हेटेरोरैबडाइटिस सूत्रकृमि के *फोटोरैबड्स* जीवाणु के साथ सहजीविता में लिप्त कारकों का आण्विक अनुसंधान

A molecular investigation of *Heterorhabditis* nematode factors involved in symbiosis with *Photorhabdus* bacteria

CHAITRA GANAPATI BHAT

Division of Nematology
ICAR- Indian Agricultural Research Institute
New Delhi -110012

2022

A molecular investigation of *Heterorhabditis* nematode factors involved in symbiosis with *Photorhabdus* bacteria

By

CHAITRA GANAPATI BHAT

A thesis submitted to the faculty of the Post-Graduate School,
ICAR-Indian Agricultural Research Institute, New Delhi,
in partial fulfilment of requirements
for the award of the degree of

DOCTOR OF PHILOSOPHY

in

NEMATOLOGY

2022

Approved by:

Advisory Committee

Chairman	: Dr. Uma Rao	Uma Rao
Co-Chairman	: Dr. Vishal Singh Somvanshi	Singh Somvanshi
Member	: Prof. Matiyar Rahaman Khan	
Member	: Dr. Vandna Rai	Vandna Rai
Member	: Dr. Bikash Mandal	

Division of Nematology
ICAR- Indian Agricultural Research Institute
New Delhi – 110 012, India

Dr. (Mrs.) Uma Rao
Head & Principal Scientist

CERTIFICATE

This is to certify that the thesis entitled "**A molecular investigation of *Heterorhabditis* nematode factors involved in symbiosis with *Photorhabdus* bacteria**" submitted to the Faculty of the Post-Graduate School, ICAR- Indian Agricultural Research Institute, New Delhi, in partial fulfilment of the requirements for the award of the degree of **Doctor of Philosophy** in **Nematology** embodies the results of bona-fide research work carried out by **Ms. Chaitra Ganapati Bhat** under my guidance and supervision, and that no part of this thesis has been submitted for any other degree or diploma.

The assistance and help availed during the course of investigation as well as source of information have been duly acknowledged by her.

Place: New Delhi, India
Date:

Dr. (Mrs.) Uma Rao
Chairperson
Advisory Committee

ACKNOWLEDGEMENT

It is a great privilege for me to express my sincere gratitude to my advisor Dr. Uma Rao (Head and Principal Scientist, Division of Nematology, ICAR-IARI, New Delhi) for the continuous guidance during my Ph.D. research program. I am highly indebted for her support, constant encouragement, scholarly suggestions, affectionate behavior and useful discussions during the course of investigation and preparation of manuscripts.

I owe my admirable indebtedness to Dr. Vishal Singh Somvanshi, Principal Scientist, Division of Nematology, ICAR-IARI, Co-chairperson of my advisory committee, for his constant support, versed advice, encouragements and generous guidance during the course of investigation.

Dr. Uma Rao and Dr. Vishal S. Somvanshi's innovative ideas and vast subject expertise have aided me in overcoming the challenges faced while conducting my research. Their great passion for research has motivated me a lot during my study.

My sincere thanks also goes to Dr. Matiyar Rahaman Khan, Professor, Division of Nematology, ICAR-IARI, New Delhi for providing guidance regarding both academics and research during the whole course of my study.

I would thank my advisory committe members Dr. Vandna Rai, Principal Scientist, ICAR- NIPB, New Delhi; and Dr. Bikash Mandal, Principal Scientist, Division of Plant Pathology, ICAR- IARI, New Delhi, for their useful suggestions and support whenever needed.

It is my privilege to express my profound gratitude to Dr. Adler R. Dillman, Department of Nematology, University of California, Riverside, USA who provided me an opportunity to join his team at UCR, USA and offered me the kind cooperation and guidance during the research program and preparation of manuscript.

I would also like to extend my gratitude to all the esteemed faculty members of the Division of Nematology, ICAR-IARI, New Delhi for their constant encouragement and suggestions during my study. Notable mention should include Dr. Gautam Chawla, Principal Scientist, Division of Nematology, for his instructions to capture excellent photomicrographs. I am also thankful to Dr. Anil Sirohi, Principal Scientist, Division of Nematology, for his kind help regarding both academics and hostel accommodations.

My deepest gratitude goes to my labmates Artha Kundu, Victor Phani, Manoranjan Dash, Khushbu Chauhan, Jyoti Kushwah, Chetna Mathur, Amit Ahuja, Madhurima Chatterjee, Jyoti Yadav, Nisha Jaiswal, Rami Kassam, Alkesh Hada, Divya Singh, Jyoti Antil, Gagandeep Singh, Sachin, Vyshali, Santhosh, Prakash, Shubham, Snehalata, Naveen, Aabid, Vivek Bhaiya, Puneeth Bhaiya,

Rajender Ji, Purshotam sir and all other lab members for their support and immense encouragement in each step of my project work. I am especially thankful to Roli Budhwar, Jeffrey Godwin for their support and co-operation during my research work.

My special thanks to Shipra Shahi and Maahi Somvanshi for their support and joyful company. I am thankful for the moral support and help form my friends, Snow, K. R. Rakshita, Krutika Patil, Soumya Chitnis, Gangubai Manguli, Chaitra Ternmakki, Bolli Venu Babu, Shivaji, Priyanka di, Ankia Paul, Bharati Hegde, Parvati Hegde, Amit, Praveen, Vimal.

I wish to express my heartfelt regards to Dr Varghese Thomas, Harpal Dhillon, Aklima Lima, Heiley Patel, Rithika Nair, Kyle, Sophia, Jacob, Kingslie, Chel Washigton, Jedeliza, Maha, Ayusshh, Lenore Davison, Mark Davison and Fulbright cohort (2020-2021) for their support during my stay at UC, Riverside.

I am also grateful to Triveni Mess for providing me nice food and joyful company during my study period at Delhi. My special thanks to one and all at the Division of Nematology for their help provided to carry out my studies in a successful manner.

I am grateful to the Director, Dean & Jt. Dir. (Edn) and Jt. Dir. (Res) of the institute for providing me necessary facilities and monetary support to carry out my research work. I also acknowledge the 'Genomics assisted crop improvement and management project' under World-bank and Indian Council of Agricultural Research sponsored NAHEP-CAAST (National Agricultural Higher Education Project-Centre for Advanced Agricultural Science and Technology) project for financial support. I am also grateful to the United States-India Educational Foundation (USIEF), Fulbright Foundation and Fulbright-Nehru Doctoral Research Fellowship for providing me a great opportunity to visit and learn new skills and gain new knowledge at UC Riverside, USA.

I also acknowledge the favours of numerous persons who, though not been individually mentioned here, have all directly or indirectly contributed during the course of the study.

Last but not the least I convey my hearty thanks and regards to my family, all my teachers and God Almighty for their constant support, wishes and encouragement, which has constantly motivated me to follow the path towards completion of this thesis. I am greatly beholden of vocabulary and owe a deep sense of honor to my beloved father for his unending love, sacrifices and moral support and motivation to follow my dreams.

Place: New Delhi, India

Date: March 30, 2022

Chaitra Ganapati Bhat

CONTENTS

List of Tables

List of Figures

CHAPTER 1 INTRODUCTION

Interactions between animals and microbes are ubiquitous in the biological world. Animal-microbe interactions range from facultative to obligate associations, and from pathogenesis to mutualistic relationships. The mutualistic associations can be simple mono-specific associations (for e.g., nematodes-bacteria symbiosis), relatively simple consortia (2-25 bacterial species in an animal, for e.g., leach gut consortium, insect gut consortium) and highly complex consortia (for e.g., vertebrate guts colonized by $> 10^2$ to 10^3 species of bacteria) (Ruby, 2008). It is well established that symbionts can affect the physiology, nutrition, metabolism, immunity, behaviour, growth and development of an animal host, and can provide protection to the host. An emerging awareness of the impact of beneficial microorganisms on animal biology including human health has increased interest in animal-microbe symbioses and the initiation of mega projects such as the Human-Microbiome project (McFall-Ngai, 2008). However, the mechanisms governing the patterns of host-symbiont associations are largely unknown. The microbial complexity of the insect and vertebrate guts is a major limiting factor in understanding the microbe-animal interactions in these models. Several invertebrates including nematodes are used as models of microbial symbiosis (Chaston and Goodrich-Blair, 2010; Murfin et al., 2012).

Species diversity, numerical abundance, pervasiveness, experimental tractability and applied importance of nematodes makes them excellent test subjects to study diverse biological phenomenon (Mckenzie-Bird, 2005; Nigon and Félix, 2018; Sommer and Bumbarger, 2012). Nematodes exhibit diverse trophic lifestyles ranging from free-living forms to parasites of plants and animals. Animal-parasitic and plant-parasitic nematodes are major threats to the health of human and veterinary animals and to agricultural crop production. However, entomopathogenic nematodes (EPNs) are distinct insect-parasitic nematodes that are considered as a beneficial group of worms. Nematodes belonging to the genera *Heterorhabditis*, *Steinernema* and some species of *Oscheius* are categorized as entomopathogenic nematodes (Dillman et al., 2012; Gaugler, 2002). These nematodes are mutualistic hosts for entomopathogenic bacteria (EPB). Along with their bacterial partner, they are obligate insect parasites (Kaya et al., 2006). EPNs have a significant economic and biological impact as biocontrol agents for agriculturally important insect pests and as a genetically tractable model system for studying a wide range of biological phenomena from mutualism to parasitism

(Campos–Herrera et al., 2012; Lacey and Georgis, 2012; Murfin et al., 2012; Stock, 2005).

EPNs have several advantages over chemicals as biocontrol agents. They are highly virulent and kill the host within 48–72 hours of infection (Kaya and Gaugler, 1993). Their broad host range, ease of multiplication and application, compatibility with chemical pesticides, active host-seeking abilities, and activity in cryptic habitats plus environmentally safe nature makes them superior to chemicals in the management of target pests (Gozel and Gozel, 2016; Lacey and Georgis, 2012). Several countries are working to develop safer bioinsecticides out of these beneficial microorganisms (Lacey et al., 2015). These attempts include searching for indigenous strains to include in pest management programs. More than a hundred species of *Steinernema* and nearly twenty-one species of *Heterorhabditis* have been described from various parts of the world. Several species of *Heterorhabditis* and *Steinernema* are commercially exploited for insect pest control (Bhat et al., 2020; Lacey et al., 2015).

EPNs are considered as an evolutionary bridge between free-living and vertebrate parasitic nematodes. *Heterorhabditis*, in particular, is regarded as a complementary model to *Caenorhabditis elegans* for studying ecology and interspecies interactions (Bai et al., 2013; Ciche, 2007). Phylogenetically, *Heterorhabditis* belongs to the same clade as that of *C. elegans* (clade 9), whereas, *Steinernema* belongs to clade 10 (Megen et al., 2009; Stock, 2015). The *Heterorhabditis* nematode lives in a mono-specific symbiotic relationship with the enterobacterium *Photorhabdus*. It is emerging as an experimental model system to understand beneficial symbioses. These nematodes are co-evolving with their bacterial symbionts for millions of years and form very specific symbiotic relationships, and mostly do not colonize other species of *Heterorhabditis* and *Photorhabdus*. There is functional specificity in the interactions, as laboratory cross-colonization experiments demonstrated that the most successful symbioses are produced by the natural pairs (Han and Ehlers, 2000). Virulence, developmental and reproductive fitness and field performance of these insect-parasitic nematodes largely depend on their bacterial symbionts.

Recent progress has led to the identification of the morphology and physiology of the symbiont colonization sites in the nematode intestine and the developmental progression of this symbiotic relationship (Ciche et al., 2008). *Heterorhabditis* selectively pass on their symbionts to their offspring by a sophisticated maternal

transmission process. *Photorhabdus* symbionts attach and form biofilm at posterior nematode intestinal cells IN9L and IN9R when nematode is at early adult stage and marks the beginning of transmission process (Ciche et al., 2008). Some of the bacterial genes and processes necessary for symbiotic colonization of the nematodes have also been identified (Easom et al., 2010; Easom and Clarke, 2012; Somvanshi et al., 2010). The bacterial and nematode genomes have been sequenced. Genome information of *Photorhabdus* bacteria (~80 genome assemblies of various *Photorhabdus* species and strains) and *H. bacteriophora* are available in the public domain, thus greatly facilitating functional genomics (Bai et al., 2013; Duchaud et al., 2003; Somvanshi et al., 2019).

In spite of all the recent developments in our understanding of the *Heterorhabditis-Photorhabdus* symbiosis system - nothing is known about the host nematode factors needed for symbiosis or host specificity. A deeper understanding of this symbiosis, including the relative roles played by the microbe versus those played by the host, will yield insights into fundamental processes underlying the ubiquitous association of microbes with animals. Additionally, omics studies in *Heterorhabditis* are still in its infancy. When the present study was commenced, whole-genome information was available for only one species of heterorhabditid nematode i.e, for *H. bacteriophora*. Available transcriptomic data of *Heterorhabditis* are focused only on IJ stage (Somvanshi et al., 2016; Vadnal et al., 2017). Generating quality genome and transcriptome information on these beneficial nematodes combined with genetic interrogation tools such as RNAi and CRISPR-Cas9 are potent methods to explore EPN biology and lay the groundwork for improving their bio-control traits and to realize their full potential as model systems (Lu et al., 2016).

Keeping in view the above-mentioned gaps in knowledge, the following objectives were formulated.

1. Whole genome sequencing of nematode *Heterorhabditis indica*
2. Elucidation of genes involved in nematode-bacterium symbiosis by transcriptome sequencing of *Heterorhabditis* at biofilm formation stage of symbiosis
3. Standardization RNAi gene silencing in *Heterorhabditis* nematode

2.1 Animal-microbe symbiosis

Symbiosis is a biological phenomenon in which distinct organisms live together for their mutual intrinsic benefits. Symbiotic interactions between animals and microorganisms are ubiquitous in nature. Almost all animals live with beneficial microbiota, which can be found in the digestive tract, on the tissue surfaces, and/or in the specialized organs of animals. Animals have established a wide range of symbiotic relationships with microbes which have often lasted millions of years (McFall-Ngai, 2008). Animal-microbe associations can be simple mono-specific associations (e.g., nematodes-bacteria symbiosis), relatively simple consortia (2-25 bacterial species in an animal, e.g., leach gut consortium, insect gut consortium), and highly complex consortia (e.g., vertebrate guts colonized by $> 10^2$ to 10^3 species of bacteria) (Ruby, 2008). Beneficial symbionts are typically transmitted from generation to generation. Based on the inter-host transmission, symbiotic systems are classified as open (horizontal transmission), closed (vertical transmission), and mixed symbioses (both kinds of transmissions) (Perreau and Moran, 2022).

Symbionts can influence host biology and evolution in incredibly diverse ways. Symbiotic associations with microbes have contributed to the immense diversity in development, morphology, and lifestyles seen across animal phyla. Several theories such as serial endosymbiotic theory (SET), hologenome theory of evolution, it's the song, not the singer (ITSNTS) theories have emphasized the role of microorganisms in the evolution of animals and plants. Animals and plants evolved by acquiring additional structures and functions, either by changing their DNA or by acquiring symbionts (Doolittle and Booth, 2017; Margulis and Fester, 1991; Zilber-Rosenberg and Rosenberg, 2008). Symbionts affect the physiology, immunity, metabolism, behavior, growth, and development of the host, and offer protection to the host. Astonishingly, trillions of bacteria live symbiotically in a vertebrate gut without eliciting any potentially harmful host immune response. An emerging awareness of the impact of beneficial microorganisms on animal biology including human health has increased the interest in animal-microbe symbioses and the initiation of mega projects such as the Human-Microbiome project (McFall-Ngai, 2008).

The majority of information on host-microbial mutualism is focused on bacterial genes and factors, whereas very limited information is available on molecular determinants of host animals involved in microbial symbiosis. Signaling pathways and other regulatory mechanisms involved in mutualistic associations are poorly understood as compared to host-pathogenic microbe interactions. Only a few studies have characterized the host factors involved in regulating symbiotic microbiota of the gut in higher animals. Such studies highlighted the importance of maintaining spatial segregation between microbiota and the host tissues to prevent infection and elicitation of potentially harmful immune responses. The role of the mammalian resistin-like molecule (RELM- a bactericidal protein) and RegIIIγ (a secreted antibacterial lectin) in fostering the host-bacterial mutualism was demonstrated. These proteins are part of host immune mechanisms and modulate the spatial segregation between the microbiota and the intestinal epithelium. (Propheter et al., 2017; Vaishnava et al., 2011). Frantz et al., (2012) showed that MyD88-dependent Toll-like receptor signaling is a key mediator of microbial-host cross-talk and concluded that MyD88 signaling in intestinal epithelial cells is crucial for the maintenance of gut homeostasis. The production of RegIIIγ and many other effectors regulating the gut bacteria is regulated through the MyD88 pathway (Frantz et al., 2012).

The microbial complexity of the insect and vertebrate guts limits our ability to understand the molecular mechanisms underlying the interactions in these models (Ruby, 2008). Some specific invertebrates are exploited as models of microbial symbiosis due to their natural and specific association with relatively simple microbiota and due to the experimental accessibility of such systems. This list includes Hawaiian bobtail squid, hydra, medicinal leeches, earthworms, aphids as well as nematodes (Chaston and Goodrich-Blair, 2010).

2.2 Nematodes as models of microbial symbiosis

Nematodes belong to the phylum 'Nematoda' and are among the most abundant multicellular animals on Earth. They are second to Arthropoda in species diversity with an estimated total number of a million species. Currently, around 30,000 species are described and classified under two classes and 12 clades in the phylum Nematoda (De Ley, 2006; Holterman et al., 2006; Kiontke and Fitch, 2013; Smythe et al., 2019). These ubiquitous organisms exhibit diverse lifestyles. The majority of the reported nematode species are free-living forms, found feeding on microorganisms and organic debris in

soil and aquatic ecosystems (~50% are marine and 25% dwell in soil and freshwater). About 15% of the recognized species are parasites of animals, including humans and insects, whereas about 10% are plant parasites (Lambert and Bekal, 2002). While plant and animal-parasitic forms are targeted for management due to their menace, insect-parasitic forms are considered beneficial and utilized in biological insect pest management programs. Free-living forms are also an important part of the soil and aquatic ecosystems and play a vital role in mineralization and nutrient recycling (Hay, 2008; Hodda et al., 2009). Thus, nematodes significantly impact human health, agriculture, ecosystem, and economies in one or another way.

Many species of nematodes are popular as model systems in a variety of biological research areas because of their practical importance and experimental tractability. The free-living nematode *C. elegans* is a remarkable model frequently chosen for basic molecular biological research and it has the distinction of being the first multicellular organism to have its whole genome sequenced (Meneely et al., 2019). Nematodes have been used as test subjects to decipher many of the biological questions in the field of genetics, developmental and evolutionary biology, ecology, omics research, etc. (Hallem et al., 2007; Mckenzie-Bird, 2005; Sommer and Bumbarger, 2012; Sommer and McGaughran, 2013). Nematodes are also the greatest model for investigating a variety of biological interactions, including animal-bacterium interactions. Nematodes have been observed to form a range of associations with bacteria, ranging from transitory to persistent associations and pathogenic/parasitic interactions to mutually benefiting symbiotic relationships (Bulgheresi, 2016; Murfin et al., 2012).

Among nematodes, entomopathogenic nematodes (EPNs) symbiotic with entomopathogenic bacteria (EPB), marine nematode *Laxus oneistus* associated with thiotrophic bacteria, animal-parasitic filarial nematodes colonized by *Wolbachia* symbionts are exclusively studied to understand microbial symbiosis. Nematode models of microbial symbiosis are presented in table 2.1. Insights of nematode-bacterial interactions can be extrapolated to higher animals since there is a degree of conservation in biological aspects between nematodes and higher animals including humans (Bulgheresi, 2016; Chaston and Goodrich-Blair, 2010). EPNs of the genera *Steinernema* and *Heterorhabditis* are known to form symbiotic associations with insect-pathogenic gammaproteobacteria (Burnell and Stock, 2000).

Phylogenetically *Steinernema* and *Heterorhabditis* belong to distinct clades (*Steinernema* belongs to clade 10 and *Heterorhabditis* belongs to clade 9 (Holterman et al., 2006)). However, both of the nematodes exploited insects as a nutritional niche due to their symbiotic relationships with bacteria (Gaugler, 2002). Steinernematids are primarily associated with *Xenorhabdus* bacteria. Recent evidence suggests that *Steinernema*-bacteria associations may be more complex. Steinernematids may have a frequently associated microbiota (FAM) that includes several Proteobacterial species in addition to *Xenorhabdus* as core symbionts (Ogier et al., 2020). However, to the best of our knowledge, the insect parasitic nematode *Heterorhabditis* primarily lives in a mono-specific symbiotic relationship with the enterobacterium *Photorhabdus* (Chaston and Goodrich-Blair, 2010; Ciche, 2007). Thus, *Heterorhabdtis-Photorhabdus* symbiosis is a natural binary system i.e., one host and one symbiont species are involved. Ruby (2008) highlighted the characteristics which are valuable in using an organism as a genetic model of symbiosis. As per the criteria, *Heterorhabdtis-Photorhabdus* is the powerful and genetically tractable genetic model for the study of animal-microbe relationships (Table 2.2).

