete phenotypic description of the organism at cellular, molecular and developmental level and a well defined genetics, including saturation of the genome, epistasis of gene interactions, discovery of suppressors, linkage mapping, molecular identification on genes by using transposon mutagenesis, chromosome walking, cosmid cloning, YAC (BAC) cloning, identification of contigs and contins, including new vectors, EST libraries, transformation, ectopic gene expression, and mosaic techniques provides an excellent data set in which to compare the *H. bacteriophora* genome (www.wormbase.org). As for the phenotypes and related genes, the dauer larva formation and recovery, sex determination, environmental tolerance and the innate immune response are probably the relevant items to the *H. bacteriophora* project, since the physiological basis of parasitism, symbiosis, and environmental tolerance are likely to involve the dauer pathway which is related to sex determination.

With the complete genome of *H. bacteriophora* we can begin to answer two important questions using a genome-wide analysis: What genes make a nematode a parasite, and what genes make a nematode a host in a bacterial-animal mutualism? A genome wide comparison will reveal common and unique genes in *C. elegans* and *H. bacteriophora* and should also reveal the extent of synteny and conservation of trans-splicing, regulatory DNA and RNA, metabolic and stress response genes between a parasite and symbiont host and a bacterial grazer. Although few if any genes in *C. elegans* were acquired by horizontal transfer, some genes in *Meloidogyne spp.* were probably acquired from a rhizobial ancestor (Scholl *et al*., 2003). By being in an obligate mutualism one would predict that certain redundant genes would be lost in the nematode. Similarly, genes from the symbiont or other bacteria may have been transmitted to the nematode and function in insect parasitism or the interaction with symbiotic bacteria. Thus, the comparison of the two genomes should shed light on the evolution of symparasitism.

Nematodes outside the *Caenorhabditis* lineage might contain genes of biological relevance that have been lost in *Caenorhabditis*, possibly due to its short generation time. The recently identified breast and ovarian cancer protein EMSY, whose amplification is correlated with increased cancer risk (Hughes-Davies *et al*., 2003), is present in EST sequence from *Meloidogyne javanica* and *Strongyloides ratti*, but absent from the *C. elegans* genome (Erich Schwarz, personal communication). It remains to be seen to what extent this and other biologically important genes have been lost in the *Caenorhabditis* lineage, but present in other nematodes. Regardless, this illustrates the utility of having a complete genome sequence of a genetically tractable nematode closely related to, yet outside the *Caenorhabditis* lineage.

5. The *H. bacteriophora* genome sequence will enhance utility of the *P. luminescens* TTO1 genome sequence to elucidate host-bacterial interactions

The completed genome sequence of *P. luminescens* contains a plethora of genes putatively involved in eukaryotic-bacterial interactions. This information will serve as a powerful tool for transcriptional analysis such as serial analysis of gene expression (SAGE). This will allow a more efficient analysis of changes in gene expression than subtraction of EST libraries and microarray analysis. A recent technique with several advantages to SAGE was recently developed to identify 5' sequence tags of sufficient length to uniquely identify the tags in a genomic sequence (Huang, Muller and Sternberg, submitted). This technique is called TEC-RED, trans-spliced exon coupled-RNA end determination and depends on SL1 and SL2 trans-splicing. SL1 is present in all nematodes (Evans et al., 1997) and SL2 likely occurs in *H. bacteriophora* due to the phylogenetic proximity to *C. elegans* (T. Blumenthal, personal communication). This technique is more efficient than EST analysis, more sensitive than microarray experiments and because the 5' end of transcripts are identified, the technique greatly aids genome annotation. One limitation is that not all transcripts are trans-splice, 70% of mRNA in *C. elegans* is trans-spliced using either SL1 or SL2. These techniques, along with functional analysis and reverse genetics will allow detailed analysis of nematode genes involved in parasitism and mutualism, while simultaneously occurring in the bacteria.