In addition to mutualism, they are an exemplary model for the study of parasitism, ecological adaptation, etc. *Heterorhabditis* is regarded as a complementary model to *C. elegans* and considered a bridge between free-living and vertebrate parasitic nematodes (Bai et al., 2013; Campos–Herrera et al., 2012; Ciche, 2007). The symbiotic pair of EPNs and entomopathogenic bacteria has a significant economic value as biocontrol agents of insect pests that are well suited for pest management in cryptic and underground soil habitats, as well as integrated pest management programs (Lacey and Georgis, 2012).

2.3 *Heterorhabditis-Photorhabdus* symbiotic pair

2.3.1 Nematode partner: *Heterorhabditis*

Entomopathogenic *Heterorhabditis* nematodes belong to phylum-Nematoda, class-Chromadorea, order-Rhabditida, suborder-Rhabditina, superfamily-Rhabditoidea, family-Heterorhabditidae and genus-*Heterorhabditis* (Poinar et al., 1992; Poinar, 1975). *H. bacteriophora* is the type *s*pecies. Poinar (1975) isolated it from *Heliothis punctigera* in South Australia and described it as a new genus, species, and family Heterorhabditidae under Rhabditida order. According to recent data, 21 species

of *Heterorhabditis* have been identified from different parts of the world (Bhat et al., 2020). Among *Heterorhabditis*, three indigenous species have been reported from India. They are *H. indica* (Poinar et al., 1992), *H. bacteriophora* (Bhat et al., 2019) and *H. baujardi* (Vanlalhlimpuia et al., 2018). *H. indica* was the only new species described from India till date and it was described from populations recovered from the sugarcane top borer, *Scirpophaga excerptalis* (Pyralidae: Lepidoptera) from Coimbatore, India (Poinar et al., 1992).

2.3.2 Bacterial partner: *Photorhabdus*

Photorhabdus are bioluminescent, gram-negative bacilli (photo - light producing and rhabdus -rod shape), asporogenous and motile bacteria. They belong to the phylum: Proteobacteria, class: Gammaproteobacteria, order: Enterobacterales, and family: Morganellaceae. This bacterium was first isolated from *H. bacteriophora* and it was included in the genus *Xenorhabdus,* as *X. luminescens* (Thomas, 1979). The genus *Photorhabdus* was proposed by Boemare et al. in 1993 to accommodate symbiotic bacteria of *Heterorhabditis* EPNs and only one species *P. luminescens* was described (Boemare et al., 1993). Later, Fischer-Le Saux et al., (1999) split the *Photorhabdus* genus into three species, *P. luminescens*, *P. temperata* and *P. asymbiotica*. There are presently 21 species in the genus *Photorhabdus* with valid published names (Machado et al., 2021). *Photorhabdus* are normally found as symbionts that colonize the gut of *Heterorhabditis* nematodes. For example, *P. luminescence* subsp. *laumondii* forms mutual association with *H. bacteriophora* whereas *P. akhurstii* is symbiont of *H. indica* (Abd-Elgawad, 2021). *Photorhabdus* are pathogenic to wide range of insects. While one species i.e., *P. asymbiotica* has been shown to be associated with sporadic human infections in Australia and USA (Hapeshi and Waterfield, 2017).

2.4 Life cycle of *Heterorhabditis* and transfer of *Photorhabdus* symbionts to offspring

Life cycle of *Heterorhabditis* nematode includes the egg stage, four juvenile stages, and the adult stage. Specialized third stage juvenile is the infective stage and is termed as infective juvenile (IJ or IJ3). It is the only free-living stage found outside of the insect host in nature. The IJ stage is analogous to the dauer L3s of *C. elegans,* infective J3s of *Steinernema* and other animal-parasitic nematodes. These soil dwelling,

non-feeding IJs are developmentally arrested and environmentally resistant as compared to other life stages. IJs carry symbiotic bacteria in their gut which play a vital role in successive developmental stages of the nematode life cycle. IJs offer protection to the bacteria when in the soil environment. The IJs hunt for a suitable host and initiate the infection process by invading the insect hosts through natural body openings and/or through the cuticle. Upon invasion, they release their symbionts in the insect haemocoel. The bacterial symbionts help to modulate insect immune systems, rapid killing of the host, release of nutrients from the cadaver, and guard the cadaver from opportunists. The infectivity, virulence and reproductive fitness of nematodes largely depend on their symbionts. Nematodes complete two or three generations inside the insect cadaver depending on the nutritional resources. In case of *Heterorhabditis,* first generation is hermaphroditic which is followed by amphimictic/sexual generations. Eventually a new generation of IJs exit the cadaver in search of new insect hosts. IJs contain maternally transmitted symbiotic bacteria (Ciche, 2007; Shapiro-Ilan et al., 2012).

Thus, symbionts have a distinct role in the life cycle of EPNs, where they facilitate insect parasitism, support proper growth, development and reproduction of nematodes. Axenic nematodes do not cause insect mortality and do not proliferate inside the insect cadavers. L1s generated fail to develop further in the absence of symbionts (Han and Ehlers, 2000). EPNs must selectively transmit their symbiont to infect and proliferate inside insects. *Photorhabdus* symbionts are maternally transmitted to offspring by a sophisticated process in order to achieve symbiont-specific transmission. Symbiont transmission is developmentally regulated. Synchronous with nematode development, cell and stage-specific adherence or invasion of host cells occurs which helps symbionts to persist as nematodes develop and move to a new host.

A brief overview of the symbiont transmission process in *Heterorhabditis* nematodes is as follows: Upon invasion of an insect host, IJs regurgitate bacteria within 4-6 hours of entry. IJ recovers from the non-feeding dauer stage and continues its normal development. When nematode is at the early adult stage (36-40 h post IJ recovery), symbionts adhere and form biofilm in the posterior intestine of the maternal nematode. Symbionts gradually invade rectal glands and continue intracellular growth leading to vacuole multiplication. Coinciding with the adult stage (112-120 h post-recovery), symbionts are released into the pseudocoelom. At the same time, juveniles

start developing inside maternal hermaphrodite nematode causing *endotokia matricida*. Symbionts released into the maternal body cavity become available to the offspring developing in the pseudocoelom. Symbionts adhere to the cells of pharyngeal intestinal valve of pre-IJ, followed by colonization of the IJ intestinal lumen (Ciche et al., 2008; Waterfield et al., 2009).

2.5 Molecular genetics and genomics of *Heterorhabditis-Photorhabdus* symbiosis

Molecular to omics level studies have been applied to diverse nematode-bacterial interactions to decipher the processes of communication and/or exchange between the symbiotic partners. These investigations have contributed to our understanding of mutualism, parasitism, innate immunity and genome evolution (Chaston and Goodrich-Blair, 2010). Several genomics and proteomics scale studies have been conducted on *Heterorhabditis* and *Photorhabdus*. Availability of omics data such as genome sequences of both, the bacteria and nematodes greatly facilitate functional, comparative genomics and other molecular investigations to decipher the intricacies of interactions (Bai et al., 2007; Duchaud et al., 2003; Somvanshi et al., 2016, 2019). In comparison to the nematode partner, there is ample information about bacterial determinants involved in symbiosis.

2.5.1 Molecular insights on mutualistic *Photorhabdus*:

The first *Photorhabdus* genome, *P. laumondii* TTO1 was published in 2003 and 4,839 protein-coding genes were predicted in 5.68 Mb genome of the bacteria (Duchaud et al., 2003). It was shown that *Photorhabdus* genome contains a large number of genes encoding adhesins, toxins, hemolysins, proteases and lipases as well as a wide array of genes (approximately 6% of the genome) involved in the synthesis of antibiotics such as polyketide synthases (PKSs) and non-ribosomal peptide synthases. The possible function of these genes in mediating interactions between *Photorhabdus* and other organisms, such as nematode host colonization, invasion and bioconversion of the insect cadaver, and in the elimination of competitors was anticipated (Duchaud et al., 2003). There are now 81 genome assemblies (complete and incomplete) available in GenBank (as on February, 2022) across 19 of the 21 different species defined in this genus. This list includes genome assemblies of 12 *Photorhabdus* isolates which are clearly distinct from currently recognized species and tentatively designated at the species level (https://www.ncbi.nlm.nih.gov/datasets/genomes/? taxon=29487&utm_

source=genome&utm_medium=referral). The published *Photorhabdus* genomes vary in size from 4.70 to 5.83 Mb, and only *P. asymbiotica* has been shown to contain plasmid DNA (Wilkinson et al., 2009, 2010). Genes related to both the nematode and insect host specificities were found in the genomes of bacteria associated with EPNs (Chaston et al., 2011). For example, in a recent study, genetic regulatory networks controlling symbiosis with the EPN, pathogenicity with the insect host were elucidated in *P. heterorhabditis* strain ETL (nematode partner *H. zealandica* strain ETL) based on genome analysis (Lulamba et al., 2021). Several genes associated with the biosynthesis of invasion and intracellular resistance proteins, fimbriae and type IV pili-like proteins were reported in ETL genome (Lulamba et al., 2021). It was postulated that such genes may facilitate not just insect host invasion but also nematode gut colonization (Waterfield et al., 2009). Genomic resources lay the foundation for further genetic validation studies.

Much progress has been made in identifying and characterizing bacterial virulence determinants. A cascade of molecules such as degradative enzymes, toxins and antibiotic compounds are involved in the pathological interaction of *Photorhabdus* with insects, resulting in immunological suppression and eventual death of the insect hosts. However, as a symbiont of *Heterorhabditis*, *Photorhabdus* contributes in IJ recovery, proper growth and development of nematodes and colonize the nematode offspring to continue their specific mutualistic association generations after generation. Some studies suggested the possible overlap between molecular factors and processes involved in pathogenic and mutualistic interactions (Easom and Clarke, 2012; Waterfield et al., 2009). Various bacterial genes and processes necessary for symbiotic interaction with the nematodes such as factors involved in nutrition, surface structures, gene regulation and stress responses have been identified.

Food signal produced by the *Photorhabdus* bacteria is known to stimulate IJ recovery in *Heterorhabditis* nematodes (Strauch and Ehlers, 1998). A stilbene compound termed (E)-3,5-dihydroxy-4-isopropylstilbene (IPS) was shown to be a crucial part of this food signal. *Photorhabdus* is the only non-plant organism known to produce a stilbene (Joyce et al., 2008; Williams et al., 2005). Biochemical pathways and genes (e.g., *stlA, stlB*) involved in IPS production in *Photorhabdus* have been characterized. IPS is a multifunctional molecule and play role in both mutualism and pathogenicity. It acts as an antibiotic, an insect immune system inhibitor, as well as an

inter-kingdoms signal and a factor in nematode growth, mostly in IJ recovery (Chalabaev et al., 2008; Joyce et al., 2008; Williams et al., 2005).

Some studies have proved the temporal separation in the role of bacteria as a pathogen of insects and as a supporter of nematode growth and development. This transition in the role appears to be bacterial growth phase-dependent (Clarke, 2020; Joyce and Clarke, 2003). During exponential growth, bacteria have been shown to produce a majority of virulence related factors, whereas during the post-exponential phase i.e., during the stationary phase, bacteria support the growth of nematode partner. High-level bacterial secondary metabolism including production of the IPS metabolite, anthraquinone (AQ) pigments and bioluminescence has been observed during the stationary phase of bacteria (Clarke, 2020; Hu and Webster, 2000). Earlier *hexA* gene (LysR-type transcriptional regulator/LTTR) was shown to play a significant role in the regulation of this pathogen to symbiont transition in *P. temperata* strain K122 (Joyce and Clarke, 2003). *HexA* has also been shown to control secondary metabolism and phenotypic heterogeneity in *P. luminescens* (Engel et al., 2017; Langer et al., 2017). The role of TCA cycle-dependent metabolic switch as well as (p)ppGpp alarmone in the transition from pathogen to mutualist was demonstrated (Bager et al., 2016; Lango and Clarke, 2010). *RelA* and *SpoT* genes are associated with (p)ppGpp production. The *ΔrelA spoT* mutant of *Photorhabdus* was unaffected in virulence to the insect host, but attenuated in its ability to produce IPS, AQ and light as well in its ability to support nematode growth and development, suggesting these two roles of bacteria are not coupled (Bager et al., 2016). Mutation of the post-transcriptional global regulator *Hfq* in *Photorhabdus* resulted in the abolishment of secondary metabolite production. There was no production of IPS or AQ. Though *Δhfq* mutant was unaffected in virulence, failed to support nematode development. RNA-Seq and double mutation analysis revealed the epistatic link between *HexA* and *Hfq*. These studies suggested that *HexA/Hfq* as an important node in the regulatory network in modulating the transition from pathogen to the mutualistic state of bacteria (Tobias et al., 2017). All these studies in part explained the regulatory mechanisms surrounding the production of secondary metabolites and suggested a strong role of bacterial secondary metabolism in maintaining a symbiotic relationship with nematode where secondary metabolites assist the development and reproduction of nematodes. These studies reinforce earlier studies where the dependence of *Heterorhabditis* nematodes on symbiotic bacteria for

the provision of secondary metabolites was suggested (Joyce et al., 2011). For example, phosphopantetheinyl transferase involved in the biosynthesis of antibiotic molecules was shown to be facilitating normal growth and reproduction of nematode (Ciche et al., 2001).

Information on nutritional interaction between *Photorhabdus* and the nematode suggested the importance of bacteria in providing some essential amino acids and vitamins for the growth and development of nematodes. The *cob/cbi* operon in the *Photorhabdus* genome codes for proteins involved in the production of adenosylcobalamin, an important precursor of the necessary vitamin B12. The ability of bacteria to produce adenosylcobalamin was proposed to be important for symbiosis (Clarke, 2020). Phase I cells of *Photorhabdus* have been shown to produce two different forms of intracellular protein inclusions (type 1 and type 2) during the stationary phase of growth. These proteins are enriched in amino acids that are essential for nematode growth. Genes *cipA* and *cipB* encoding protein components that form inclusion bodies were characterized. Mutation in these bacterial genes affected the bacterial ability to support nematode growth and development (Bintrim and Ensign, 1998; Bowen and Ensign, 2001). Watson et. al., (2005) identified that *P. temperata* K122 *ΔexbD* mutant did not support the growth of nematode partner *H. downesi*. *ExbD* encodes an important component of the ExbB-ExbD-TonB siderophore complex and they suggested the importance of siderophore-mediated iron uptake in maintaining the bioavailability of iron for nematodes (Watson et al., 2005). In contrast, *P. luminescens* TT01 *ΔexbD* mutant was unaffected in its symbiotic interaction with *Heterorhabditis*, suggesting variation of iron uptake in different *Photorhabdus-Heterorhabditis* systems (Clarke, 2020; Watson et al., 2010). Additionally, involvement of yfeABCD operon of *P. luminescens* TT01 (encoding the Yfe divalent cation transporter) in Mn2+-uptake during colonization of the gut of the IJs was predicted. This transporter was shown to be involved in iron uptake during pathogenicity to *Manduca sexta,* suggesting the invertebrate host-dependent substrate specificity of Yfe transporters in *Photorhabdus* (Watson et al., 2010).

The only study conducted to identify the genomic loci of *Photorhabdus* involved in the specificity of nematode interaction. Gaudriault et al. identified 8 genomic regions possibly involved in host specificity by comparing gDNA of *P. luminescens* subsp. *laumondii* TT01 with *P. temparata* XlNach strains in a microarray

hybridization experiment. These analyses suggested the potent involvement of *lsr* locus in interaction with *H. bacteriophora*. The *lsr* locus is known to encode ABC transporters and control cell-cell signalling in quorum sensing in enterobacteria such as *Salmonella enterica* serovar *typhimurium* and *Escherichia coli*. It was suggested that termination of cell-cell signaling mediated by *lsr* locus could be important for the physiological shift of *Photorhabdus* in insect cadaver to recolonize the nematode intestine (Gaudriault et al., 2006).

Recent progress has led to the identification of the morphology and physiology of the symbiont colonization sites in the nematode intestine and the developmental progression of this symbiotic relationship. Bacteria adhere, persist, invade and grow inside nematode cells, breaching the alimentary tract to gain access to the developing IJs in the mother's body. The IJs select for bacteria that adhere to pharyngeal-intestinal valve cells, which colonize the intestinal lumen of IJs (Ciche et al., 2008). Subsequent genetic analyses revealed bacterial determinants important for the symbiont transmission process. Somvanshi *et al.,* (2010) showed that *Mad* (maternal adherence) fimbriae as the primary bacterial factor responsible for adherence to the maternal nematode gut (Somvanshi et al., 2010). An invertible promoter switch called the *madswitch* control the expression of *Mad* genes. The *madswitch* can be in either an ON (Mad pili produced) or an OFF (Mad pili not produced) orientation. In the primary form of bacteria (P-form) *madswitch* is in OFF orientation and these forms are involved in insect pathogenicity and aid in nematode growth and development. M-form cells *(madswitch* ON bacteria) are a subpopulation of *Photorhabdus* with small colonies that do not produce antibiotics, no bioluminescent, no involvement in nematode growth and are avirulent to insects. M-form cells are capable of colonizing the maternal nematode, thereby play a unique role in the transmission process (Somvanshi et al., 2012).

Bennet and Clarke (2005) showed that mutation in the *pbgE1* (fifth gene in the pbgPE operon) of *P. luminescens* TT01 attenuated the ability of bacteria to colonize the nematode as well as virulence against insects (Bennett and Clarke, 2005). Easom et al., (2010) also screened for transmission mutants in *Photorhabdus* and identified six genetic loci (*pbgPE*, *galE* and *galU proQ*, *asmA* and *hdfR)* involved in assembly, maintenance of lipopolysaccharides (LPS) and cell surface associated factors of bacteria regulating symbiont colonization in nematodes (Easom et al., 2010). Transmission mutant *Photorhabdus*, defective in LPS biosynthesis were avirulent to

insects implying that LPS is involved in both mutualism and pathogenicity (Easom et al., 2010). In addition, the LysR-type regulator *HdfR* was shown to be essential for proper transmission in *P. luminescens* TTO1. Transcriptome studies revealed that *HdfR* regulates the expression of nearly 124 genes, including genes involved in arginine metabolism, hydroxyphenylacetate catabolism and pigment production. *HdfR* was shown to play an important role in coordinating the temporal progression of the symbiont attachment and invasion in the gut of the adult nematode. Symbionts must be present in body cavity of maternal nematode coinciding with the developing IJs which are in colonization-permissive state. This guarantees that IJs choose for symbionts with the best fitness. It was shown that mutation in *hdfR* caused delayed entry of symbionts into nematode body cavity, consequently resulting in failed transmission into the IJs (Easom and Clarke, 2012). These studies highlight the importance of bacterial fimbriae and cell surface factors as well as temporal synchronization in mediating the symbiot and nematode interactions.

An and Grewal (2010) investigated the molecular basis of *Photorhabdus* bacterial persistence in the guts of IJs. Gene expression of *P. temperata* residing in the infective juveniles of *H. bacteriophora* as well as under in-vitro simulated starvation conditions were profiled using SCOTS (Selective capture of transcribed sequences) technique. This study demonstrated that symbiotic bacteria lower their nutritional reliance on the nematode partner and halt motility to ensure persistence in the nematode IJs. It has been demonstrated that mechanisms such as starvation, cellular acidification, and switching to the pentose phosphate pathway can enable bacteria to slow down the growth and overcome oxidative damage and nutrient constraint in the gut lumen of nematode IJs (An and Grewal, 2010).

2.5.2 Molecular insights on nematode host *Heterorhabditis*:

The first genome of an EPN to be released was that of *H. bacteriophora* strain TTO1 in 2013 (Bai et al., 2013). *H. bacteriophora* was one of the targets selected by the National Human Genome Research Initiative for whole genome sequencing (Ciche, 2007). Genome of an inbred line of TT01 nematode was sequenced using Roche 454 technology. Draft genome size was 77 Mb and assembly prediction using JIGSAW showed the presence of 21,250 putative protein coding genes. Almost half of the predicted proteins were suggested to be novel as they did not have noteworthy homologs in GenBank nonredundant protein database. A fewer KEGG orthologs (KO)

in majority of metabolic pathways were predicted in *H. bacteriophora* as compared to nine other nematodes. Relatively lower number of secreted proteins were predicted compared to free-living nematodes and it was postulated that nematodes depend on their symbionts for secreted effectors. An abundance of mariner transposons in *H. bacteriophora* genome was reported (Bai et al., 2007, 2013). Predicted proteome of *H. bacteriophora* had fewer shared orthologues with other rhabditids and animal parasitic *Strongylomorpha* (Blaxter and Koutsovoulos, 2015). An unusually high proportion (8.9%) of noncanonical splice sites (5′ GC splice donor site) were predicted in these gene models, whereas most of the metazoans including nematodes are known to have low proportions of noncanonical introns (less than 1%) (McLean et al., 2018).