IV. Preliminary data

Homology between *C. elegans* and *C. briggsae* and *H. bacteriophora* chemoreceptor genes

Significant progress has been made in unravelling the steps involved in olfactory signalling in *C. elegans*. Many of these steps have been identified when genes responsible for various chemotaxis deficient mutants were cloned and characterized. For example, *odr-10* mutants of *C. elegans* are unable to chemotax to diacetyl. When the *odr-10* gene was cloned it was found to be a divergent G-protein coupled receptor with a weak homology to vertebrate olfactory receptors. The downstream effectors of the *C. elegans* chemoreceptors are heterotrimeric G proteins, comprised of α, β and γ subunits, each subunit encoded by a different gene. There are 21Gα, 2Gβ and 2Gγ genes in *C. elegans* (Jansen *et al*., 1999). We designed degenerate primers to conserved motifs identified from multiple alignments of Gα amino acid sequences from a variety of organisms. Using these degenerate PCR primers we have successfully cloned three candidate G protein α subunit gene fragments from *H. bacteriophora*. To clone more divergent members of the G protein α family from *H. bacteriophora* we aligned homologous *C. elegans*/*C. briggsae* G protein α gene pairs and gene-specific degenerate primers were designed. In this way 5 further G protein α subunit gene fragments were cloned from *H. bacteriophora* (Fig. 1). Using a similar approach we have recently cloned a homologue of *odr-10* gene from *H. bacteriophora*.

The availability of the completed *C. elegans* and *C. briggsae* genomes has provided a very valuable opportunity to study the comparative evolution of chemoreceptor genes and chemosensory transduction pathways in nematodes. Since chemoreception and olfaction are critical components of the infection process for parasitic nematodes, chemorecptor genes and signal transduction pathways have the potential to be important antihelminthic targets. Because olfactory repertoire of parasitic nematodes is likely to be more targeted and more specific than in free living nematodes, it is possible that the olfactory receptors will comprise smaller gene families in parasitic nematodes. While it is possible to use a comparative genomics approach to isolate individual chemoreceptor genes in parasitic nematodes, a complete understanding of chemoreceptor genes in *H. bacteriophora* will require access to the genomic DNA sequences.

2. cDNA EST sequences from recovering dauer juveniles of *H. bacteriophora*

Some of the toxins produced by the symbiont bacterium *Photorhabdus* have been cloned, patented and transferred into transgenic plants. It is known that *H. bacteriophora* also contributes to the pathogenenesis, but the genes that are involved have not been investigated. We have found that during the early stages of insect infection there is intensive synthesis of mRNA in the pharyngeal glands of *H. bacteriophora* (Dolan *et al.,* 2002). We have carried out an EST (expressed sequence tag) analysis of the genes which are transcribed early in the infection process in *H. bacteriophora.* Of 861 distinct gene sequences which have been identified as being expressed by the nematodes during the early stages of infection, our data set 154 novel sequences not previously recorded in genetic databases. These data are important because they support our hypothesis that the *H. bacteriophora* genomes harbor several genes important for infection of the insect host. These genes may be absent from free-living nematodes such as *C. elegans* and hence could be unique for *H. bacteriophora* or other parasitic nematodes. We are currently cataloguing these sequences and hope to be able to submit them to the EST database early in 2004.

3. Genetic Linkage map of expressed genes in *H. bacteriophora.*

We have begun to use our EST data set to construct a genetic linkage map of expressed genes in *H. bacteriophora.* We are currently identifying introns in *H. bacteriophora* housekeeping genes from our EST set and using single strand conformation polymorphism (SSCP) gels to identify polymorphisms across the twelve *H. bacteriophora* strains that we are using as putative mapping strains. We have decided that *H. bacteriophora* strain TT01 will be one of the mapping strains.

V. Methods

We have already established and also distributed *H. bacteriophora* TTO1 inbred lines to researchers. The wild type strain has been cryopreserved in Drs. Burnell (University of Ireland, Maynooth, Ireland), Sternberg (Cal Tech, Pasadena, CA) and Grewal (Ohio State University, Wooster, OH) laboratories. For isolation of DNA, the nematodes will be reared in last instar *G. melloenella* larvae following the methods described by Kaya and Stock (1997).