McLean *et al.,* (2018) revisited the *H. bacteriophora* genome assembly and gene predictions. The original predicted proteome had poor completeness score (47.8% complete and 34.7% fragmented BUSCO). They used illumina short-read re-sequencing data generated from *H. bacteriophora* strain G2a1223 and presented the revised gene annotations for *H. bacteriophora* (McLean et al., 2018). G2a1223 an inbred derivative of *H. bacteriophora* strain Gebre and it has about 1 single-nucleotide change per ~2,000 nucleotides compared to the originally sequenced TT01 strain. They used an improved pipeline to re-predict gene sets using BRAKER1 and published transcriptome data. Newly predicted proteome contained 15,747 protein-coding genes with improved completeness value (BUSCO score for proteome: Complete-94%, Fragmented- 4.3%) and reduced complement of novel proteins (2889 unique proteins as opposed to 8962 in old predicted proteome). More number of secreted proteins were predicted as compared to original gene models. McLean *et al.,* (2018) proposed that the excess of unique genes, lack of secreted proteins, and other unusual features observed in the original gene set might be due to methodological issues, and recommended using improved tools on available genome data to validate the correctness and credibility of data (McLean et al., 2018). Despite the fact that *Heterorhabditis* has nearly 21 described species, genomic information for only one was available in the public domain until August 2021, namely the genome of *H. bacteriophora* TTO1-M31e (BioProject PRJNA13977).

Expressed sequence tags (ESTs) of *H. bacteriophora* strain GPS11 (Bai et al., 2007; Sandhu et al., 2006) and transcriptome of *H. bacteriophora* TTO1 adults (Bai et al., 2009), transcriptomics of trait deterioration in *H. bacteriophora* IJs (Adhikari et al.,

2009), transcriptome of *H. indica* IJs (Somvanshi et al., 2016), hemolymph-activated *H. bacteriophora* IJs (Vadnal et al., 2017), as well as transcriptomics of *Heterorhabdits* nematodes in response to abiotic stress (Levy et al., 2020; Sumaya et al., 2020) are available. However, most of the transcriptomic studies are highly focused on the IJ stage to decipher the parasitism genes as well as abiotic stress tolerance of the IJ stage. Transcriptomic data on other developmental stages are still lacking.

Some attempts have been made to deduce signalling pathways that control IJ recovery in *Heterorhabditis* based on genome and transcriptome data. It was shown that *Heterorhabditis* genome contains homologs of proteins known to be important in the regulation of dauer formation and/or recovery in *C. elegans*. Components of Insulin/IGF-1, cGMP-PKG, dafachronic acid, TGF-beta pathways were detected in *Heterorhabditis* (Bai et al., 2013; Somvanshi et al., 2016). In *C. elegans* it is proven that dauer formation is critical for longevity, immunity and abiotic stress resistance. However, dauer formation in *C. elegans* depends on adverse environmental conditions, whereas, in parasitic nematodes, it is an obligate part of the life cycle (Fielenbach and Antebi, 2008; Karp, 2018). So, it demands for further validation of such pathways in *Heterorhabditis* nematodes. Some attempts have been made to characterize molecular processes of IJ recovery in *Heterorhabditis*. It was demonstrated that muscarinic acetylcholine receptors and cGMP signalling mediated by the ASJ neuron are important in IPS-mediated IJ recovery in *Heterorhabditis* (Hallem et al., 2007). Another study on the recovery mechanism in the *H. bacteriophora* showed that genes belonging to metabolic pathways are differentially expressed between IJs and recovering nematodes (Moshayov et al., 2013).

Moreover, specific immune signaling and stress pathways are yet to be described in *Hetrerorhabditis*. The C-type lectins are a superfamily of proteins that recognize a broad range of ligands and function in innate and adaptive immune responses. Fewer C-type lectin domain proteins in *H. bacteriophora* compared to *C. elegans* were correlated to the mutualistic relationship with *P. luminescens* bacteria. It was predicted that these C-type lectins may function in interactions with insect hosts or bacterial symbionts (Bai et al., 2013). In our previous study, we demonstrated the possible role of *H. bacteriophora* C-type lectin *Hb-clec-78* in the context of symbiosis with *Photorhabdus* bacteria (Bhat et al., 2019).

Several genetic diversity assessments, genetic selection, breeding and DNA transformation studies have been done in *Heterorhabditis,* particularly in relation to virulence and abiotic stress tolerance traits (Anbesse et al., 2013; Levy et al., 2020). First successful transformation of an entomopathogenic nematode was proved in *Heterorhabditis* where *C. elegans* heat shock protein (*hsp70A*) gene transcriptionally fused to beta-galactosidase was delivered into the germline of *H. bacteriophora* using microinjection (Hashmi et al., 1995; Hashmi et al., 1998). In another study, GFP (Green Fluorescent Protein) under the control of a mechanosensitive promoter *mec-4* from *C. elegans* was used to transform *H. bacteriophora* (Hasi-Imi et al., 1997). Forward genetic studies using ethyl methane sulfonate (EMS) induced mutagenesis were conducted in *Heterorhabditis* (O'Leary and Burnell, 1997; Rahimi et al., 1993; Zioni, 1992). Reverse genetics using RNAi where soaking of post-embryonic developmental stage in dsRNA solution and microinjection-based delivery of dsRNA into the gonad of adult hermaphrodites for RNAi induced gene silencing has been demonstrated in *H. bacteriophora* (Ciche and Sternberg, 2007; Ratnappan et al., 2016).

Several of these selection, breeding, and genetic engineering efforts, however, did not yield the expected results in the field performance of the strains investigated. A comprehensive understanding of underlying genomic regulatory network of the nematode biology is very essential for trait improvement in EPNs (Lu et al., 2016). It is critical to investigate both the parasitic as well as mutualistic lifestyles of this worm. Many important traits of EPNs which are of practical importance in mass multiplication and field efficacy are influenced by their symbiotic bacteria. Specificity of mutualistic association between nematode and bacteria influences the recovery of infective juveniles, proper growth and development, virulence properties of entomopathogenic nematodes. It is suggested that genetic study of EPNs and further practical implications requires more advanced information and tools. Reportedly, genome information is one such fundamental resources to understand specific lifestyle of nematodes and their symbiotic bacteria. It may help to develop tools to render the bacteria and nematodes to enhance their virulence against specific insect hosts, mass production of quality EPN products, environmental resilience, and improve overall field performance and to further develop these organisms as model systems.

Table 2.1: Models of nematode-bacterial symbiosis

Nematode host	Symbiotic bacteria	Host structures harboring symbionts	Benefits of association	Symbiont Transmission strategy
Heterorhabditis Entomopathogenic nematode (Heterorhabditidae, Clade 9)	Gamma-proteobacteria *Photorhabdus*	Intestine	Insect pathogenicity, nutrition, growth and development	Vertical/maternal transmission
Steinernema Entomopathogenic nematode (Steinernematidae, Clade 10)	Gamma-proteobacteria *Xenorhabdus*	Intestinal receptacle	Insect pathogenicity, nutrition, growth and development	Horizontal transmission
Brugia Animal-parasitic nematodes (Onchocercidae, Clade 8)	Alpha-proteobacteria *Wolbachia*	Hypodermal cells of the lateral cords of both sexes and in the ovaries, oocytes and developing embryos within females	Normal development and fertility	Vertical/maternal transmission
Laxus oneistus Marine thiotrophic nematode (Desmodoridae, Clade 4)	Gamma-proteobacteria Sulfur-oxidizing bacteria	Whole cuticle of the nematode, except its anterior-most region	Nutrition	Likely horizontal

Table 2.2: Valuable characteristics of *Heterorhabditis* nematode and *Photorhabdus* bacteria as models of symbiosis

S.No.	**Valuable characteristics in a genetic model of symbiosis** (Ruby, 2008)	***Heterorhabditis-Photorhabdus* symbiosis**
1.	An inexpensive and easily collected or bred host	The nematode *Heterorhabditis* is an inexpensive host and can be multiplied *in-vitro* or *in-vivo* to produce a large number of nematodes.
2.	A small host with a simple morphology	*Heterorhabditis* nematodes are transparent and morphologically resembles well established model nematode *C. elegans.*
3.	Easily imaged symbiotic structures	*Heterorhabditis* nematodes harbour symbionts in their gut and it can be imaged easily.
4.	Known nutritional characteristics	*Heterorhabditis* can be easily cultured *in-vivo* or *in-vitro* on their symbiont bacterial lawn on nutrient growth media (NGM) supplemented with cholesterol.
5.	A range of ecologically or evolutionarily distinct associations	*Heterorhabditis-Photorhabdus* association has facilitated the ability of nematodes to exploit insects as food sources.
6.	An association that is economically important	Nematodes along with bacterial partners are pathogenic to a broad range of insect pests and are utilized as biocontrol agents in agriculture.
7.	An association that is representative of a general biological principle (or principles)	Highly specific association between *Heterorhabditis* and *Photorhabdus*. Sophisticated maternal transmission of symbiotic bacteria to achieve successful symbiosis generations after generations.
8.	A host and bacterium that can be cultured separately (grown in the laboratory)	*Heterorhabditis* and *Photorhabdus* can be cultured separately. It also possible to generate symbiont free axenic nematodes using standard procedures.
9.	A host and bacterium for which genome sequences are available	Genome sequences are available for both the symbiont and host partners.
10.	An association in which one or both partners are amenable to genetic manipulation	Genetic manipulations such as RNAi and mutation/genetic transformations have been standardised in *Heterorhabditis*. *Photorhabdus* is also amenable for genetic manipulations/transformations/mutations.

3.1 Nematode population and development of an inbred line *Heterorhabditis indica* for whole-genome sequencing

H. indica strain IARI-EPN-Hms1 used in this study was originally isolated from Nagpur, Maharashtra, India (ITS accession No. HQ637414) (Kumar et al., 2015). Nematode stocks of *Heterorhabditis* were maintained in the laboratory by infecting the last instar larvae of the greater wax moth, *Galleria mellonella* with infective juveniles (IJs) using standard procedures (Kaya and Stock, 1997). Freshly emerged IJs were collected in Ringer's solution (100 mM NaCl, 1.8 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$ and 5 mM HEPES, pH 7.0) using modified White's trap. The IJs were stored in tissue culture flasks by maintaining a nematode density of 5000 IJs per ml at 15 °C and used within 10 days.

Symbiotic bacteria of *H. indica* i.e., *Photorhabdus akhurstii* was isolated from infect juveniles. Briefly, symbiotic bacterial isolation was as follows: IJs were surface sterilized using 2% NaOCl (bleach) for 2 min and subsequently washed with sterile distilled water thrice to remove residues of NaOCl. Crushed homogenate of surface sterilized IJs was plated on NBTA (in one litre- nutrient broth 8g, agar 15g, bromothymol blue 0.25 g, TTC/triphenyltetrazolium chloride 0.04g) plates under aseptic condition and incubated at 28 °C for 24-48 hr. *Photorhabdus* produces greenish-brown colonies on NBTA. Fresh pure colony of bacteria was inoculated into modified Luria broth i.e., LBP media (in one litre - casein enzymic hydrolysate 10 g, yeast extract 5 g, sodium chloride 10 g, sodium pyruvate 1 g) to prepare starter cultures of *Photorhabdus*. Freezing stock of bacteria containing 5% of dimethyl sulfoxide (DMSO) was prepared and stored at -80°C. Starter culture (i.e., ~100-150 µl of overnight grown bacterial inoculum) was seeded on the NGM or nematode growth medium (in one litre - peptone 5 g, sodium chloride 5 g, meat extract B 1.50 g, yeast extract 1.50 g, agar 16 g) supplemented with 5 mg/ml cholesterol (NGM+Chol) in 60*15 mm plates and incubated for 24-48 h at 28°C to prepare the symbiotic bacterial lawns. Such symbiotic bacterial lawns were used for inbreeding and propagation of nematodes.

Development of the inbred line *H. indica* Hms1-i20: The authentic population of *H. indica* strain IARI-EPN-Hms1 was inbred for 20 generations to generate

genetically homogenous nematodes for whole-genome sequencing. Inbreeding involved self-fertilizing the hermaphrodites by placing a single L4 nematode onto the lawns of *P. akhurstii* on NGM+Chol plates and then inoculating new cultures with a single L4 nematode from the previous generation and this process was repeated 20 times (Figure 3.1). This inbred line is designated as *H. indica* Hms1-i20. Inbred nematodes were multiplied on NGM+Chol plates to produce a large number of worms.

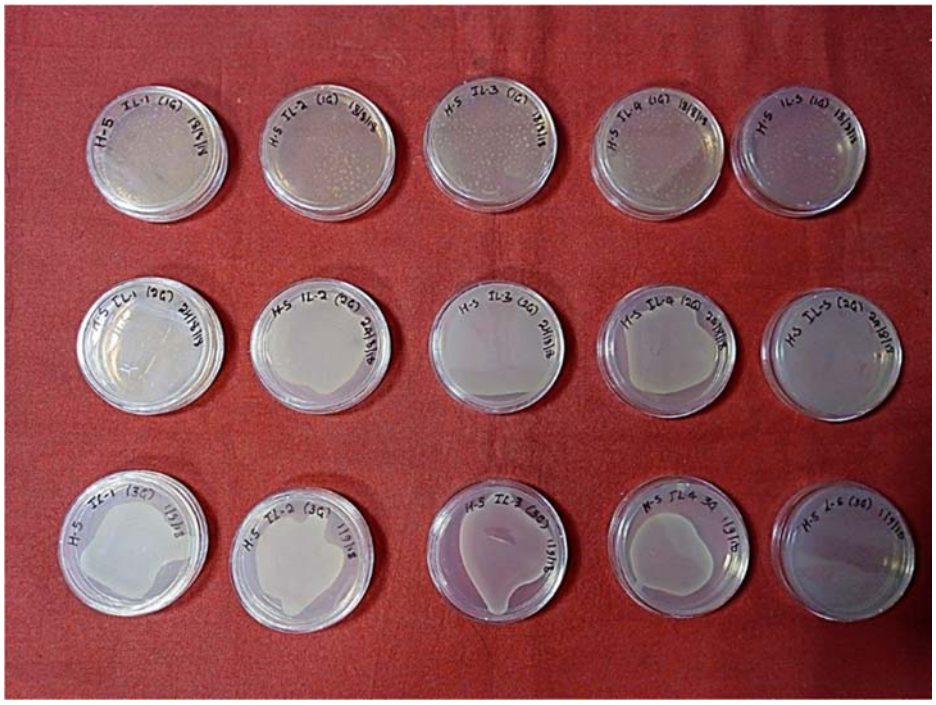


Figure 3.1: Development of *H. indica* inbred line. Self-fertilizing the hermaphrodite by placing a single L4 onto the lawns of *P. akhurstii* on NGM+Chol plates and then inoculating new cultures with a single L4 nematode.

3.2 Isolation of high molecular weight DNA from the inbred line, genomic library preparation and whole genome sequencing using Illumina platform

Inbred nematodes were lysed using Qiagen buffer G2 (Catalogue no. 19060, Qiagen, USA), and high molecular weight DNA was extracted using the phenol-chloroform method. Briefly, ~100 mg of surface sterilized worms were homogenized thoroughly using tissue homogenizer. Homogenate was suspended in 2 ml of digestion buffer G2 which contains RNase pre-added to a final concentration of 200 µg/ml. Proteinase K was added to a final concentration of 400 µg/ml and incubated at 55°C for 3-4 hours until the lysates turn almost clear. The lysate was centrifuged at 5000g for 10 minutes at room temperature and the supernatant was transferred to a clean tube. An

equal volume of phenol:chloroform: isoamyl alcohol (25:24:1, pH=8) was added to the lysate and gently mixed for 15-20 minutes. Then, centrifuged at 5000g for 15 minutes at 4 °C and the top aqueous phase was transferred to a new tube and phenol:chloroform: isoamyl alcohol extraction was repeated once again. The DNA was then precipitated with equal volume isopropanol by incubating at -20 °C for 30-60 minutes and centrifuged at 5000g for 15 minutes at 4°C. The resulting pellets were washed twice with 70% cold ethanol, air dried, dissolved in 50 µl of nuclease-free water.

DNA concentration and quality were estimated using Qubit 4.0 Fluorometer (Thermo Fisher Scientific, USA) and agarose gel electrophoresis (0.6% agarose gel, 120 min run time at 100 V). DNA was fragmented by sonication using Bioruptor (Diagenode, Seraing (Ougrée), Belgium). The size distribution was checked by running an aliquot of the fragmented DNA sample on an Agilent high sensitivity bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Illumina TruSeq DNA sample preparation guide (Illumina, San Diego, CA, USA) was followed to prepare two high-quality libraries of insert sizes 300 bp and 600 bp (Library name: HI-01_300 and HI-01_600). Illumina Nextera Mate Pair Library Prep Kit was used to prepare a mate-pair library of insert size 5 kb (HI-01_5kb). Paired-end reads of length 2*151, and 2*131 (Mate-pair) were generated on Illumina HiSeq 4000 platform.

3.3 Genome assembly, validation, gene prediction, annotation and further bioinformatic analysis

All the raw reads generated from sequencing of three genomic libraries were subjected to quality filtering using fastp version 0.22 to identify high quality (HQ) reads (Chen et al., 2018). The HQ reads were used to generate primary assembly using SOAPDenovo assembler Release 1.0 (Li et al., 2010). The contaminating mitochondrial and bacterial sequences of *Photorhabdus* symbiont were removed from the assembly using BlobTools v1.0.1 (Laetsch and Blaxter, 2017) and the NCBI server. Primary assembly contigs were further scaffolded using SSPACE 'Standard' version to generate the final draft assembly (Boetzer et al., 2011). Genome completeness assessment was done by BUSCO (Benchmarking Universal Single-Copy Orthologs) v5.2.2 against the 3,131 BUSCOs in the nematoda_odb10 database (Seppey et al., 2019) and using CEGMA (Core Eukaryotic Genes Mapping Approach) against 248 ultra-conserved core eukaryotic genes (Parra et al., 2007). Protein-coding genes were predicted in the

H. indica Hms1-i20 by SNAP, Augustus, GeneMark, and Maker (Besemer and Borodovsky, 2005; Cantarel et al., 2008; Korf, 2004; Stanke and Morgenstern, 2005). The NCBI Blastx+/ RefSeq/SwissProt/UniProt databases were used for annotation of the predicted gene list. Predicted genes were assigned Gene Ontology (GO) terms using Blast2GO (v1.3) and KEGG Pathway analysis was done using KAAS annotation server (Conesa et al., 2005; Moriya et al., 2007). The pipeline used to generate the final draft assembly, assembly validation, gene prediction and annotation is provided in figure 3.2.

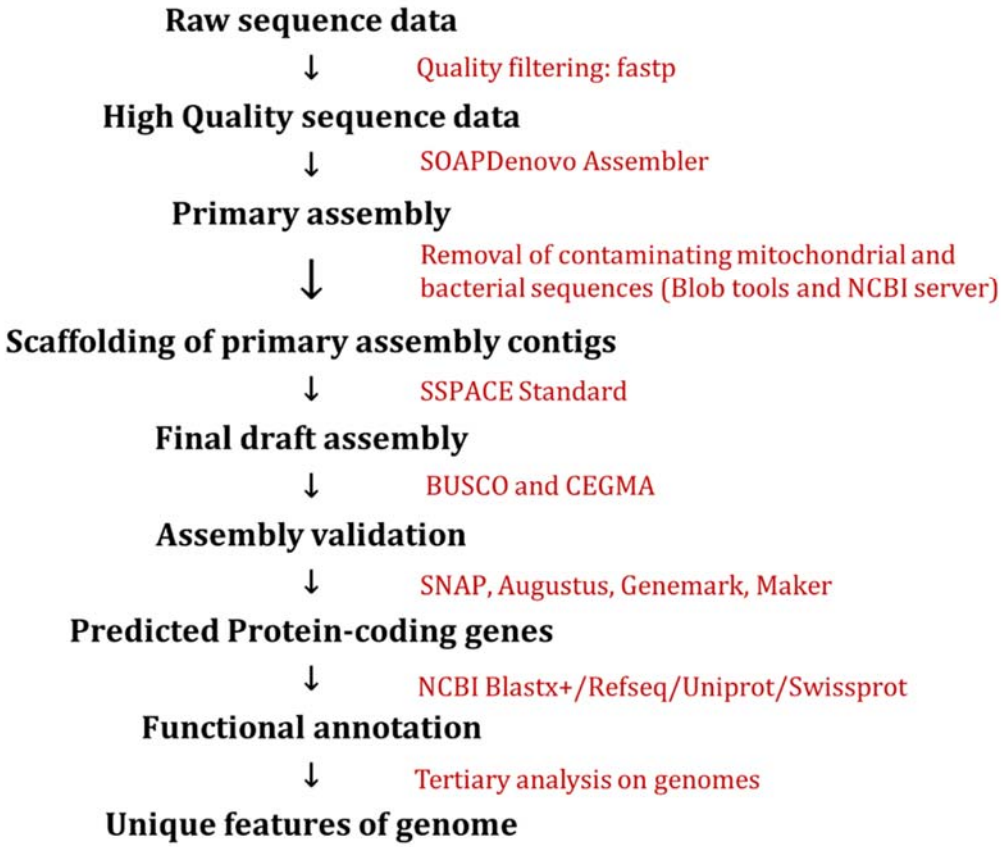


Figure 3.2: The diagrammatic representation of the pipeline used to generate the *H. indica* draft genome assembly from the Illumina high-quality reads, assembly validation, gene prediction and annotation.