Library Construction. For genome sequencing, data will be generated from a combination of libraries constructed in different vectors to confer a range of insert sizes with paired-end links in the genome sequence assembly: (i) 3-4 kb Insert Length High-Copy Plasmid Libraries, (ii) 8-10 kb Insert Length Low-Copy Plasmid Libraries, and (iii) 35-40 kb Insert Length Fosmid Libraries. The various insert length libraries are constructed in different vector systems tailored to each library type. The 3-5 kb insert length libraries and 8-10 kb insert libraries are constructed in the high-copy and low-copy pSMART vector systems (Lucigen, Middleton, WI), respectively. This vector system was developed specifically for the creation of non-biased shotgun libraries. Libraries are constructed in these vectors using an adaptor based cloning method. The cloning vector is cleaved at a symmetrical pair of BstX1 sites to produce non-complementary 4-base overhangs. Insert DNA is mechanically sheared via a Hydroshear device (Genemachines, San Carlos, CA), end-repaired and ligated to an adaptor with an overhang complementary to the vector ends, but not to itself. The processed inserts are then ligated into the vector. The use of non-self-complementary adaptors lowers the incidence of chimeric clones.

Sequencing strategy and methods. Agencourt has extensive experience in developing and implementing large-scale automated template purification systems using solid-phase reversible

immobilization. The SPRI technology uses carboxylate-coated, iron-core, paramagnetic particles to capture DNA of a desired fragment length based on tuned buffering conditions. Once the desired DNA is captured on the particles, they are magnetically concentrated and separated so that contaminants are washed away. This technology is highly automated since it eliminates the traditional hard-to-automate centrifugation steps associated with most DNA purification protocols. Agencourt Bioscience has invested significant effort to redesigning the chemistry to improve efficiency and developing a streamlined SPRI protocol termed "SprintPrep". This procedure harvests plasmid DNA directly from lysed bacterial cultures by trapping both plasmid and genomic DNA to the functionalized beads and selectively eluting only plasmid DNA.

DNA templates are sequenced in 384-well format, with the forward and reverse reactions (paired ends) done in the same plate to maximize the paired end rate. High copy plasmid shotgun clones are sequenced using BigDye Version 3.1 reactions on ABI3700 instruments. Thermal cycling is performed with approximately sixty 384-well Thermal cyclers (ABI, MJ Research). Sequencing Reactions are purified using Agencourt's CleanSeq dye-terminator removal kit. This process has been shown to significantly enhance the performance of capillary sequencing separations, both at Agencourt and at other large genome centers (E. Mardis, personal communication). Agencourt uses two 384-tip robots dedicated to 384-well CleanSeq processing, each robot is capable of purifying more than ten 384 plates per hour. For Fosmid sequencing, thermal cycling is extended to 100 cycles. With these cycling parameters evaporation becomes a major issue. While most laboratories use large amounts of BigDye terminator reagent to overcome these limitations, Agencourt's production method uses oil immersion to accomplish the same end using 2.5-5 microliter reaction volumes. Reaction miniaturization offers advantages over dilution of BigDye, by maintaining the appropriate concentrations of fluorescent ddNTPs and TaqFS enzyme through all the reaction cycles.

Agencourt's facility currently houses 43 ABI3700 DNA sequencers and 7 ABI3730xl sequencers. Software has been developed to allow all 104 capillaries to be utilized on the ABI3700, providing an 8% increase in machine efficiency over 96 capillary runs. Conditions have been optimized to complete 10 runs per day on ABI3700 instruments and 24 runs per day on ABI3730xl instruments using standard 50 cm capillaries. Thus, current capacity is approximately 20 million reads per year, assuming 90% machine up-time. The combined total number of sequence reads produced over the past 24 months by the Agencourt Team, encompassing a large number of projects, both public and private, was approximately 18 million reads. The majority of these reads, about 70%, were produced for publicly funded projects. The average pass rate across all projects at Agencourt during the past year was 86%, with an average number of Phred Q20 bases of 559 and clipped read length of 688.