The orthologous sequences among *H. indica, H. bacteriophora*, *C. elegans*, *S. carpocapsae*, and *O. tipulae* were identified using the OrthoMCL (v2.0.9) program on the predicted protein sequences from the genomes (Li et al., 2003). The nematode genomes used for comparison in this study were taken from Wormbase Parasite, namely *H. bacteriophora* (PRJNA13977), *S. carpocapsae* (PRJNA202318), *C. elegans* (PRJNA13758), and *O. tipulae* (PRJNA644888) genomes. Domain level enrichment and functional analysis of proteins were done using InterProScan 5.53-87.0 using pfam, ProSitePatterns, CDD, and SMART databases (Jones et al., 2014). Pairwise whole genome alignment between *H. indica* and *H. bacteriophora* was performed using synteny mapping and analysis program (SyMAP v5.05) (Soderlund et al., 2011).

TMHMM version 2.0 was used for prediction of proteins with transmembrane helix. TMHMM is a hidden Markov model-based membrane protein topology predictor. It accurately predicts transmembrane helices and distinguishes between soluble and membrane proteins (Krogh et al., 2001). A stringent pipeline was used to identify putative G-protein coupled receptors (GPCRs) in *H. indica* genome. The predicted proteome of *H. indica* was subjected to length-based filtering where only such sequences with more than 200 amino acid residues were retained as most of the GPCRs reported are within the range of 200-1500 amino acid residues (Takeda et al., 2002). Four different transmembrane predictors (TMHMM, Phobius, HMMTOP2 and TOPCONS) were employed to identify heptathetical transmembrane sequences (Bernsel et al., 2009; Käll et al., 2007; Krogh et al., 2001; Tusnády and Simon, 1998). Four different GPCR predictors (GPCRHMM, GPCRPipe, GPCRPred and GPCR-PenDB) were employed to identify putative GPCRs from heptathetical transmembrane sequences (Begum et al., 2020; Bhasin and Raghava, 2004; Theodoropoulou et al., 2013; Wistrand et al., 2006).

The secreted proteins were identified by using tools SignalP (v4.1), TMHMM (v2.0), Mitofates, PS SCAN, PredGPI, NucPred and Phobius (Brameier et al., 2007; Fukasawa et al., 2015; Gattiker et al., 2002; Käll et al., 2007; Krogh et al., 2001; Nielsen, 2017; Pierleoni et al., 2008). Predicted secreted proteins were subjected to MEROPS database search at cutoff of E value 0.0001 to identify secreted peptidases and peptidase inhibitors (Rawlings et al., 2018). Pipeline used to identify the *H. indica* secreted proteins are presented in the figure 3.3.

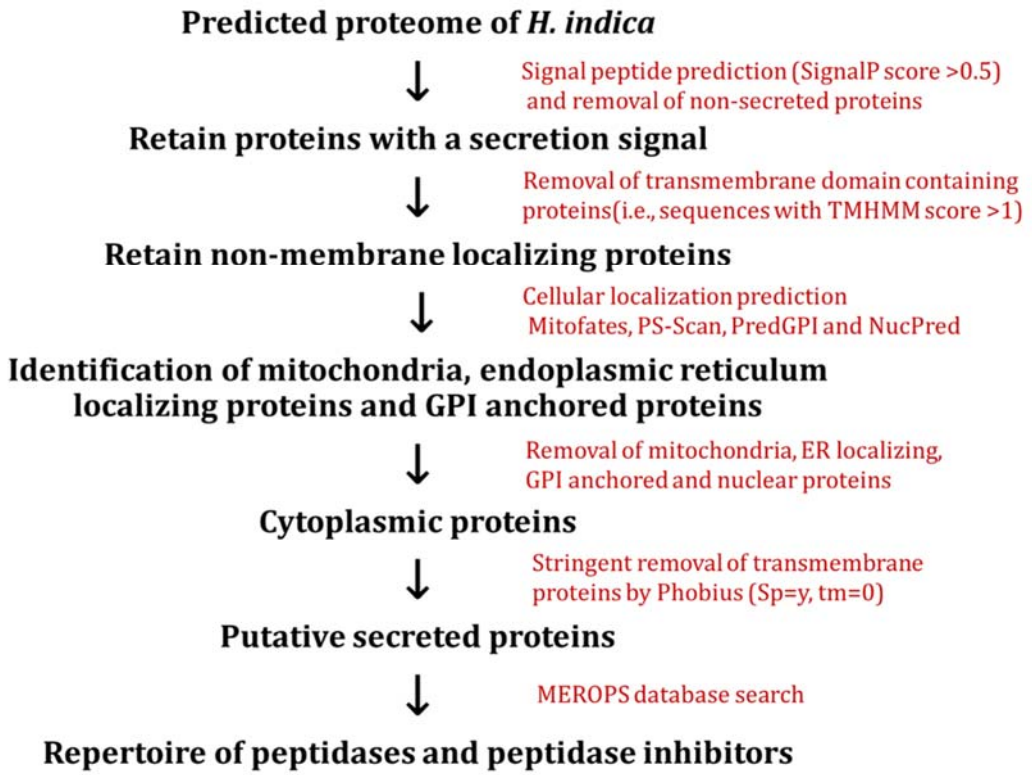


Figure 3.3: The diagrammatic representation of the pipeline and tools used for identification of *H. indica* secreted proteins.

MISA software was used to identify genomic SSR markers or microsatellite loci using default parameters (Beier et al., 2017). Transposon prediction was done with reasonaTE which is a part of the TransposonUltimate suite of tools (https://github.com/DerKevinRiehl/TransposonUltimate) (Riehl et al., 2021). The PHI-database (www.phi-base.org) Version 4.12 was used to characterize putative *H. indica* proteins involved in host-pathogen interactions (Urban et al., 2020). Pathogen-host interaction database (PHI-base) is a curated database of genes that have been shown to influence the phenotypic outcome of pathogen-host interactions. Version 4.12 of PHI-base (released in September 2021), contains data on 8,411 genes, 18,190 interactions, 279 pathogens, 228 hosts and 4,387 references.

The orthologs of *C. elegans* genes involved in the RNAi pathway in *H. indica* genome were identified by the reciprocal best hit method. Seventy-seven *C. elegans* proteins belonging to six functional groupings are known to be involved in core aspects of RNAi (Dalzell et al., 2011). Protein coding sequences of these 77 *C. elegans* genes

were retrieved from WormBase (www.wormbase.org; version WS277) and used as search strings against *H. indica* protein database in BLASTx analysis. *H. indica* proteins returning with a bit score of >45 and an expect value ≤ 0.0001 were subjected to reciprocal BLASTp against the *C. elegans* non-redundant protein databases on the NCBI BLAST server (http://www.ncbi.nlm.nih.gov/BLAST), using default settings. Retrieved *Heterorhabditis* proteins were analysed for conserved domains/domain structure using NCBI conserved domain search tool. The identity of the top-scoring reciprocal BLAST hit was accepted as the identity of the relevant primary hit, as long as that identity was also supported by domain structure analysis.

Infernal ("INFERence of RNA ALignment") tool allows searching DNA sequence databases for RNA structure and sequence similarities. This tool was used to identify tRNA and other non-coding RNA (ncRNA) genes (Nawrocki and Eddy, 2013). The Alienness tool (http://alienness.sophia.inra.fr) is a taxonomy-aware web tool that is useful in the identification of HGT of non-metazoan origin to a metazoan species (Rancurel et al., 2017). The BLAST results of the whole proteome of *H. indica* against NCBI non-redundant(nr) protein database were used as input to Alienness. Alienness calculates an Alien Index (AI) for each query protein. Putative HGT events were identified based on the stringent criteria of AI > 45. The mitochondrial genome was assembled separately from complete HQ reads using SPAdes assembler (Bankevich et al., 2012). The assembled genome was further annotated using MITOS (Bernt et al., 2013) and Mitfi (Jühling et al., 2012).

3.4 Nematode and bacterial strains used in transcriptomic studies

The nematode *H. bacteriophora* strain TTO1 used in this study was a gift from Dr. Byron J. Adams (Brigham Young University, USA) to (Late) Dr. (Mrs.) Sudershan Ganguly. Nematode stocks were maintained as explained in section 3.1. Bacterial strains used in this study were *P. luminescens* ssp. *laumondii* strain TTO1 (wild type symbiont of *H. bacteriophora* strain TTO1), *P. luminescens* ssp. *laumondii* strain TTO1GFP (TTO1 containing Tn7-GFP (Ciche et al., 2008)), and *P. luminescens* ssp. *laumondii* strain TTO1 *ΔmadA* (GFP labelled transmission defective mutant that is unable to colonize the nematodes (Somvanshi et al., 2010)). Freezing stocks of bacterial strains were stored at -80 C. LBP media was used for preparation of starter cultures of *Photorhabdus.*

3.4.1 Production of axenic nematodes

Heterorhabditis IJs carry *Photorhabdus* symbionts in their gut. Such worms are referred to as symbiotic or monoxenic IJs (Figure 3.4). To obtain the nematode eggs, symbiotic *H. bacteriophora* strain TTO1 IJs were cultured on lawns of wild-type *P. luminescens* ssp. *laumondii* strain TTO1. Around 400 surface-sterilized IJs were added to bacterial lawns and incubated at 28°C for 3-4 days. Gravid adult hermaphrodites were washed off the lawns with sterile distilled water and added to axenizing solution (2.4% (v/v) NaOCl, 0.25 N KOH). The eggs were washed and collected as described earlier (Stock and Goodrich-Blair, 2012). *Heterorhabditis-Photorhabdus* symbiosis is obligate in nature, however, it is possible to culture both the partner separately under the laboratory condition. These nematodes cannot be multiplied/cultured on any other bacteria such as *E. coli*. One of the reported methods to obtain axenic *Heterorhabditis* is to grow them on transmission defective symbiont bacterial strain (for example, Bai et al., 2013; Somvanshi et al., 2010). Here isolated nematode eggs were placed on the lawns of the transmission mutant *Photorhabdus* strain *ΔmadA* to generate symbiont free axenic IJs (Somvanshi et al., 2010). To confirm the sterility of axenic nematodes, the IJs were surface sterilized in 1% commercial bleach for 5 min, homogenized, and plated on Luria Bertani agar (in one litre - casein enzymic hydrolysate 10 g, yeast extract 5 g, sodium chloride 10 g, agar 10 g) as well as screened under the fluorescence microscope to confirm the absence of symbionts inside the gut (Figure 3.4). Antibiotics were added to maintain axenic nematodes stock at the following concentrations- streptomycin 100 mg/ml, ampicillin 100 mg/ml, kanamycin 30 mg/ml and gentamicin 10 mg/ml.

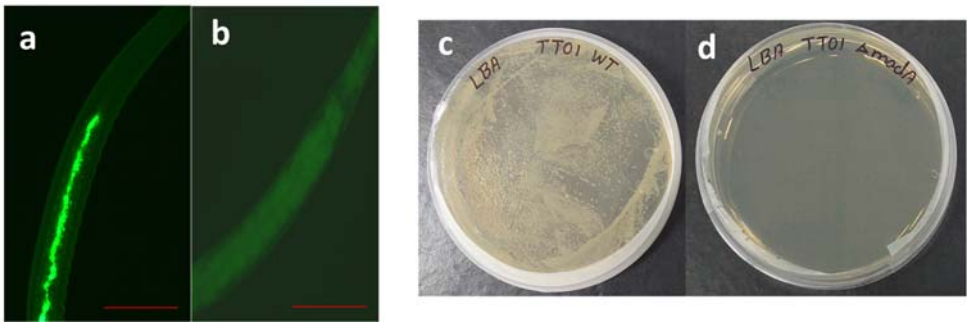


Figure 3.4: Monoxenic and axenic IJs of *H. bacteriophora.* a) Monoxenic *Heterorhabditis* IJ containing *Photorhabdus* symbionts inside the gut; b) Axenic *Heterorhabditis* IJ devoid of bacterial symbionts in the gut; c) *Photorhabdus* colonies on LBA plate from the homogenate of surface sterilized monoxenic IJs; d) No colonies on LBA plate from the homogenate of surface sterilized axenic IJs.

3.5 Preparation of nematodes for RNA-Sequencing experiment

Symbiont bacterial attachment and biofilm formation takes place in posterior nematode intestinal cells when the nematode is at the early adult stage, i.e., 36-40 h after IJ recovery. Pulse-chase experiment was performed to re-confirm the biofilm formation time point during nematode development as described earlier (Ciche et al., 2008). Briefly, nematodes were pulsed for 36-40 h with GFP-labelled symbionts and then starved for 4 h to eliminate transient cells in the nematode intestine before imaging. The presence of symbiotic bacterial biofilm (cells of persistent green colonizing bacteria) in the nematodes was determined by fluorescent microscopy using a ZEISS microscope (ZEISS Axio Imager 2 Research Upright Microscope from Carl Zeiss, Germany). The symbiotic (S) and axenic (A) IJs were cultured on the lawns of wild type *P. luminescens* ssp. *laumondii* and *ΔmadA* strain for 36-40 h, respectively. The nematodes were washed from the plates, filtered using 20-micron sieves to get rid of transient bacteria, collected in trizol and snap-frozen for RNA isolation. The experiments were repeated thrice.

3.6 RNA isolation, cDNA synthesis, library preparation and RNA-sequencing

Total RNA was extracted from a 100 µl pellet of symbiotic and axenic early adults of *H. bacteriophora* separately by TRIzol Plus RNA Purification Kit (Thermo Fisher Scientific, Massachusetts, USA) according to the manufacturer's instructions. RNA was treated with DNase I, amplification grade (Thermo Fisher Scientific, Massachusetts, USA) to remove any genomic DNA contamination. The integrity of the extracted RNA was tested on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and RNA with an RNA integrity number (RIN) of 8.0 was used for messenger RNA (mRNA) purification. The mRNA was purified from approximately 5 µg of intact total RNA using oligodT beads (Illumina® TruSeq® RNA Sample Preparation Kit v2). The purified mRNA was fragmented in the presence of bivalent cations and first strand cDNA was synthesized using Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA, USA) and random hexamer primers (Invitrogen, Carlsbad, CA, USA). Second strand cDNA was synthesized by DNA polymerase I and RNaseH following standard protocol (Illumina). The cDNA was cleaned using Agencourt AMPure XP purification kit (Beckman-Coulter, Brea, CA, USA), amplified, quantified using a Nanodrop spectrophotometer and checked for

quality with a Bioanalyzer. In total, 6 libraries were prepared for symbiotic and axenic nematode samples (3 each) as per the Illumina protocols. cDNA libraries were sequenced on Illumina HiSeq platform (150x2 Paired-End format) by outsourcing to Bionivid Technologies Pvt. Ltd., Bengaluru, India.

3.7 Bioinformatic analyses

H. bacteriophora genome available on Wormbase Parasite (BioProject ID: PRJNA13977, Bai et al., 2013) is fragmented (BUSCO score for assembly-Complete 87.1%, Fragmented 9.6%). In addition, gene prediction and annotation of this assembly is poor (BUSCO score for annotation – Complete-31.4%, Fragmented 14.9%). One other genome assembly published for *H. bacteriophora* strain G2a1223 (McLean et al., 2018) is not available in the public domain. Under these circumstances, we generated a *de-novo* transcriptome assembly for the present study. The bioinformatic analysis methods pipeline is provided in figure 3.5. All the Paired-end fastq reads were subjected to adapter trimming and quality filtering using fastp (v.0.20.0) (Chen et al., 2018) at a minimum length cutoff of 70 bp and phred quality score of 30. High Quality (HQ) reads from all the samples were pooled together to generate a *de-novo* primary assembly using Trinity Assembler (v.2.4.0) (Grabherr et al., 2011). Reads were *in-silico* normalized and the maximum read coverage was set at 80. A minimum count of Kmer for Inchworm assembly was set at 7 (parameters: --normalize_max_read_cov=80 and –min_kmer_cov=7). Nucleotide sequences (transcripts) which were similar in length and identity were clustered together using CD_HIT_EST tool (Li and Godzik, 2006) and only one representative sequence (unigene) per cluster was retained. The assembly was curated further using Bionivid's in-house amelioration pipeline Transimprove (Bankar et al., 2015) to get high quality transcripts supported by adequate depth (5x) and coverage (70%). To make sure that no contaminating sequences are used for downstream analysis, the sequences were screened using NCBI taxonomy server. Additionally, FASTQC screen against *Photorhabdus* genome was done to check for possible symbiont bacterial contamination (Andrews, 2010, Duchaud et al., 2003). Assembly validation was done by mapping the reads against the *de-novo* transcriptome using Kallisto (Bray et al., 2016). The completeness of transcriptome assembly was assessed by BUSCO against eukaryota_odb10.2019-11-20 and nematoda_odb10.2019-11-20 databases (Seppey et al., 2019).

The transcripts were quantified with the RSEM method (Li and Dewey, 2011) and the expression was normalized using fragment per kilobase million (FPKM) metric (Mortazavi et al., 2008). The expression data from each of the replicates were analysed using the GGally R package which utilizes Pearson correlation to determine the correlation between the replicates. Differential gene expression analysis was performed using DESeq package in R language (Anders & Huber, 2010). DESeq provides a method to test differential expression by use of the negative binomial distribution. Read counts taken from alignment bam files were taken as input, and differential analysis was performed by comparing the transcriptome of symbiotic nematodes to axenic nematodes. Transcripts with $\log_2$ fold change ≥ 2 and *p-value* ≤ 0.05 were considered as significantly up-regulated, while those with $\log_2$ fold change ≤ -2 and *p-value* ≤ 0.05 were treated as significantly down-regulated. Hereafter, the up-regulated or down-regulated transcripts refer to transcripts up- /down-regulated in symbiotic nematodes (treatment) as compared to the axenic nematodes (control). Sample-specific transcripts were identified based on the normalized expressions. A FPKM >1 cut-off was used to filter the false positives and identify the uniquely expressed transcripts in the symbiotic and axenic nematodes. The transcripts were annotated by blastx search against protein sequences from UniProt, SwissProt and NCBI RefSeq databases. Further, Trinotate (Trinotate/Trinotate.github.io), a tool for functional annotation of the transcriptome was also used to annotate the un-annotated transcripts. Filtration criteria used for blastx were: E-value ≤ 0.001, query coverage ≥ 60 and percentage identity ≥ 40. BLAST results were further processed using Blast2GO (v1.3) and KEGG Pathway analysis was done on KAAS annotation server (Conesa et al., 2005; Moriya et al., 2007). DAVID Bioinformatics Resources (v.6.7) was used for functional annotation and pathway enrichment analysis (Huang et al., 2007).

3.8 Validation of gene expression patterns by qRT-PCR

The expression pattern of 27 randomly selected transcripts was validated by quantitative real-time PCR (qRT-PCR). cDNA was prepared by reverse transcription of 500 ng of the RNA (same as used for RNA-sequencing) by Superscript VILO (Invitrogen, Carlsbad, CA, USA). qRT-PCR was performed on a Realplex2 thermal cycler equipment (Eppendorf, Hamburg, Germany) using SYBR Green Supermix Kit (Eurogentec, Liege, Belgium). Each reaction mixture contained 1.5 ng of cDNA, 750 nM of gene-specific forward and reverse primers, and 5 µl of SYBR Green PCR master

mix in a final volume of 10 µl. The amplification reactions were carried out in the thermal cycler with a hot start of 95°C for 5 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. Amplification specificity using melt curve analysis was performed as follows: 95 °C for 15 s, 60 °C for 15 s, followed by a slow gradient from 60 °C to 95 °C. A constitutively expressed nematode gene, 18S rRNA, was used as an internal reference. Three biological and three technical replicates were maintained for each sample. Data were analyzed by the ΔΔCt method (Livak and Schmittgen, 2001), and the results were expressed as log2-transformed fold change values. The oligonucleotide primers used for qRT-PCR are listed in table 3.1.

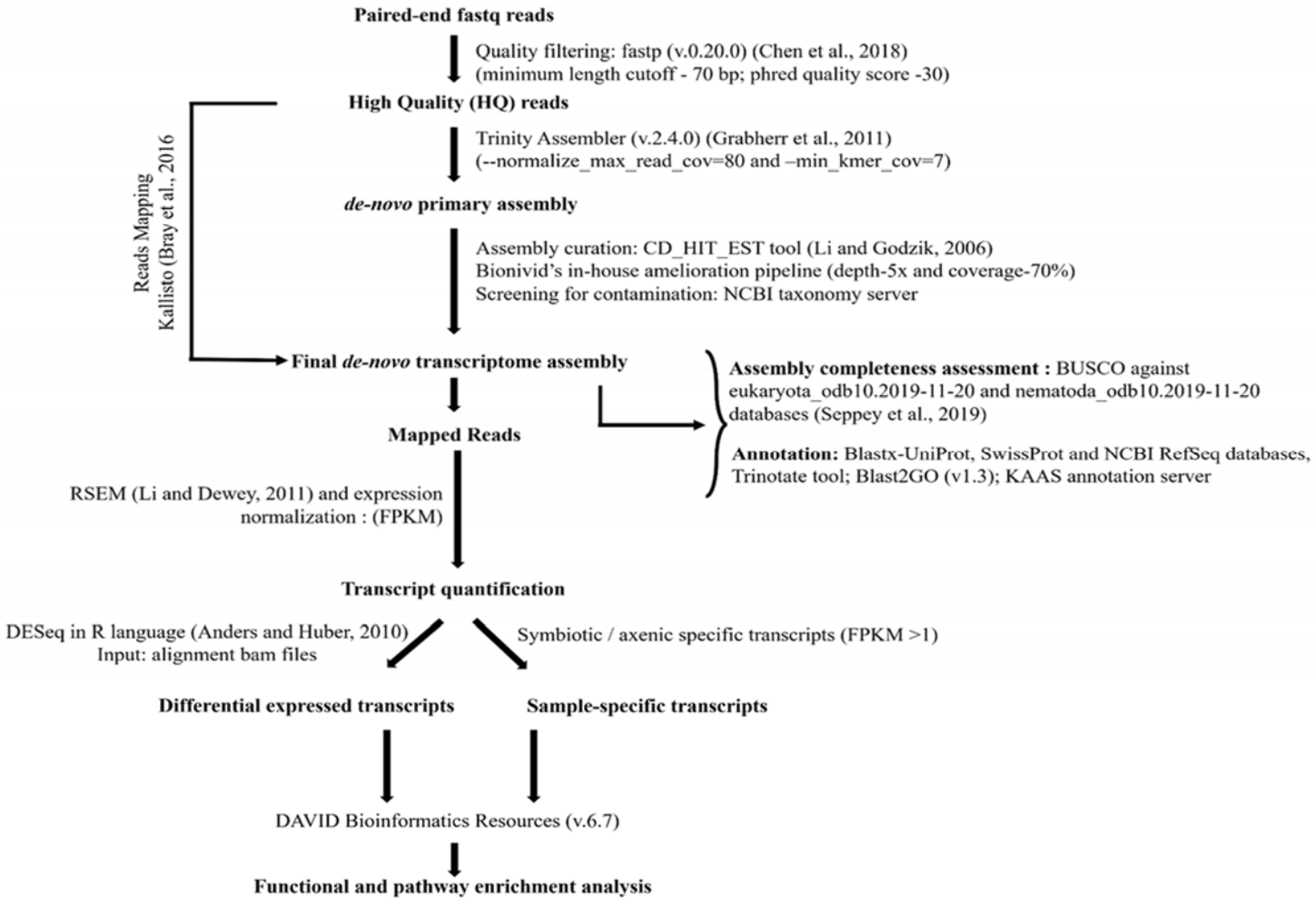


Figure 3.5: Bioinformatic analysis pipeline used to analyze the RNA-Seq data.

Table 3.1: List of primers used in the qRT-PCR validation of expression pattern of RNA-Seq data

S. No.	Gene annotation	Primer sequence (5'-3')	Product length (bp)	Tm °C	Transcript ID
1	*clec-87*	FP- CGACTGAGGAGCCAACTACC	181 bp	60	TRINITY_DN1416_c0_g2_i1--LEN=1071
		RP- TTCGCATAAACCGATGATGA			
2	*dpy-5*	FP- ACAAATGGCCGTTGATAAGG	153 bp	60	TRINITY_DN5772_c0_g1_i1--LEN=1068
		RP- CCTGTCCGGCTATCATTCAT			
3	*col-19*	FP- GCTTTGTGTTTCGTTGACCA	180 bp	60	TRINITY_DN7746_c0_g4_i1--LEN=1358
		RP- GGATTTCCAGGCCTACCTTC			
4	*mex-5*	FP- AGAAAGAAAACGGACCACGA	156 bp	60	TRINITY_DN5388_c0_g2_i2--LEN=2237
		RP- CACGCAGATCTTCAGTTCCA			
5	*plx-2*	FP- CGTGGGAGACTCTGGTCAAT	203 bp	60	TRINITY_DN4092_c0_g1_i1--LEN=4657
		RP- ATTTTTGCACTGTCCCCTTG			
6	*WDR26*	FP- GGTCTGTAAACCACCGCCTA	240 bp	60	TRINITY_DN7542_c0_g1_i2--LEN=1259
		RP- TTGGCAGCATTCAGTTCTTG			
7	*LRR20*	FP- CACTTTCCCTTGTTGGATGG	225 bp	60	TRINITY_DN6804_c0_g1_i6--LEN=3224
		RP- GGTTAACAGGAGGCACCGTA			
8	*IGSF9B*	FP- AGCGATGGACCATTCAGTTC	236 bp	60	TRINITY_DN6322_c0_g2_i1--LEN=2942
		RP- TTACGATGGGAGCCAGTACC			
9	*RPLP0*	FP- GGGAAGACAGATCTACCTGGAA	168 bp	60	TRINITY_DN3587_c0_g1_i1--LEN=1046
		RP- CCCCCTCCAAGTTCAAATC			
10	*COII*	FP- AGGTTACAGGGCATCAGTGG	197 bp	60	TRINITY_DN6295_c0_g10_i1--LEN=742
		RP- TCAAGCATGAATAACATCAGCA			
11	*ced-1*	FP- TCCAACAGCTGTCCTGATTG	169 bp	60	TRINITY_DN7259_c0_g1_i27--LEN=2471
		RP- CCCAGTAGCCCGGTATACAA			

12	*ced-3*	FP- CAAGGCACAGAAGTCGATCA RP- CATGGGTAAGAACCACGACA	175 bp	60	TRINITY_DN5119_c0_g4_i1--LEN=522
13	*ced-4*	FP- GATCTGCCACATTTCAAGCA RP- GCATCACGCGTAGTAGCAAG	170 bp	60	TRINITY_DN7359_c0_g1_i10--LEN=717
14	*LBP*	FP- GGATCAGCAACACGCAGATA RP- TCCCTTGCATAGTTGGCATA	189 bp	60	TRINITY_DN5483_c0_g1_i3--LEN=1784
15	*DMBT1*	FP- CACACCCATATGCCAGTGAT RP- TGTGATCCCGATAATTTTGC	172 bp	60	TRINITY_DN7168_c2_g5_i1--LEN=832
16	*daf-16*	FP- TCCTGATGGCAACAACACAT RP- TCGGCGTAGCTCATATTTCC	159 bp		TRINITY_DN7792_c0_g1_i15--LEN=2589
17	*tir-1*	FP- CTCCATTGGCTGGATACGAT RP- TTTCCATCATCTGTCGGTGA	148 bp		TRINITY_DN7733_c0_g4_i1--LEN=3597
18	*sma-3*	FP- AAACCTCTCGGAGTCGTCAA RP- GGCACGTTGTAGCATGTGAT	148 bp	60	TRINITY_DN6489_c0_g1_i4--LEN=1758
19	*sma-4*	FP- TGCCTAATGGTTGTGTCGAG RP- GCCATCCTTGGCTGTTAGAC	128 bp	60	TRINITY_DN8082_c0_g7_i2--LEN=1719
20	*daf-4*	FP- AAGAGGAAAGGCTGGAGAGG RP- AGCAGCCAATCCAAAATCAG	130 bp	60	TRINITY_DN6869_c0_g1_i2--LEN=2335
21	*gsnl-1*	FP- CGGCGGATATGATACTGGTT RP- ATATCGCGACCCAAATCAAG	168 bp	60	TRINITY_DN4551_c0_g2_i3--LEN=723
22	*daf-7*	FP- TTTATCGTCGCACCACTCAG RP- CAGCATTGCATAAGGTGCTC	152 bp	60	TRINITY_DN5392_c0_g2_i5--LEN=597
23	*ilys-3*	FP- TCTATCGCTACCGCTGACAA RP- TAATATCCTCGCCCGATTTG	172 bp	60	TRINITY_DN6977_c0_g2_i1--LEN=2116
24	*age-1*	FP- ACATGGGTGCCTCAGATCAT	120 bp	60	TRINITY_DN7607_c0_g1_i7--LEN=613

3.9 Standardization RNAi gene silencing in *Heterorhabditis* nematode

3.9.1 *In-silico* characterization of RNAi pathway in *Heterorhabditis*

RNAi pathway effector complements in the genome of *H. bacteriophora* (BioProject: PRJNA13977) and *H. indica* (PRJNA720543) were identified by reciprocal best hit method using *C. elegans* RNAi pathway genes as search strings.

3.9.2 Screening of *Heterorhabditis* developmental stages for dsRNA uptake using Fluorescein isothiocyanate (FITC) as marker of uptake

The objective was to determine the soaking efficiency in different developmental stages of *Heterorhabditis* i.e., maximal uptake with minimal lethality. Details of nematodes used and collection of different developmental stages were as follows: Nematode stocks of *H. bacteriophora, H. indica, S. abbasi* were reared individually *in vivo* in the laboratory using *Galleria mellonella* as a host. Live active IJs were collected in Ringer's solution and stored at 15°C and used within 10 days of collection. The IJs were plated on their respective symbiotic bacterial lawn on NGM+Chol plates to induce recovery and post IJ recovered stages were collected. L4 stage, early adult stages, and adult hermaphrodites were collected 10-12 h, 36-40 h and 108-116 h post-recovery respectively.

3.9.2.1 Soaking Protocol

FITC was used as a marker of uptake to show the uptake of the soaking solution, and thus dsRNA by nematodes. Fluorescent dye fluorescein isothiocyanate (FITC) is a non-toxic chemical marker having a molecular weight of 389.38. The stock solution of 5 mg/ml FITC (Sigma Co., St. Louis, Catalogue number: F7250) was prepared by dissolving FITC in dimethylformamide (DMF). Stock solutions were stored at 4°C and used within 1 week. M9 buffer (22 mM KH2PO4, 42.3 mM Na2HPO4, 85.6 mM NaCl, 1 mM MgSO4) was used as soaking solution. The final exposure concentration was 1 mg/ml FITC. Two soaking approaches were tested. In the first approach, only FITC was in the soaking solution. In a second approach, 50 mM octopamine was added to the soaking solution for neuro-stimulant-enhanced soaking. Approximately 20 mg nematodes were soaked in ~100 µl soaking solution for 16 h at 25°C in the dark and then washed three times with milliQ water. Fluorescence and viability of nematodes were analyzed immediately (0 h) and after 16 h, using a fluorescence microscope. Nematodes not exposed to FITC were included as negative controls for fluorescence

measurements. This experiment was repeated three times. The incorporation of FITC by *Steinernema* IJs was tested using the same concentration and methods as used for *Heterorhabditis* nematodes. To confirm the route of entry of FITC molecules into the nematode body, cephalic openings of *Heterorhabditis* IJs were blocked using cyanoacrylate tissue adhesive and then screened for the FITC uptake using the protocol published for *Heterodera glycine* (Schroeder and MacGuidwin, 2010)

Desiccation prior to soaking was tested in the case of *Heterorhabditis* IJs to see if it enhanced FITC uptake using the method described earlier for *Panagrolaimus* nematodes (Seybold et al., 2016). Briefly, nematode IJs were transferred to a microscope slide, the surface water was removed, and the slides were placed in a sealed desiccator containing saturated potassium sulphate solution. This induces slow dehydration of worms by lowering relative humidity to ~ 98%. Nematodes were incubated for 16 h at 25°C and then rehydrated with soaking buffer containing 1 mg/ml FITC.

3.10 Testing the efficacy of RNAi by soaking PIJR stage in dsRNA solution

3.10.1 Amplification and cloning of target genes

Genes *Hb-dpy-7*, *Hb-dpy-13* and *Hb-cct-2* which have proven RNAi phenotype in *H. bacteriophora* nematode were selected for validation of RNAi-soaking technique in the PIJR stage. Total RNA was isolated from PIJR nematodes using TRIzol Plus RNA Purification Kit (Thermo Fisher Scientific, USA, Catalogue no. 12183018A) as per the manufacturer's protocol. RNA was subjected to cDNA preparation using SuperScript® III Reverse Transcriptase (SuperScript® VILO™ cDNA Synthesis Kit catalogue No.11754-050) as mentioned in the previous section 3.8. Synthesized complementary DNA was used as the template for PCR amplification of target genes. Details of the gene primers are provided in table 3.2. Details of the PCR reaction mixture and the PCR cycling conditions followed are given in table 3.3. The amplified PCR products were visualized on 1.2% agarose gel for purity and integrity. PCR products were purified by column purification using Wizard® SV Gel and PCR Clean-Up System (Promega Corporation, USA, Catalogue no. A9281) following the manufacturer's instructions. The concentration of purified PCR products was measured using NanoDrop2000/2000c (Thermo Scientific, USA).

Table 3.2: List of primers used in RNAi experiments

Gene name	**Forward primer**	**Reverse Primer**
Hb-dpy-13	AGCCCGGAGCTAAAGGTAAC	TACGAGTCATCAATGGCACA
Hb-dpy-7	GGTAGACCAGGTCGTCCAGT	ACCAGGCAAACCAGGACTT
Hb-cct-2	AACAGACCAATCGTCCTTGG	CTTGGATGTGGTAGCCGTTT
Hb-18s	CAGCCAAAGAGGATGGAGAA	CCTCCGAGAACAAGTGCAAG
M13	GTAAAACGACGGCCAG	CAGGAAACAGCTATGAC

Table 3.3: Details of PCR reaction mixture and PCR cycling conditions used for amplification of target genes from cDNA of post-IJ recovery stage of *H. bacteriophora*

Components		**Volume(µl)**	**Final concentration**
10X PCR buffer		5 µl	1X
10mM dNTPs		1 µl	0.2mM
50mM MgCl2		1.5 µl	1.5mM
Forward primer		1 µl	0.2 µm
Reverse primer		1 µl	0.2 µm
Template DNA		2 µl	100 ng
Platinum Taq DNA polymerase		0.2 µl	1.0 unit
DNAase free water to the final volume		50 µl	
PCR cycling conditions			
PCR steps	**Temperature**	**Time**	**No. of cycles**
Initial denaturation	94°C	5 min	1
Denaturation	94°C	45 sec	40
Annealing	57 °C	45 sec	
Extension	72 °C	45 sec	
Final extension	72 °C	10 min	1

The purified PCR products were cloned into the pGEM-T easy vector (pGEM®-T Easy Vector Systems - Promega Corporation, USA, Catalogue No A1360). Vector to insert ratio was maintained at 1:3 based on the size of the pGEM-T vector i.e. 3015 bp. The ligation reaction was set up by adding specific amplicons and other components provided with the kit in a 500 µl micro-centrifuge tube. The reaction mixture consists of 5 µl of 2X rapid ligation buffer, 1 µl of pGEM-T Easy vector (50ng), 1 µl of purified PCR product (100ng), 1 µl of T4 DNA ligase and nuclease-free water to a final volume of 10 µl. The reaction mixture was mixed well by short spin and incubated for ligation at 4°C overnight. The competent cells of *Escherschia coli* strain DH5α were prepared by $CaCl_2$ method and stored at -80 C for future use (Mandel and Higa, 1970). The transformation was carried out by the heat shock method (Froger and Hall, 2007). Briefly, 3-5µl of the ligated product was added to the 100µl competent cells, mixed gently and subjected to heat shock in the water bath at 42ºC for 90 seconds. After that immediately transferred to ice and chilled for 3-5minutes. 1ml of LB was added to the cells and incubated at 37ºC in a shaker incubator for 90 minutes with continuous shaking at 180 rpm. Cells in LB were plated aseptically on Luria Bertani agar plate containing IPTG, X-gal and ampicillin (100 ml LBA contained 100 µl ampicillin (100 mg/ml), 200 µl X-gal (20 mg/ml) and 100 µl IPTG (0.1M stock i.e. 0.238g in 10ml). The plates were incubated overnight at 37ºC. The selection of positive transformants containing the inserts was done based on blue/white colony screening. The positive white colonies were picked and streaked onto a LBA + ampicillin plate and incubated overnight at 37ºC. Target gene insertion was confirmed by colony PCR using gene-specific primers. The confirmed positive recombinant colonies were inoculated in LB with ampicillin at 37°C and the overnight grown fresh cultures were used for preparing plasmid. The plasmid isolation was done using FavorPrep Plasmid Extraction kit (Catalogue Number: FAPDE 100) as per the manufacturer's protocol. Cloned gene target was validated by sequencing using universal M13 forward and reverse primers (Table 3.2).

3.10.2 Double-stranded RNA (dsRNA) preparation and soaking

Plasmids containing confirmed target genes were used to prepare the DNA template for in-vitro transcription. The gene sequences were PCR amplified from recombinant plasmids using M13 forward and reverse primers using the amplification conditions of initial denaturation: 94 ºC for 4 min, 35 cycles of denaturation 94 ºC for

1 min, annealing 57 ºC for 45 s, extension 72 ºC for 1 min and final extension at 72 ºC for 10 min. The target gene in the pGEMT Easy vector is flanked by T7 and SP6 promoters. Single strand RNA (ssRNA) synthesis reaction was carried out with T7 and SP6 MEGA script RNAi kit (MEGAScript SP6, Catalogue number AM1330, MEGAScript T7, catalogue number, Invitrogen, USA). The reaction mixture using T7 and SP6 were incubated separately at 37 ºC for 20-23 h followed by addition of 30 ul of lithium chloride and nuclease-free water each and incubated at -20 ºC for 3 h. After incubation, the microcentrifuge tubes were centrifuged at 10000g for 20 min at 4 ºC. Pellet of ssRNA was washed with chilled 70% ethanol and dissolved in 20 µl of nuclease-free water. The two ssRNA stands were mixed in equal concentration and incubated at 65 ºC for 10 min followed by 37 ºC for 30 min. The quantity of dsRNA was determined by Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, USA) and analyzed by agarose gel electrophoresis. dsRNA was stored at -80 ºC for future use. PIJR nematodes were collected, washed with sterile M9 buffer thrice. Nematodes were soaked in 200 µl of the soaking solution containing 1mg/ml dsRNA for 24 h at 25 ºC in dark using 24 multi-well plates for invitro RNAi. Nematodes incubated in M9 buffer were used as control. Post incubation nematodes were washed with sterile M9 buffer and plated on the symbiotic bacterial lawn for further growth and phenotyping. A subset of treated worms was used for RNA isolation. Expression of the target gene in the dsRNA treated and control worms was assessed using qRT-PCR.

4.1 Whole-genome sequencing of nematode *Heterorhabditis indica*

4.1.1 Developing an inbred line of *Heterorhabditis indica*, high molecular weight DNA extraction, genomic library construction, whole-genome sequencing using Illumina platform and draft genome assembly

The *H. indica* strain IARI-EPN-Hms1 was inbred for 20 generations to obtain genetically homogenous nematodes for genome sequencing. The inbred line is designated as *H. indica* Hms1-i20. High molecular weight DNA extracted from the inbred nematodes was used for the preparation of genomic libraries. A total of three high-quality genomic libraries were prepared. This included two libraries of insert sizes 300 bp and 600 bp (designated as HI-01_300, HI-01_600), and a mate-pair library of insert size 5 kb (HI-01_5KB). Paired-end reads of length 2*151, and 2*131 (mate-pair) were generated on Illumina HiSeq 4000 platform.

A total of 22.74 Gb sequence data comprising of ~160 million reads were generated. The reads provided 750X coverage of the *H. indica* genome. Quality filtering of raw reads identified 155 million high quality (HQ) reads. The raw and filtered read statistics for *H. indica* are presented in table 4.1. The raw sequence data has been deposited in GenBank under BioProject No. PRJNA720543, BioSample No. SAMN18671197 and SRA IDs SRR14181568, SRR14181569 and SRR14181570. The HQ reads were assembled. The final *H. indica* genome assembly was of 91.3 Mb size having 3,538 scaffolds with an N50 of 587 kb and an average scaffold length of 25.79 kb. The minimum and maximum scaffold length was 501 bp and 2,409 Kb, respectively. The GC content of the assembled genome was 35.31%, and there were 7.33% N's in the assembly. Assembly statistics for the *H. indica* Hms1-i20 are provided in table 4.2. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAJAVP000000000. GenBank assembly accession is GCA_020740425.1 (https://www.ncbi.nlm.nih.gov/assembly/ GCA_ 020740425.1)

Table 4.1: Read statistics for *H. indica* Hms1-i20 genome

	Raw Reads			**Filtered high quality (HQ) Reads**			
Genomic library (Insert size)	**Total reads (Million)**	**Total bases (Giga bases)**	**GC content (%)**	**Total HQ reads (Million)**	**Total HQ bases (Giga bases)**	**GC content (%)**	**% HQ reads**
HI-01_300 (300 bp)	48.37	7.30	40.92	47.15	7.02	40.62	97.47
HI-01_600 (600 bp)	39.71	5.99	41.12	38.08	5.67	40.89	95.89
HI-01_5KB (Mate pair library, 5 Kb)	72.27	9.44	39.65	69.75	8.77	38.83	96.51

Table 4.2. Assembly statistics of *H. indica* Hms1-i20 genome at contig and scaffold level

Assembly statistics	**Contig Level**	**Scaffold Level**
Total sequences	92,372	3,538
Total bases (bp)	95,526,713	91,263,125
Min sequence length (bp)	88	500
Max sequence length (bp)	435,793	2,409,384
Average sequence length (bp)	1,034.15	25,795.12
Median sequence length (bp)	145	3410.50
N50 length (bp)	33,662	587,367
As (%)	30.74	28.71
Ts (%)	30.5	28.64
Gs (%)	19.28	17.67
Cs (%)	19.48	17.64
(A + T)s (%)	61.24	57.35
(G + C)s (%)	38.76	35.31
Ns (%)	0	7.33

Genome completeness assessment was done by Benchmarking Universal Single-Copy Orthologs (BUSCO) v5.2.2 (Seppey et al., 2019) against the 3,131 BUSCOs in the nematoda_odb10 database as well as against the 303 BUSCO markers in the eukaryota_odb9 database. The *H. indica* Hms1-i20 genome showed the presence of 2,630 (84%) complete BUSCOs (complete and single-copy- 2617 (83.6%); complete and duplicated- 13 (0.4 %)), whereas 205 (6.5%) BUSCOs were fragmented, and 296 (9.5 %) were missing. Out of 303 conserved eukaryote genes, 266 genes (87.78%) were fully present, 4.95% were partially mapped and 7.26% were missing. CEGMA analysis considering 248 ultra-conserved core eukaryotic genes showed 97.17% completeness of the genome considering partial genes and 95.96% for complete genes. Assembly completeness assessment statistics are presented in table 4.3. Anchoring of *H. indica* genome scaffolds to that of *H. bacteriophora* by synteny mapping showed that the share of the collinear region between two genomes is 35.37% (Figure 4.1).