All reads and associated data (library information, sample plates, project history, sequence traces, etc.) are tracked though our Oracle 9i driven Laboratory Information Management System (LIMS). Sample data are tracked in the laboratory by means of color-coded barcodes on microtiter plates, and the appropriate sample-processing data are automatically downloaded from the lab robots and stored in the LIMS. Sample sheets for the sequencing instruments are automatically generated from the barcode-associated data, and the resulting sequence traces are automatically uploaded and processed to identify vector sequence and other contaminants (eg., *E. coli*, phage sequences, etc). The data are then subjected to quality clipping, processed using the Phred base calling software, and the resulting files are stored in the LIMS. At regular intervals chromatographic data will be deposited with the NCBI Trace Data archive to ensure public access to the basic data.

During the raw data generation process, a QC assembly is done initially on the first 1000 reads and BLAST analysis is performed to determine the quality and randomness of the library. Once the QC assembly and analysis is deemed successful, raw sequence data is generated for 8-10x coverage. At this point the sequence data is assembled to produce contiguous sequence elements (*contigs*) that can be arranged within larger units (*supercontigs* or *scaffolds*) wherein the gaps between the contigs are bridged by paired reads (or *mates*) from clone insert ends whose average spacing is known experimentally. For routine assembly of BAC and microbial genome sequences we use Phrap (http://www.phrap.org/), the Paracel GenomeAssembler (http://www.paracel.com/products/pga.html), and Arachne (Whitehead Institute for Biomedical Research). The sequence data will be loaded into ERGO for annotation. Integrated Genomics' ERGO database includes one of the largest, if not the largest collection of non-redundant protein sequences in the world comprising over 2.2 million sequences from public and private genomic sequence repositories throughout the world. Our experts continuously manually curate the collection to improve the quality of our bioinformatics products. ORF calling and automatic annotations use various combinations of public and proprietary tools. A comparative approach is utilized at each step of genome analysis to ensure that the annotations are complete.

Project Management

Summer workshop for the consortium members. A 1.5 day Summer Workshop will be held in Wooster for all the consortium participants at the end of the first year of the project. All participants will receive hands-on training on annotation and the use of ERGO database in addition to the update on the sequencing effort. Educational, outreach, and extension strategies for implementation will be discussed and modules developed for implementation in their own research, teaching, and extension programs at their home institutions and states or countries.

Education. There will be two post-doctoral researchers on this project, one at the Ohio State University, Wooster, OH and the other (Todd Ciche) at the California Technology Institute, Pasadena, CA. Both Post-doc's will be fully involved in all aspects of the project including DNA isolation construction of genomic libraries, annotation, workshops, supervision of summer students, and outreach at local high schools. The PIs will incorporate genomic, bioinformatics and annotation modules in their graduate courses. As highlighted by studies such as "Bio2010: Transforming Undergraduate Education for Future Research Biologists" an enormous disconnect exists between the current state of biological research and the content and teaching methods used in our undergraduate biology curricula (National Research Council, 2003). Hence, there is an urgent need to strengthen biology education through more interdisciplinary, hands-on learning, and a focus on experimental thinking and problem solving. Currently, undergraduates at most universities complete exercises in their lab courses that are similar to or complement the same activities that are involved in "genome finishing" (characterization, assembly & annotation), but with "canned" *in silico* and wet lab exercises and data. Direct involvement of these students with "real" data provides an opportunity for them to make meaningful contributions to science and take part in the publication process. Brigham Young University has an established undergraduate degree program in bioinformatics (one of only a handful in the US), and has long-standing molecular and cell biology labs that are geared to provide just such "cutting-edge" opportunities. Proof of concept has been aptly demonstrated in a similar program at Hiram College, where completion of the *Agrobacterium tumefaciens* C58 genome project involved over 200 undergraduates, both within courses and through independent research (Goodner *et al*., 2001; Wood *et al*., 2001). Similar projects utilizing undergraduate involvement in sequencing