Table 4.3: Genome assembly completeness of *H. indica* Hms1-i20 by BUSCO and CEGMA

BUSCO analysis	**Dataset: nematoda_odb9**		**Dataset: eukaryota_odb9**	
	Number	**%**	**Number**	**%**
Complete BUSCOs (C)	2,630	84	266	87
Complete and single-copy BUSCOs (S)	2,617	83.6	NA	NA
Complete and duplicated BUSCOs (D)	13	0.4	NA	NA
Fragmented BUSCOs (F)	205	6.5	15	4.9
Missing BUSCOs (M)	296	9.5	22	7.2
Total BUSCO groups searched	3,131		303	
CEGMA analysis	**Number**	**%**		
Complete	238	95.96		
Partial	241	97.17		
Missing	10	4.03		
Total groups searched	248			

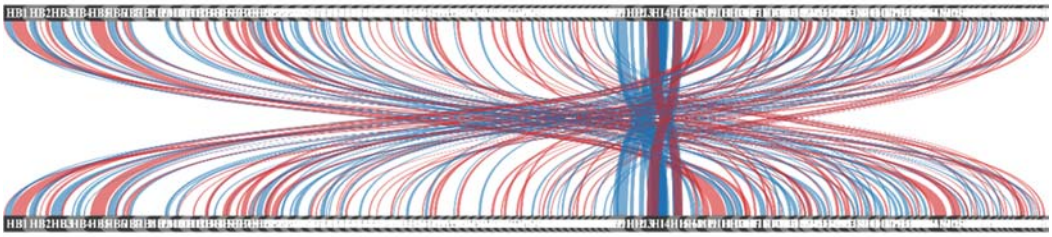


Figure 4.1: A comparison of *H. indica* Hms1-i20 genome assembly to *H. bacteriophora* TTO1-M31e by synteny analysis. Anchoring of *H. indica* genome scaffolds to that of *H. bacteriophora* showed that 35.37% of the scaffold regions were collinear between two genomes.

4.1.2 Gene prediction and annotation

A total of 10,494 protein-coding genes were predicted in the *H. indica* Hms1-i20 draft genome, out of which 9,596 (91.44%) could be annotated using various databases. 8,893 genes found a hit in the RefSeq database, whereas 9,596 genes had UniProt IDs. Out of the total predicted 10,494 protein-coding genes, 7,631 genes were annotated by gene ontology (GO) analysis. Functional classification of *H. indica* genes using eukaryotic cluster of orthologs (KOG) database could assign a function to 5,847 genes. The GO analysis showed that 'cell', ‘cell part’ and ‘membrane’ were most abundant terms in the cellular component category, 'binding ' and ‘catalytic activity’ in molecular functions category, whereas ‘cellular process’ and ‘metabolic process’ were most abundant in biological process category (Figure 4.2). The KOG analysis revealed that the three most plentiful gene classes were class T (signal transduction), class O (post-translational modification, protein turnover, chaperones) and class K (transcription). In contrast, class M (cell wall/membrane/envelope biogenesis) and class N (cell motility) were the least represented classes (Figure 4.3). KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis could annotate 2,090 genes mapping to 191 pathways, where protein processing in endoplasmic reticulum (PATH:ko04141), endocytosis (PATH:ko04144) and lysosome (PATH:ko04142) terms were most abundant. The top 20 enriched KEGG pathway terms are represented in figure 4.4.

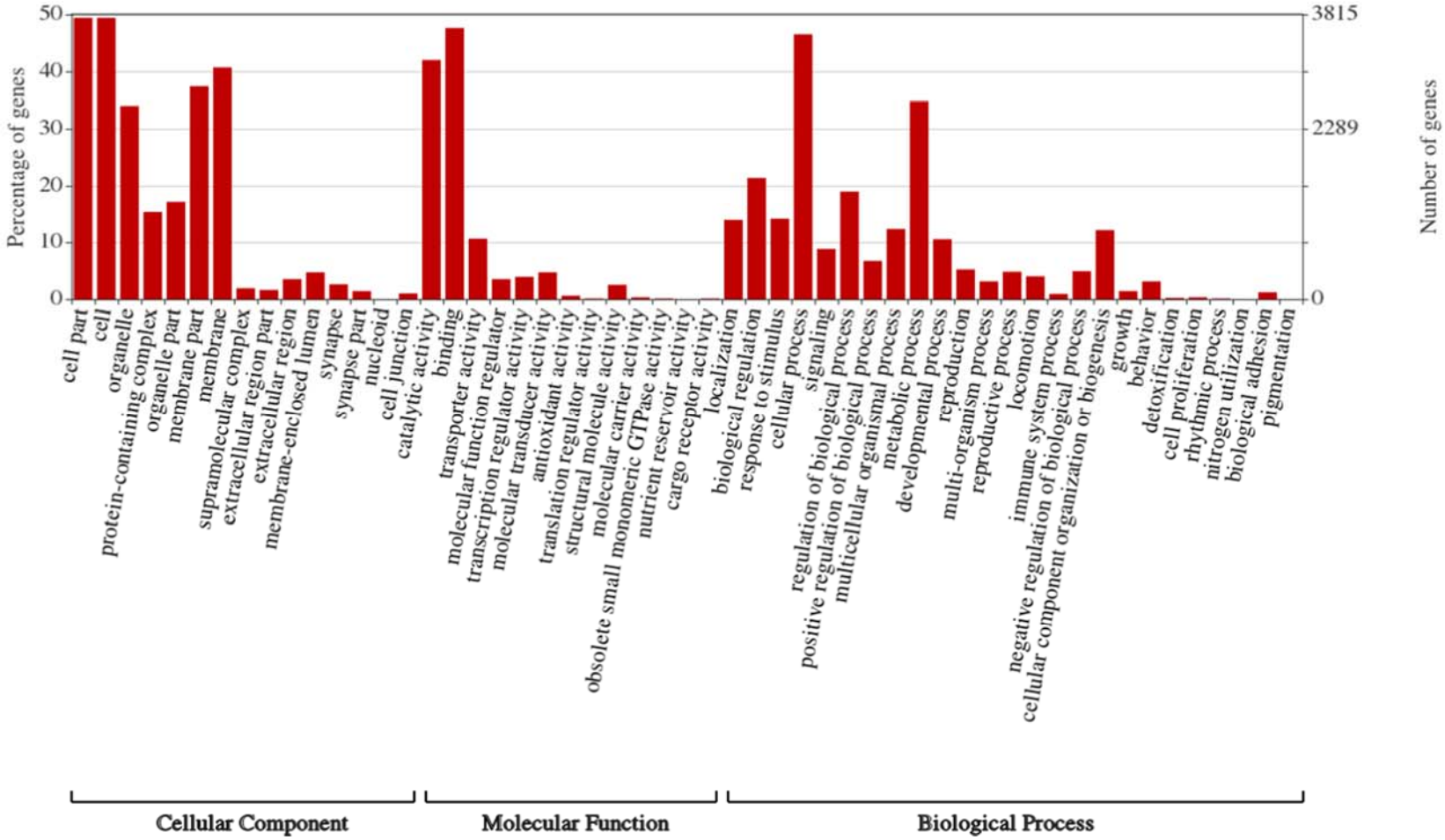


Figure 4.2: The annotation of *H. indica* genome by Gene Ontology (GO) analysis. Web Gene Ontology (WEGO) annotation plot representing the number of genes present under various GO categories of cellular component, molecular function and biological process.

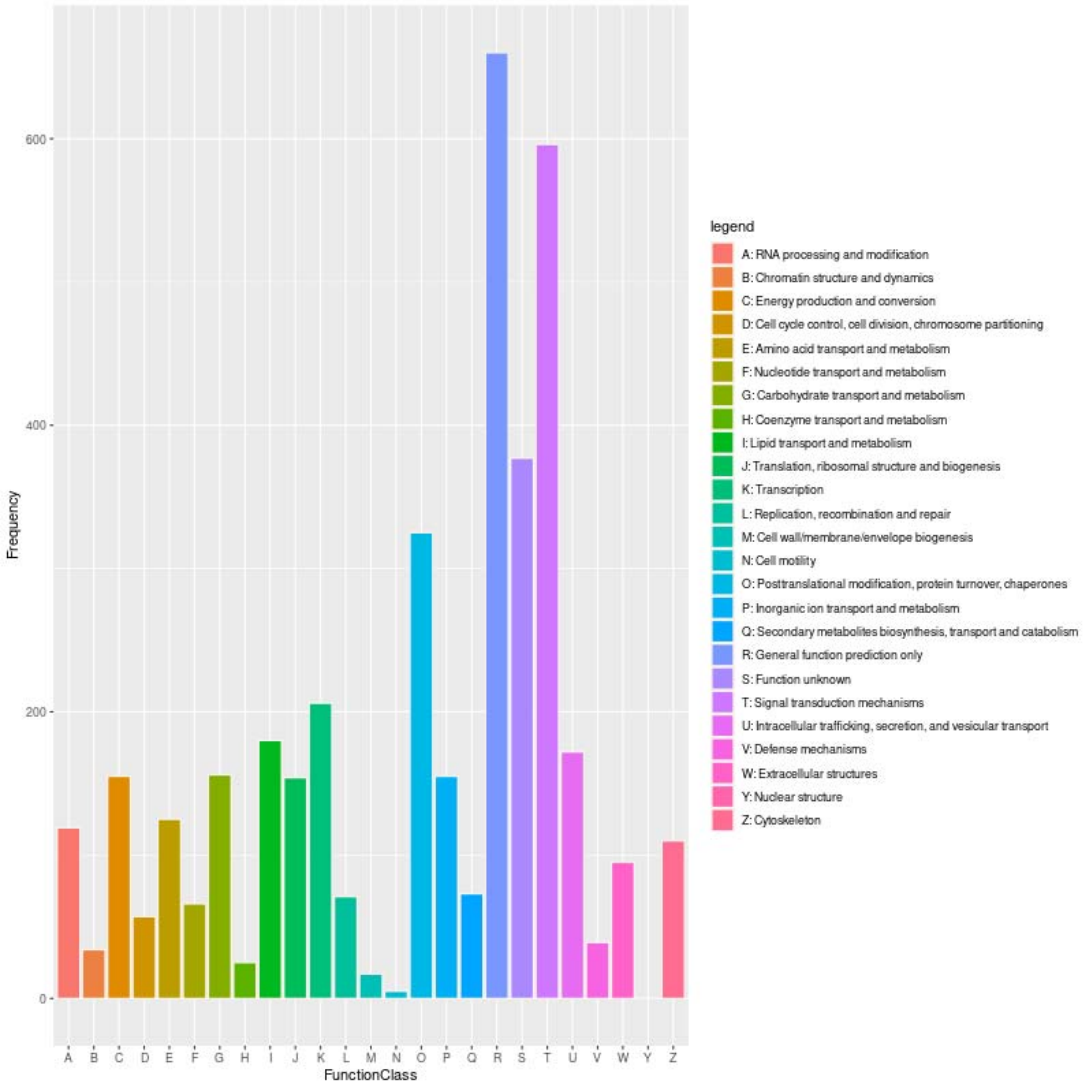
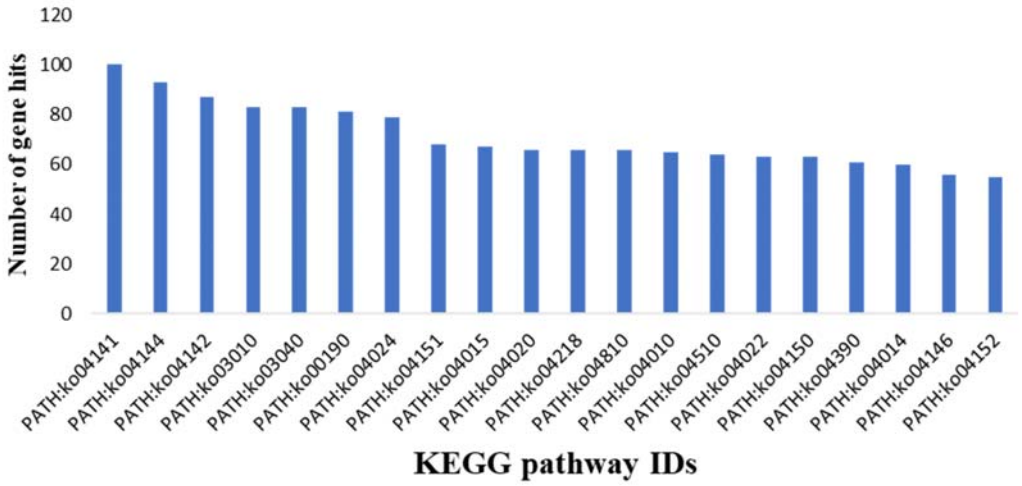


Figure 4.3: The annotation of *H. indica* genome by eukaryotic cluster of orthologous groups (KOG) analysis. The plot represents various KOG functional classes present in the *H. indica* genome.



KEGG ID	Description
PATH:ko04141	Protein processing in endoplasmic reticulum
PATH:ko04144	Endocytosis
PATH:ko04142	Lysosome
PATH:ko03010	Ribosome
PATH:ko03040	Spliceosome
PATH:ko00190	Oxidative phosphorylation
PATH:ko04024	cAMP signaling pathway
PATH:ko04151	PI3K-Akt signaling pathway
PATH:ko04015	Rap1signaling pathway
PATH:ko04020	Calcium signaling pathway
PATH:ko04218	Cellular senescence
PATH:ko04810	Regulation of actin cytoskeleton
PATH:ko04010	MAPK signaling pathway
PATH:ko04510	Focal adhesion
PATH:ko04022	cGMP-PKG signaling pathway
PATH:ko04150	mTOR signaling pathway
PATH:ko04390	Hippo signaling pathway
PATH:ko04014	Ras signaling pathway
PATH:ko04146	Peroxisome
PATH:ko04152	AMPK signaling pathway

Figure 4.4: Top 20 enriched KEGG pathway terms in *H. indica* genome.

4.1.3 Orthologs

The orthologous sequences among *H. indica*, *H. bacteriophora*, *C. elegans*, *Steinernema carpocapsae*, and *Oscheius tipulae* were identified using the OrthoMCL program on the predicted protein sequences from the genomes (Figure 4.5). A total of

8,641 orthologous groups were detected in the *H. indica* genome as compared to five other *nematode* genomes. *H. indica* shared 6,574 orthogroups with *H. bacteriophora*, 6,635 orthogroups with *C. elegans*, 6,228 with *S. carpocapsae* and 6,669 with O. *tipulae*. A total of 4,824 orthogroups were common between these five nematodes analyzed. *H. indica* shared the highest number of orthologous groups with *H. bacteriophora* (435), followed by *C. elegans* (55) and *S. carpocapsae* (49). It shared the lowest number of orthologous groups with *O. tipulae* (43).

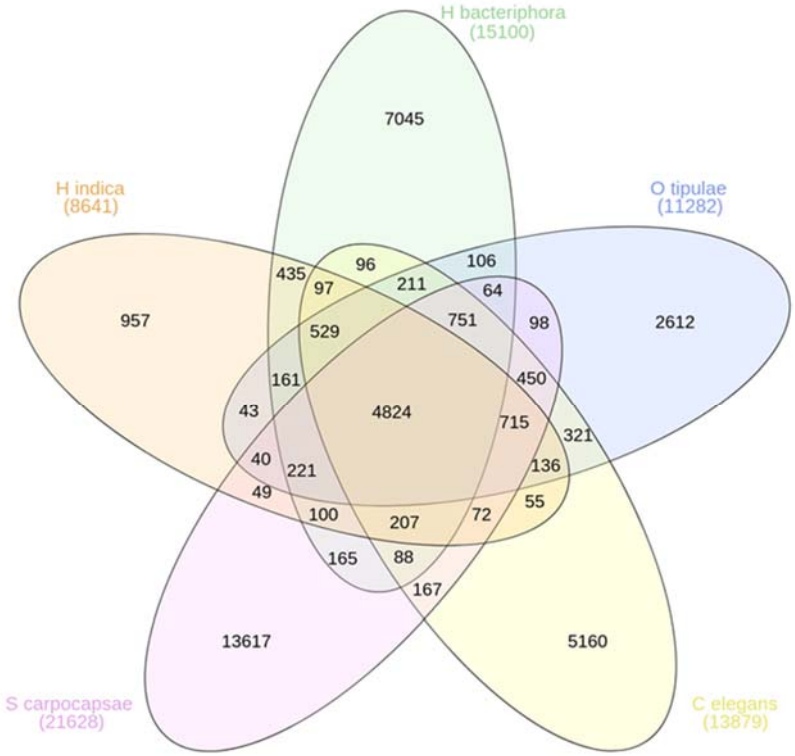


Figure 4.5: The orthologous gene groups identified in the genomes of *H. indica*, *H. bacteriophora*, *C. elegans*, *S. carpocapsae* and *O. tipulae* by OrthoMCL analysis.

4.1.4 Enriched protein domains in *Heterorhabditis indica* genome

Protein domain analysis using InterProScan 5.53-87.0 using pfam, ProSitePatterns, CDD, and SMART databases showed presence of as diverse as >6900 protein domains in *H. indica*. Zinc finger C2H2-type, WD40 repeat, EGF-like domain, ankyrin repeat, and fibronectin type III were the top enriched protein families. The 20 most abundant domains with their protein family details are provided in table 4.4. The protein domains enriched in four other nematodes compared to *H. indica* are

represented in Figure 4.6. As compared to *H. indica*, various protein domains were present in higher abundance in *C. elegans, S. carpocapase*, and *O. tipulae*, whereas the domain counts in *H. bacteriophora* was lesser. Specifically, domains of immunoglobulin subtype and immunoglobulin I-set families were highly enriched in *C. elegans* as compared to *H. indica*.

Table 4.4: Top 20 most abundant protein domains in *H. indica*

S. No.	Interproscan ID	No. of domains detected	Domain Description	Family
1.	IPR013087	854	c2h2final6	Zinc finger C2H2-type
2.	IPR001680	825	WD domain, G-beta repeat	WD40 repeat
3.	IPR000742	674	egf_5	EGF-like domain
4.	IPR002110	538	ANK_2a	Ankyrin repeat
5.	IPR003961	442	Fibronectin type III domain	Fibronectin type III
6.	IPR000719	370	Protein kinase domain	Protein kinase domain
7.	IPR001881	284	Calcium-binding EGF domain	EGF-like calcium-binding domain
8.	IPR003591	279	LRR_typ_2	Leucine-rich repeat, typical subtype
9.	IPR003599	274	IG_3c	Immunoglobulin subtype
10.	IPR002049	264	EGF_Lam	Laminin-type EGF domain
11.	IPR000504	241	RNA recognition motif	RNA recognition motif domain
12.	IPR003598	230	igc2_5	Immunoglobulin subtype 2
13.	IPR002223	226	KU	Pancreatic trypsin inhibitor Kunitz domain
14.	IPR003593	226	AAA_5	AAA+ ATPase domain
15.	IPR002048	220	EF hand	EF-hand domain
16.	IPR019734	216	Tetratricopeptide repeat	Tetratricopeptide repeat
17.	IPR013098	203	Immunoglobulin I-set domain	Immunoglobulin I-set
18.	IPR003439	189	ABC transporter	ABC transporter-like, ATP-binding domain
19.	IPR001356	176	homeodomain	Homeobox domain
20.	IPR006150	170	WR1_2	Cysteine-rich repeat

A scan for the trans-membrane (TM) helices in predicted protein models identified a total of 9,162 TM helices in 2,525 proteins in the *H. indica* genome. A single TM domain was detected in 901 proteins, whereas up to 27 TM domains were detected in one protein. The incidence of TM domains in various predicted proteins is provided in table 4.5.