and annotating the genomes of two strains of the entomopathogenic bacteria *Xenorhabdus* are currently underway (Goodner and Forst, pers. Comm.). The successes of each of these projects provide a ready model for our proposed efforts. Accordingly, undergraduates at Brigham Young University, and eventually undergraduate serving institutions of several other collaborators involved in the project (College of Wooster, Hiram, Marian, and Cal Tech), will aid in the completion of the project at three important junctures. Each opportunity provides novel and meaningful learning experience, and produces a valuable contribution to project completion.

As part of the bioinformatics training described above, undergraduates will help with the first-pass annotation of each complete genome sequence. Similar to exercises in the current curricula, students will be assigned a large segment of genomic sequence to fully annotate, subject to checking by the project directors. Such efforts are entirely within the scope of well-trained students, as evidenced by recent efforts at Hiram College where individual students annotated 50-70 kbp segments of another genome during an intensive 3-week course on bioinformatics (Brad Goodner, personal communication). In other cases, teams of students will be assigned whole-genome analyses to complement the standard first-pass annotation. For example, annotation of protein-coding regions that take into account predicted protein size, pI, predicted signal sequence cleavage sites, and predicted protease digestions patterns are available for very few completed genomes. However, it is exactly this type of information that will allow genomics to interface with future proteomics efforts. Such an effort is within the reach of undergraduate teams that include both biology, computer science, and bioinformatics majors as evidenced by a database of proteomics annotation constructed at Hiram College for *A. tumefaciens* C58 (accessible at http://hgi.hiram.edu).

Outreach. Outreach will consist of two components aimed at different audiences. Additional workshops will be held one each at the annual meetings of the Society of Nematologists, Society of Invertebrate Pathology, and the EPN-EPB international meeting in 2005 to engage the scientific community at large in the genome annotation. At the undergraduate level, model syllabi, laboratory exercises, and ideas for independent research incorporating genomics and bioinformatics in current use at Brigham Young University will be modified (after the Hiram College model) and disseminated to other collaborators in the *Heterorhabditis* genome project with undergraduate cell/molecular biology and bioinformatics. This effort will be coordinated through the workshop portions of the proposed project. At the high school level, model lesson plans and laboratory exercises incorporating bioinformatics will be developed through a partnership with the BYU science education program and disseminated to local high school biology teachers in the areas surrounding each project director as well as through presentations at biology education conferences (e.g., NABT). Outreach to high schools will build on the current success of the BYU Science Education Outreach program, which has fostered learning collaborations between faculty and students at BYU and local high school biology classes.

Extension. EPN community excels in delivering the extension related information. Four years ago the EPN community lead by one of the PI's (Gaugler) created a tool box of extension information on entomopathogenic nematodes which includes a Video, website, poster, and a slide set. The tools have been used extensively to deliver information on various aspects of entomopathogenic nematodes worldwide. We will create a new module on the EPN-EPB website (www.oardc.ohio-state.edu/nematodes) to disseminate easy to use information related to Hb genome sequence to the public at large. In addition we will conduct a national Extension-in-service workshop in which extension agents, crop consultants, and industry personnel will be given an update on the genome sequencing projects in addition to information on the biology and

application of nematodes. All participants will receive an easy to understand slide set for their use in training of the farmers. Many of the members of the Entomopathogenic Nematode Genome Consortium have Research and Extension/Teaching appointments so will be able to rapidly include the nematode genomic information in their current programs.