Table 4.5: The presence of transmembrane helices in predicted protein models of *H. indica* genome

Number of predicted transmembrane helices	**Number of protein models**
1	901
2	381
3	276
4	211
5	176
6	148
7	106
8	87
9	52
10	71
11	46
12	37
13	17
14	6
15	2
16	1
17	3
20	1
21	1
23	1
27	1
Total	2,525

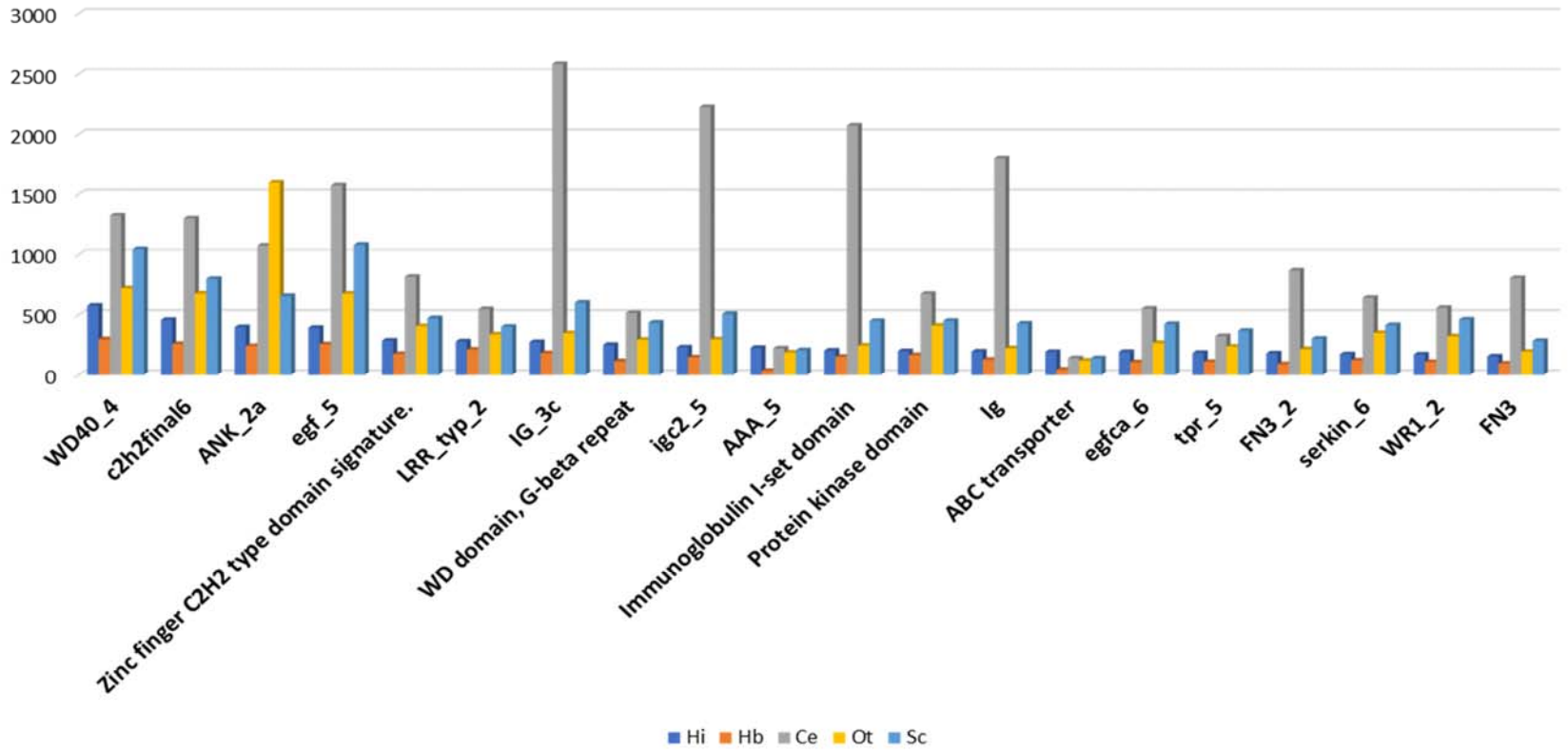


Figure 4.6: Comparison of top 20 protein domains in *H. indica* genome with those in the genome of *H. bacteriophora*, *C. elegans*, *S. carpocapsae* and *O. tipulae*. (Hi- *H. indica*, Hb- *H. bacteriophora*, Ce-*C. elegans*, Ot- *O. tipulae*, Sc- *S. carpocapsae*)

H. indica predicted proteome was scanned to detect G-protein-coupled receptors (GPCRs) which are the largest and most diverse group of membrane receptors in eukaryotes. Predicted protein sequences were subjected to length-based filtering (>200 amino acid residues). A total of 6,790 sequences had more than 200 amino acid residues. Out of 6,790 proteins, 68 sequences were confirmed as heptathetical transmembrane sequences by at least three out of four transmembrane predictors (TMHMM, Phobius, HMMTOP2 and TOPCONS) used in this study. A total of 56 gene sequences were identified as putative GPCRs using four different GPCR predictors, where GPCRHMM, GPCRPipe, GPCRPred and GPCR-PenD detected 22, 27, 55 and 18 sequences as GPCRs, respectively (Figure 4.7). Details of the putative GPCRs which were predicted by at least three out of the four tools, as putative GPCRs are presented in table 4.6. The majority of the predicted GPCRs were belonging to Class A or rhodopsin-like class. Two GPCRs from the frizzled class, one from each of the glutamate-like and secretin-like classes, were discovered.

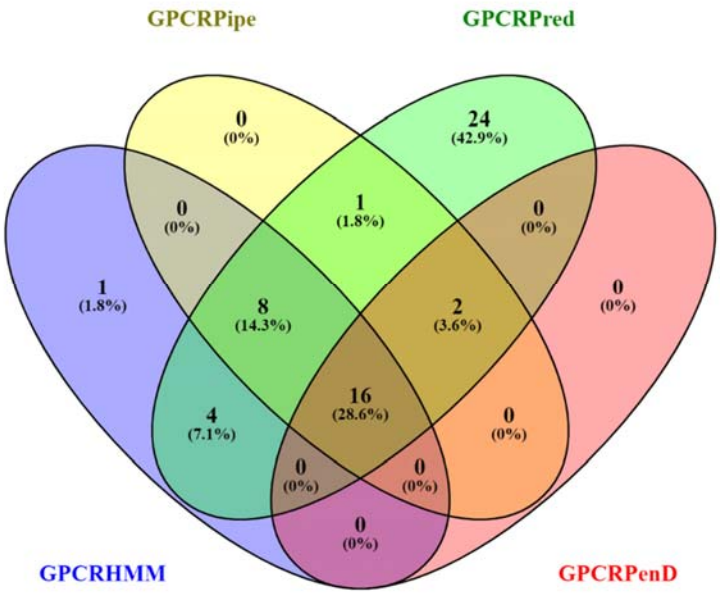


Figure 4.7: Number of putative GPCRs identified in the *H. indica* genome using four different GPCR predictors. GPCRHMM, GPCRPipe, GPCRPred and GPCR-PenD identified 22, 27, 55 and 18 sequences as GPCRs, respectively.

Table 4.6: GPCRs predicted in the *H. indica* genome using a combination of GPCRPipe, GPCRHMM, GPCRPenD, and GPCRPred tools

S.No.	Protein	Length (AA)	GPCRPred Subclass	Blast Match	Blast Evalue	Pfam Family	Pfam Score
	GPCRPred Class:A GPCR tm+7						
1.	HIND-IARI-GENE-6595	313	Lysospingolipid	-	-	7TM_GPCR_Srd	1.90E-88
2.	HIND-IARI-GENE-10050	808	Peptide	Glutamate-like	2.00E-162	-	-
3.	HIND-IARI-GENE-9776	501	Rhodopsin	Frizzled	4.00E-104	-	-
4.	HIND-IARI-GENE-8817	325	Olfactory	-	-	7TM_GPCR_Srx	6.50E-15
5.	HIND-IARI-GENE-6218	352	Peptide	Rhodopsin-like	2.00E-167	7tm_1	3.00E-42
6.	HIND-IARI-GENE-4006	497	Amine	Rhodopsin-like	2.00E-32	7tm_1	2.00E-37
7.	HIND-IARI-GENE-3008	520	Peptide	Rhodopsin-like	6.00E-12	7TM_GPCR_Srw	5.20E-32
8.	HIND-IARI-GENE-2412	366	Amine	Rhodopsin-like	1.00E-13	7tm_1	2.40E-21
9.	HIND-IARI-GENE-8642	412	Peptide	Rhodopsin-like	2.00E-14	7TM_GPCR_Srw	1.10E-57
10.	HIND-IARI-GENE-8582	731	Amine	Rhodopsin-like	5.00E-30	7tm_1	2.10E-61
11.	HIND-IARI-GENE-7501	439	Peptide	Rhodopsin-like	9.00E-32	7tm_1	1.90E-19
12.	HIND-IARI-GENE-5120	488	Amine	Rhodopsin-like	7.00E-64	7tm_1	3.50E-58
13.	HIND-IARI-GENE-1708	533	Peptide	Rhodopsin-like	2.00E-37	7tm_1	3.70E-39
14.	HIND-IARI-GENE-1327	474	Peptide	Rhodopsin-like	5.00E-59	7tm_1	3.10E-15

15.	HIND-IARI-GENE-9481	403	Peptide	Rhodopsin-like	0	7tm_1	1.00E-27
16.	HIND-IARI-GENE-9966	368	Peptide	Rhodopsin-like	2.00E-67	7tm_1	8.90E-13
17.	HIND-IARI-GENE-10489	307	Lysospingolipid	Rhodopsin-like	4.00E-20	7tm_1	3.50E-21
18.	HIND-IARI-GENE-9867	397	Peptide	Rhodopsin-like	3.00E-12	7TM_GPCR_Srw	2.20E-35
19.	HIND-IARI-GENE-6991	414	Peptide	Rhodopsin-like	9.00E-75	7tm_1	9.70E-45
20.	HIND-IARI-GENE-3827	599	Amine	Glutamate-like	1.00E-78	7tm_3	2.70E-42
21.	HIND-IARI-GENE-348	681	Amine	Rhodopsin-like	2.00E-34	7tm_1	1.50E-22
22.	HIND-IARI-GENE-7838	310	Peptide	-	-	7TM_GPCR_Srw	2.40E-15
23.	HIND-IARI-GENE-6408	334	Peptide	-	-	-	-
24.	HIND-IARI-GENE-596	560	Amine	Frizzled	2.00E-150	-	-
25.	HIND-IARI-GENE-439	353	Nucleotide Like	-	-	Srg	5.20E-21
	GPCRPred Class: B						
26.	HIND-IARI-GENE-6994	470	-	Secretin-like	2.00E-39	7tm_2	3.50E-46

4.1.5 Characterization of *Heterorhabditis indica* secretome

A total of 654 proteins containing a secretion signal were detected in the predicted proteome of *H. indica*. Out of these 654 proteins, 199 were found to contain one or more transmembrane (TM) helix domains. Cellular localization analysis identified 36 proteins localizing to mitochondria, endoplasmic reticulum, nucleus and glycophosphatidylinositol (GPI) anchored proteins (Table 4.7). A final list of 370 putative secreted proteins was obtained. Among the 370 proteins, 339 had significant similarity to the proteins in the Uniprot database. Among annotated proteins, 122 proteins matched 'uncharacterized proteins and/or hypothetical proteins'. Two genes i.e. HIND-IARI-GENE-6694 and HIND-IARI-GENE-1861 encoding nematode FAR (fatty acid retinoid-binding) proteins were detected in *H. indica* secretome. FAR proteins are candidate effectors characterized in nematodes of diverse trophic groups. A search of the MEROPS database containing peptidases and peptidase inhibitors revealed the presence of 38 peptidases (1 aspartic, 2 threonine, 6 serine, 7 cysteine, 18 metallo peptidases, and 4 peptidase of unknown catalytic type) and 31 peptidase inhibitors in *H. indica* secreted proteins (Table 4.8-4.9).

Table 4.7: Predicted *H. indica* proteins localizing to mitochondria, endoplasmic reticulum, nucleus and glycophosphatidylinositol (GPI) anchored proteins

S. No	**Endoplasmic reticulum targeting proteins- Prosite Scan (PS00014)**	**S. No.**	**Nuclear proteins with NLS >0.8**
1.	HIND-IARI-GENE-1718	20.	HIND-IARI-GENE-2244
2.	HIND-IARI-GENE-1951	21.	HIND-IARI-GENE-3855
3.	HIND-IARI-GENE-4653	22.	HIND-IARI-GENE-110
4.	HIND-IARI-GENE-5327	23.	HIND-IARI-GENE-838
5.	HIND-IARI-GENE-5987	24.	HIND-IARI-GENE-8203
6.	HIND-IARI-GENE-6497	25.	HIND-IARI-GENE-2420
7.	HIND-IARI-GENE-6767	26.	HIND-IARI-GENE-3736
8.	HIND-IARI-GENE-6894	27.	HIND-IARI-GENE-4645
9.	HIND-IARI-GENE-7542	28.	HIND-IARI-GENE-6693
10.	HIND-IARI-GENE-9768	29.	HIND-IARI-GENE-3958

	GPI anchored proteins -NetGpi1.1	30.	HIND-IARI-GENE-4057
11.	HIND-IARI-GENE-6831	31.	HIND-IARI-GENE-4750
12.	HIND-IARI-GENE-6126	32.	HIND-IARI-GENE-5242
13.	HIND-IARI-GENE-7708	33.	HIND-IARI-GENE-7761
14.	HIND-IARI-GENE-8743		**Mitochondria targeting proteins -mitofates**
15.	HIND-IARI-GENE-1564		
16.	HIND-IARI-GENE-593	34.	HIND-IARI-GENE-9820
17.	HIND-IARI-GENE-3990	35.	HIND-IARI-GENE-2840
18.	HIND-IARI-GENE-2609	36.	HIND-IARI-GENE-10001
19.	HIND-IARI-GENE-2896		

4.1.6 Pathogen-host interaction database (PHI-base)

The pathogen-host interaction database (PHI-base) is a curated database of genes that have been shown to influence the phenotypic outcome of pathogen-host interactions. Version 4.12 of PHI-base (released in September 2021), contains data on 8,411 genes, 18,190 interactions, 279 pathogens, 228 hosts and 4,387 references. Search against PHI-base identified 2,139 *H. indica* proteins likely to be involved in host-pathogen interactions. The larger portion of genes matched with fungal proteins curated in the database (1,354) followed by the bacterium (717) and nematodes (29). Of these 2,139 genes, 143 genes matched with pathogen proteins interacting with insects. A total of 820 and 1,078 genes matched with proteins involved in interaction with vertebrates and plants respectively. The majority of these genes (1,219) are involved in interactions leading to 'reduced virulence' phenotypic outcome on the host, followed by 525 proteins related to unaffected_pathogenicity. Up to 193 proteins suggested to be involved in loss_of_pathogenicity, and 95 proteins suggested to have a lethal phenotype were identified in *H. indica*. The summary of pathogen-host interaction analysis on *H. indica* genome is provided in table 4.10.

Table 4.8: List of secreted peptidases identified in *H. indica* genome

S.No.	QueryName	Identifier	Description	E-value
	Aspartic Peptidases			
1.	HIND-IARI-GENE-6125	MER1020334	asp-6 g.p. (*Caenorhabditis elegans*)	9.4e-166
	Cysteine Peptidases			
2.	HIND-IARI-GENE-24	MER0697836	family C15 unassigned peptidases	1.4E-58
3.	HIND-IARI-GENE-1335	MER0459293	glutamate synthase (NADPH) large chain	7.8e-216
4.	HIND-IARI-GENE-1234	MER0725955	qua-1 g.p. (*Caenorhabditis elegans*)	9.1E-94
5.	HIND-IARI-GENE-1098	MER0694370	subfamily C1A unassigned peptidases	3.4e-178
6.	HIND-IARI-GENE-2970	MER0694370	subfamily C1A unassigned peptidases	7.2e-163
7.	HIND-IARI-GENE-3710	MER1104934	subfamily C1A unassigned peptidases	1.4e-149
8.	HIND-IARI-GENE-5284	MER1105699	subfamily C1A unassigned peptidases	2.5E-78
	Metallo Peptidases			
9.	HIND-IARI-GENE-10243	MER0831264	family M13 unassigned peptidases	5.8E-08
10.	HIND-IARI-GENE-4379	MER0827005	nas-10 g.p. (*Caenorhabditis elegans*)	6.8E-41
11.	HIND-IARI-GENE-671	MER0826778	nas-32 g.p. (*Caenorhabditis elegans*)	4.6E-21
12.	HIND-IARI-GENE-3182	MER0826736	nas-32 g.p. (*Caenorhabditis elegans*)	5.7E-67
13.	HIND-IARI-GENE-9976	MER0826736	nas-32 g.p. (*Caenorhabditis elegans*)	1.1E-53
14.	HIND-IARI-GENE-7572	MER0002287	nas-38 g.p. (*Caenorhabditis elegans*)	1.0E-98
15.	HIND-IARI-GENE-1896	MER1127465	NEP-1 peptidase (*Caenorhabditis* sp.)	6.6e-264
16.	HIND-IARI-GENE-743	MER1104997	subfamily M12A unassigned peptidases	3.0e-110
17.	HIND-IARI-GENE-3875	MER1096136	subfamily M12A unassigned peptidases	7.8E-83
18.	HIND-IARI-GENE-6405	MER1120940	subfamily M12A unassigned peptidases	9.0e-113
19.	HIND-IARI-GENE-6838	MER1210997	subfamily M12A unassigned peptidases	8.0e-116

20.	HIND-IARI-GENE-7365	MER1104779	subfamily M12A unassigned peptidases	3.5E-64
21.	HIND-IARI-GENE-6270	MER0832952	subfamily M14A unassigned peptidases	1.5e-202
22.	HIND-IARI-GENE-8366	MER1095494	subfamily M14A unassigned peptidases	4.6e-205
23.	HIND-IARI-GENE-2298	MER0037361	subfamily M23B non-peptidase homologues	8.8E-25
24.	HIND-IARI-GENE-10185	MER0834401	Y47G6A.19 g.p. (*Caenorhabditis elegans*)	6.0e-190
25.	HIND-IARI-GENE-5446	MER1104361	Zmp-2 peptidase (*Caenorhabditis* sp.)	5.5E-98
26.	HIND-IARI-GENE-7020	MER0002295	Zmp-6 (*Caenorhabditis elegans*)	6.1E-75
	Serine Peptidases			
27.	HIND-IARI-GENE-8266	MER1007712	family S12 non- peptidase homologues	6.0E-07
28.	HIND-IARI-GENE-10271	MER1007712	family S12 non- peptidase homologues	1.0E-05
29.	HIND-IARI-GENE-4874	MER1363039	family S33 non- peptidase homologues	6.5E-47
30.	HIND-IARI-GENE-4333	MER1105322	family S9 unassigned peptidases	5.1e-243
31.	HIND-IARI-GENE-4327	MER0952407	subfamily S1A non- peptidase homologues	2.6E-36
32.	HIND-IARI-GENE-6689	MER1105092	subfamily S9C unassigned peptidases	1.9e-134
	Threonine Peptidases			
33.	HIND-IARI-GENE-9806	MER1094606	*Chryseobacterium meningosepticum*- type N4-(beta-N- acetylglucosaminyl)- L-asparaginase	3.0e-170
34.	HIND-IARI-GENE-2448	MER1073402	family T3 unassigned peptidases	2.3e-186
	Peptidase of unknown catalytic type			
35.	HIND-IARI-GENE-3408	MER0414865	family U69 unassigned peptidases	2.9E-30
36.	HIND-IARI-GENE-3469	MER0270090	family U69 unassigned peptidases	3.5E-38
37.	HIND-IARI-GENE-6361	MER0270095	family U69 unassigned peptidases	7.0E-39
38.	HIND-IARI-GENE-8406	MER0418452	family U69 unassigned peptidases	8.7E-15

Table 4.9: List of secreted peptidase inhibitors identified in *H. indica* genome

S.No.	QueryName	Identifier	Description	E-value
1.	HIND-IARI-GENE-4130	MER0023726	aspin-2	1.1E-93
2.	HIND-IARI-GENE-5291	MER1105050	cystatin Bm-CPI-1 (*Brugia malayi*)	2.5E-32
3.	HIND-IARI-GENE-3273	MER1127789	family I1 unassigned peptidase inhibitors	1.6E-37
4.	HIND-IARI-GENE-5014	MER1144738	family I15 unassigned peptidase inhibitors	1.0E-31
5.	HIND-IARI-GENE-622	MER1108296	family I17 unassigned peptidase inhibitors	1.2E-22
6.	HIND-IARI-GENE-1411	MER1121360	family I2 unassigned peptidase inhibitors	2.5E-26
7.	HIND-IARI-GENE-1544	MER0752581	family I2 unassigned peptidase inhibitors	6.8E-05
8.	HIND-IARI-GENE-6940	MER1104296	family I2 unassigned peptidase inhibitors	3.2E-21
9.	HIND-IARI-GENE-8638	MER1103376	family I2 unassigned peptidase inhibitors	2.3E-28
10.	HIND-IARI-GENE-5360	MER1128240	family I20 unassigned peptidase inhibitors	8.4E-21
11.	HIND-IARI-GENE-3130	MER0774358	family I35 unassigned peptidase inhibitors	2.4E-57
12.	HIND-IARI-GENE-10222	MER0774422	family I39 unassigned peptidase inhibitors	0.0E+00
13.	HIND-IARI-GENE-1315	MER0781466	family I43 unassigned peptidase inhibitors	4.1E-45
14.	HIND-IARI-GENE-2605	MER0149221	family I43 unassigned peptidase inhibitors	1.0E-12
15.	HIND-IARI-GENE-5548	MER0201916	family I43 unassigned peptidase inhibitors	1.9E-36
16.	HIND-IARI-GENE-8665	MER0781382	family I43 unassigned peptidase inhibitors	6.4E-44
17.	HIND-IARI-GENE-193	MER0917613	family I63 unassigned peptidase inhibitors	3.6E-82
18.	HIND-IARI-GENE-1623	MER0791667	family I63 unassigned peptidase inhibitors	6.0E-51
19.	HIND-IARI-GENE-1831	MER0156638	family I63 unassigned peptidase inhibitors	2.4E-13
20.	HIND-IARI-GENE-1862	MER0791387	family I63 unassigned peptidase inhibitors	3.1E-45
21.	HIND-IARI-GENE-2385	MER0156051	family I63 unassigned peptidase inhibitors	1.1E-10
22.	HIND-IARI-GENE-3229	MER0790338	family I63 unassigned peptidase inhibitors	1.9E-16