Participants of the Entomopathogenic Nematode Genome Consortium

Name	**Institution and E-mail**	**Major research interest**	**Role in the proposed project**
P. S. Grewal	Ohio State University, USA; grewal.4@osu.edu	Infective juvenile longevity, virulence, and stress tolerance	Overall coordination, annotation, education, outreach, and extension
T. Ciche	Cal Tech, USA; taciche@its.caltech.edu	Symbiosis	Annotation, education, and outreach
P. Sternberg,	Cal Tech, USA; pws@caltech.edu	*C. elegans* developmental biology	Annotation, outreach to *C. elegans* community and a genomics advisor
R. Gaugler	Rutgers University, USA; gaugler@rci.rutgers.edu	Host-finding and immune response	Education and outreach
B. Adams	Brigham Young University, USA; bjadams@byu.edu	Evolution of parasitism, molecular evolution and coevolution	Annotation, undergraduate education and research opportunities
S. Hogenhout	Ohio State University, USA; http://www.oardc.ohio-state.edu/phtoplasma	Immune response	Annotation and genomics advisor to the project
Ann Burnell	NUI Maynooth, Ireland; ann.burnell@may.ie	Chemoreception, dauer recovery, and genetic linkage mapping	Annotation, education, and outreach
A. Fodor	Ohio State University, USA; fodorandras@yahoo.com	Sex determination	Annotation
Hinanit Koltai	Volcani Center, Israel; hkoltai@agri.gov.il	Osmotic tolerance	Outreach, and genomics advisor
Itamar Glazer	Volcani Center, Israel	Desiccation tolerance	Outreach and extension
M. Blaxter	University of Edinburgh, UK; mark.blaxter@ed.ac.uk	Evolution of genomes, Brugia genome project	Genomics advisor
David H. Fitch	New York University; david.fitch@nyu.edu	Evolutionary developmental biology	Annotation, education, and outreach
Brad Goodner	Hiram College, USA; goodnerbw@hiram.edu	Microbial and nematode genomics	Annotation and undergraduate education
William Morgan	College of Wooster; wmorgan@wooster.edu	*C. elegans* metabolomics	Annotation and undergraduate education
Dean Fraga	College of Wooster, USA; dfraga@wooster.edu	Paramecium physiology	Annotation and undergraduate education
F. Kunst	Institut Pasteur, France; fkunst@pasteur.fr	Evolution of genomes	Outreach, and genomics advisor
N. Boemare	INRA-UMII, France; boemare@ensam.inra.fr	Coevolution	Outreach
Richard ffrench-	University of Bath, UK; bssrfc@bath.ac.uk	Pathogenicity	Education and outreach

Constant			
David Clark	University of Bath, UK	Symbiosis	Education and outreach
Louis Tisa	University of New Hampshire, USA, LST@hypatia.unh.edu	Symbiosis	Education and outreach
Steve Forst	University of Wisconsin, USA; sforst@uwm.edu	Symbiosis	Education and outreach
Heidi Goodrich-Blair	University of Wisconsin, USA; hgblair@bact.wisc.edu	Symbiosis	Education and outreach,
Creg Darby	University of Alabama at Birmingham, USA; creg@uab.edu	Symbiosis	Education and outreach
Diana Cox-Foster	Pennsylvania State University, USA	Immune response and host specificity	Education and outreach
Susan Bornstein-Forst	Marian College, WI sbornsteinforst@mariancollege.edu	Anhydrobiosis	Undergraduate education and outreach
Michael Klein	USDA, ARS, Wooster, USA Klein.10@osu.edu	Field application, genetic variability	Outreach, and extension
Albert Pye	BioLogics Co., PA, USA	Production, formulation, and field application	Extension
Simon Piggott	BeckerUnderwood, UK	Production, formulation, and field application	Workshops and extension
James Cate	IBCS, Lawrenceburg, USA	Production, formulation, and field application	Extension
Ralf Ehlers	University of Kiel,Germany; ehlers@biotec.uni-kiel.de	Dauer recovery and Extension	Outreach, and extension
Dawn Gouge	University of Arizona, USA	Urban IPM	Extension
Ed Lewis	Virginia Tech, USA	Host finding	Outreach and Extension
Harry Kaya	University of California	Nematode ecology	Education, outreach and Extension
Lerry Lacey	USDA, Washington State	Biological Control	Extension
Jim Campbell	Integrated Genomics	Sequencing and annotation	Sequencing, bioinformatics, and database development