23.	HIND-IARI-GENE-3837	MER0149830	family I63 unassigned peptidase inhibitors	3.6E-12
24.	HIND-IARI-GENE-5542	MER0147914	family I63 unassigned peptidase inhibitors	4.4E-13
25.	HIND-IARI-GENE-5594	MER0147020	family I63 unassigned peptidase inhibitors	3.2E-08
26.	HIND-IARI-GENE-7478	MER0278557	family I63 unassigned peptidase inhibitors	1.5E-48
27.	HIND-IARI-GENE-3636	MER1134178	family I8 unassigned peptidase inhibitors	3.9E-17
28.	HIND-IARI-GENE-7333	MER0760348	family I8 unassigned peptidase inhibitors	4.1E-12
29.	HIND-IARI-GENE-7898	MER1105718	family I8 unassigned peptidase inhibitors	1.1E-26
30.	HIND-IARI-GENE-9142	MER0759260	family I8 unassigned peptidase inhibitors	1.5E-14
31.	HIND-IARI-GENE-9705	MER1104905	family I84 unassigned peptidase inhibitors	1.6E-27

Table 4.10: Pathogen-host interaction (PHI-Base) analysis of *H. indica* genome

Phenotype	**Bacterium**	**Fungus**	**Protist**	**Nematode**	**Others***
Number of pathogens	87	85	2	2	7
Interactions in total	717	1,354	12	29	27
chemistry_target:_resistance_to_chemical	-	6	-	-	-
chemistry_target:_sensitivity_to_chemical	-	-	-	-	-
effector_(plant_avirulence_determinant)	18	3	-	-	1
enhanced_antagonism	-	-	-	-	-
increased_virulence_(hypervirulence)	61	23	1	-	1
lethal	5	89	-	-	1
loss_of_pathogenicity	12	171	1	2	
reduced_virulence	465	703	10	27	14
unaffected_pathogenicity	156	359	-	-	10
*Others include protozoans and parasites					
Phenotype	**Plant**	**Vertebrata**	**Insect**	**Nematode**	**Others***
Host species	50	13	10	1	3
Interactions in total	1,078	820	143	41	55
chemistry_target:_resistance_to_chemical	4	-	-	-	-
chemistry_target:_sensitivity_to_chemical	-	-	-	-	-
effector_(plant_avirulence_determinant)	6	16	-	-	-
enhanced_antagonism	-	-	-	-	-
increased_virulence_(hypervirulence)	22	53	11	-	-
lethal	43	8	-	-	37
loss_of_pathogenicity	136	48	1	1	7
reduced_virulence	603	498	79	28	11
unaffected_pathogenicity	264	197	52	12	-
*Other include fungus, protozoans, crustaceans					

4.1.7 RNAi machinery in *Heterorhabditis indica*

Seventy-seven *C. elegans* proteins known to be involved in core aspects of RNAi belong to six functional groups, namely, small RNA biosynthesis, dsRNA uptake and spreading, siRNA secondary amplification, AGOs and RISC, RNAi inhibitors, and nuclear effectors. The occurrence of orthologues of C. *elegans* RNAi pathway genes in the predicted proteome of *H . indica* was examined using BLASTx and reciprocal BLASTp approach with domain structure confirmation. We could identify 60 RNAi pathway genes in the *H. indica* (Table 4.11). Different RNAi effector genes identified were nine genes encoding for small RNA biosynthetic proteins; two genes for dsRNA uptake and spreading; six orthologues involved in siRNA amplification; twenty-one for Argonautes, four for RNA-induced silencing complex genes (RISC); seven for RNAi inhibitors, and eleven for nuclear RNAi effectors.

Table 4.11: Orthologs of *C. elegans* RNAi pathway genes identified in *H. indica* genome

Gene name	**Wormbase accession**	***H. indica* genes**	**E-value**	**Bit score**
Small RNA Biosynthesis				
dcr-1	K12H4.8	HIND-IARI-GENE-8972	0	1912
drh-1	F15B10.2	HIND-IARI-GENE-10190	4.17E-51	194
drh-3	D2005.5	HIND-IARI-GENE-10190	3.23E-141	449
drsh-1	F26E4.10	HIND-IARI-GENE-9897	0	739
pash-1	T22A3.5	HIND-IARI-GENE-2917	1.50E-56	197
rde-4	T20G5.11	HIND-IARI-GENE-7459	1.14E-25	104
xpo-1	ZK742.1	HIND-IARI-GENE-10269	0	1528
xpo-2	Y48G1A.5	HIND-IARI-GENE-1397	0	1108
xpo-3	C49H3.10	HIND-IARI-GENE-1399	0	824
dsRNA uptake and spreading				
rsd-2	F52G2.2	-		
rsd-3	C34E11.1	HIND-IARI-GENE-1775	2.23E-50	179
rsd-6	F16D3.2	-		
sid-1	C04F5.1	HIND-IARI-GENE-3667	1.25E-21	98.2
sid-2	ZK520.2.1	-		
siRNA secondary amplification				
ego-1	F26A3.3	HIND-IARI-GENE-2479	0	907

rrf-1	F26A3.8	HIND-IARI-GENE-2479	0	774
rrf-3	F10B5.7	HIND-IARI-GENE-2479	2.06E-108	377
smg-2	Y48G8AL.6	HIND-IARI-GENE-4704	0	1000
smg-5	W02D3.8	HIND-IARI-GENE-5794	1.55E-17	84
smg-6	Y54F10AL.2	HIND-IARI-GENE-10116	1.57E-110	376
RNA induced silencing complex (RISC)-argonautes				
alg-1	F48F7.1	HIND-IARI-GENE-2935	0	1673
alg-2	T07D3.7	HIND-IARI-GENE-2935	0	1565
alg-3	T22B3.2	HIND-IARI-GENE-2935	3.25E-142	451
alg-4	ZK757.3	HIND-IARI-GENE-2935	4.07E-141	448
	C04F12.1	HIND-IARI-GENE-520	6.09E-34	138
	C16C10.3	HIND-IARI-GENE-947	5.15E-90	308
csr-1	F20D12.1	HIND-IARI-GENE-520	1.90E-42	165
ergo-1	R09A1.1	HIND-IARI-GENE-2935	4.95E-49	188
	F58G1.1	HIND-IARI-GENE-9875	6.58E-174	530
nrde-3	R04A9.2	HIND-IARI-GENE-947	2.12E-69	250
ppw-1	C18E3.7	HIND-IARI-GENE-9347	9.51E-66	234
ppw-2	Y110A7A.18	HIND-IARI-GENE-9875	6.99E-162	499
prg-1	D2030.6	HIND-IARI-GENE-9953	0	860
	R06C7.1	HIND-IARI-GENE-9875	0	578
rde-1	K08H10.7	HIND-IARI-GENE-2935	4.87E-46	178
sago-1	K12B6.1	HIND-IARI-GENE-9347	1.05E-66	237
sago-2	F56A6.1	HIND-IARI-GENE-9347	6.52E-65	232
	T22H9.3	HIND-IARI-GENE-947	3.80E-90	308
	T23D8.7	HIND-IARI-GENE-2935	0	567
	Y49F6A.1	HIND-IARI-GENE-947	2.22E-85	294
	ZK1248.7	HIND-IARI-GENE-9875	0	585
ain-1	C06G1.4.1	HIND-IARI-GENE-1354	9.05E-89	299
ain-2	B0041.2	HIND-IARI-GENE-1354	4.69E-19	90.9
tsn-1	F10G7.2	HIND-IARI-GENE-4491	0	1158
vig-1	F56D12.5	HIND-IARI-GENE-3499	1.23E-39	143
RNAi inhibitors				
adr-1	H15N14.1	HIND-IARI-GENE-6812	2.68E-93	315
adr-2	T20H4.4	HIND-IARI-GENE-6812	5.75E-05	44.7
eri-1	T07A9.5	HIND-IARI-GENE-3256	6.09E-61	201

eri-3	W09B6.3	-		
eri-5	Y38F2AR.1	HIND-IARI-GENE-9042	2.40E-34	135
eri-6/7	C41D11.7	HIND-IARI-GENE-9457	3.16E-53	194
lin-15b	ZK662.4	-		
xrn-1	Y39G8C.1	HIND-IARI-GENE-430	0	1119
xrn-2	Y48B6A.3	HIND-IARI-GENE-518	0	1139
Nuclear effectors				
cid-1	K10D2.3	HIND-IARI-GENE-9193	3.94E-26	112
ekl-1	F22D6.6	HIND-IARI-GENE-9042	5.32E-95	302
ekl-4	Y105E8A.17	HIND-IARI-GENE-132	9.75E-168	482
ekl-5	Y26E6A.1	-		
ekl-6	T16G12.5	HIND-IARI-GENE-7036	5.14E-76	268
gfl-1	M04B2.3	HIND-IARI-GENE-9857	6.32E-99	285
mes-2	R06A4.7	HIND-IARI-GENE-1929	3.75E-42	163
mes-3	F54C1.3	-		
mes-6	C09G4.5.1	HIND-IARI-GENE-10098	3.01E-18	85.9
mut-2	K04F10.6	HIND-IARI-GENE-2812	2.38E-78	249
mut-16	B0379.3	-		
mut-7	ZK1098.8	HIND-IARI-GENE-434	8.08E-95	317
rde-2	F21C3.4	-		
rha-1	T07D4.3	HIND-IARI-GENE-1233	0	1256
zfp-1	F54F2.2	HIND-IARI-GENE-5727	7.15E-91	305

4.1.8 Microsatellites

Microsatellites are commonly referred to as simple sequence repeats (SSRs). They are polymorphic DNA loci with tandem repeat sequences of 1–6 bp. SSRs are co-dominant and exhibit a high mutation rate. These genetic markers are very useful in resolving relationships among closely related species. *H. indica* genome sequences (3,538 scaffolds) were subjected to *in-silico* MISA (microsatellite finder) analysis to predict microsatellites/SSRs. A total of 2,954 microsatellite loci were predicted in 459 scaffolds of the current *H. indica* draft genome (Table 4.12). Out of 459 scaffolds, 354 contained more than one SSRs. The majority of them are monomeric repeats (1,871) followed by dinucleotide repeats (903). Fifty-three of the detected SSRs were present in compound formation.

Table 4.12: Microsatellites predicted in *H. indica* genome

Total number of sequences examined	3,538
Total size of examined sequences (bp)	91,263,125
Total number of identified SSRs	2,954
Number of SSR containing sequences	459
Number of sequences containing more than 1 SSR	354
Number of SSRs present in compound formation	53
Distribution of different repeat type classes	
Monomeric repeat	1,871
Dinucleotide repeat	903
Trinucleotide repeat	170
Tetranucleotide repeat	10

4.1.9 Transposable elements

A search for transposable elements (TE) in *H. indica* genome revealed that TEs are distributed on 521 scaffolds. A total of 1,548 transposable elements belonging to 11 different classes were detected in *H. indica* genome (Table 4.13). TEs belonging to Zator/TIR/DNA transposon (class=2/1/5), hAT/TIR/DNA transposon (class=2/1/2) and, Gypsy/LTR/Retrotransposon (class=1/1/2) classes were most abundant in *H. indica* genome.

Table 4.13: Different classes of transposable elements detected in *H. indica* genome

Transposon Class	**Description of Transposon class**	**Number detected in *H. indica* genome**
class=2/1/5	Zator/TIR/DNA Transposon	492
class=2/1/2	hAT/TIR/DNA Transposon	347
class=1/1/2	Gypsy/LTR/Retrotransposon	278
class=2/1/1	Tc1-Mariner/TIR/DNA Transposon	131
class=2/1/3	CMC/TIR/DNA Transposon	98
class=1/1/1	Copia/LTR/Retrotransposon	79
class=2/1/4	Sola/TIR/ DNA Transposon	76
class=2/2	Helitron/DNA Transposon	38
class=2/3	MITE/DNA Transposon	4
class=1/1/3	ERV/LTR/Retrotransposon	3
class=2/1/6	Novosib/TIR/DNA Transposon	2

4.1.10 Noncoding RNAs (ncRNAs)

Infernal (INFERence of RNA ALignment) tool was used to identify ncRNA homologs in *H. indica* genome. A total of 631 ncRNA were identified on 249 scaffolds of *H. indica* genome. Out of them, 269 were tRNA genes. 59 CRISPR-DR4 loci were identified. The top 10 ncRNAs identified are presented in the table/figure. Moreover, 7 t-RNA-Sec loci were identified. Top predicted ncRNA which constitutes ~70% of the total detected ncRNA in *H. indica* genome are represented in table 4.14.

Table 4.14: Top ncRNA (non-coding RNA) detected in *H. indica* genome

Accession	**ncRNA gene**	**Description**	**Number of loci detected**
RF00005	tRNA	tRNA	269
RF01317	CRISPR-DR4	CRISPR RNA direct repeat element	59
RF01695	C4	C4 antisense RNA	14
RF00059	TPP	TPP riboswitch (THI element)	14
RF00004	U2	U2 spliceosomal RNA	11
RF00003	U1	U1 spliceosomal RNA	9
RF01852	tRNA-Sec	Selenocysteine transfer RNA	7
RF00012	U3	Small nucleolar RNA U3	7
RF00020	U5	U5 spliceosomal RNA	7
RF02543	LSU_rRNA_eukarya	Eukaryotic large subunit ribosomal RNA	6
RF00050	FMN	FMN riboswitch (RFN element)	6
RF00504	Glycine	Glycine riboswitch	6
RF00032	Histone3	Histone 3' UTR stem-loop	6
RF00174	Cobalamin	Cobalamin riboswitch	5
RF01766	cspA	cspA thermoregulator	5
RF01989	SECIS_3	Selenocysteine insertion sequence 3	5
RF00001	5S_rRNA	5S ribosomal RNA	4
RF00013	6S	6S / SsrS RNA	4
RF02346	ar35	Alphaproteobacterial sRNA ar35	4
RF00485	K_chan_RES	Potassium channel RNA editing signal	4
RF00391	RtT	RtT RNA	4
RF00015	U4	U4 spliceosomal RNA	4

4.1.11 Likelihood of horizontal gene transfer (HGT):

Candidate HGT events in *H. indica* genome were predicted using the Alienness tool, where the main focus was to identify HGT of non-metazoan origin to a metazoan species. Based on the Alienness index (AI) >14, we identified 29 likely cases of HGT (Figure 4.8). Major donors seem to be from Proteobacteria and Fungi taxonomic groups. Gamma proteobacteria, followed by alpha-proteobacteria in the bacterial group, and fungi *Rhizopus oryzae* and *Beauveria bassiana* appear to be major contributors to HGT events. HGT score was highest (1,621) for the potential donor *B. bassiana*. Further filtering was done based on AI> 45 to eliminate false positives and focus on candidates which are likely to be highly supported by phylogenies. Subject to further validation, a total of 21 likely HGT cases were detected in *H. indica* genome (Table 4.15).

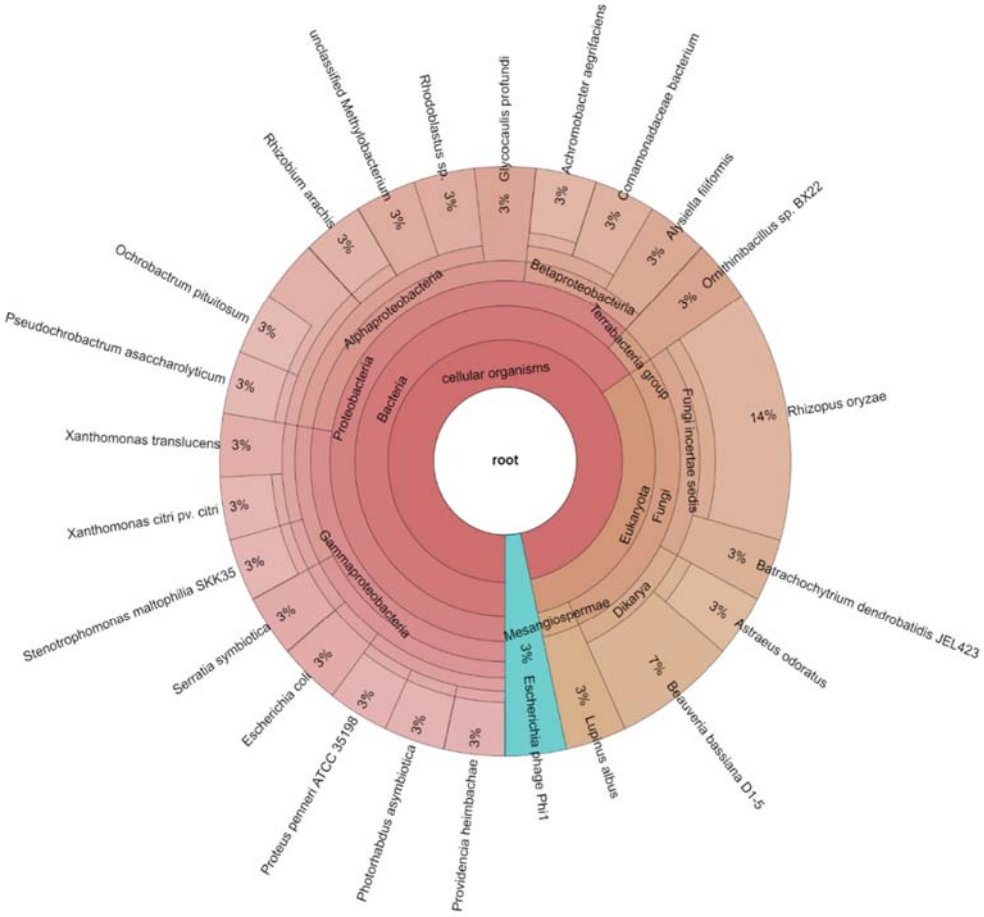


Figure 4.8: Putative donors of horizontal gene transfer events in *H. indica* genome as assessed by Alienness tool.

Table 4.15: Likely HGT events predicted in *H. indica* genome using Alienness tool based on Alienness Index (AI) of >45

S. No.	Query	AI	HGT score	Potential donor match	Donor_taxonomy
1.	HIND-IARI-GENE-10329	460.52	1621	*Beauveria bassiana* D1-5	Eukaryota@Fungi
2.	HIND-IARI-GENE-10438	460.52	542	unclassified *Methylobacterium*	Bacteria
3.	HIND-IARI-GENE-6415	460.52	595	*Rhodoblastus* sp.	Bacteria
4.	HIND-IARI-GENE-7383	460.52	582	*Beauveria bassiana* D1-5	Eukaryota@Fungi
5.	HIND-IARI-GENE-8159	460.52	523	*Ochrobactrum*	Bacteria
6.	HIND-IARI-GENE-9345	460.52	613	*Astraeus odoratus*	Eukaryota@Fungi
7.	HIND-IARI-GENE-2432	392.12	490	*Ochrobactrum pituitosum*	Bacteria
8.	HIND-IARI-GENE-6700	391.26	525	*Comamonadaceae bacterium*	Bacteria
9.	HIND-IARI-GENE-2313	309.94	396	*Pseudochrobactrum asaccharolyticum*	Bacteria
10.	HIND-IARI-GENE-2558	307.11	390	*Escherichia coli*	Bacteria
11.	HIND-IARI-GENE-8898	262.14	343	*Photorhabdus asymbiotica*	Bacteria
12.	HIND-IARI-GENE-6124	229.55	336	*Xanthomonas citri* pv. *citri*	Bacteria
13.	HIND-IARI-GENE-5234	206.56	323	*Stenotrophomonas maltophilia* SKK35	Bacteria
14.	HIND-IARI-GENE-1300	204.44	288	Escherichia phage Phi1	Viruses
15.	HIND-IARI-GENE-7350	200.49	268	*Serratia symbiotica*	Bacteria
16.	HIND-IARI-GENE-7385	187.53	283	*Rhizopus oryzae*	Eukaryota@Fungi
17.	HIND-IARI-GENE-1512	122.31	206	*Achromobacter aegrifaciens*	Bacteria
18.	HIND-IARI-GENE-2341	94.9	156	*Rhizopus oryzae*	Eukaryota@Fungi
19.	HIND-IARI-GENE-8010	91.65	147	*Providencia heimbachae*	Bacteria
20.	HIND-IARI-GENE-4449	75.03	139	*Batrachochytrium dendrobatidis* JEL423	Eukaryota@Fungi
21.	HIND-IARI-GENE-10330	45.45	54	*Proteus penneri* ATCC 35198	Bacteria

4.1.12 Mitochondrial genome of *H. indica* Hms1-i20

The mitochondrial genome was assembled separately which is 17,393 bp in size with 70.94% (A + T)s and 27.33% (G + C)s content. A total of 46 genes were detected in mitochondrial genome assembly. Out of 46 genes, 23 are tRNA, 3 are rRNA genes. Six genes are associated with the origins of replication, and the other 14 genes are related to mitochondrially encoded ATP synthase membrane subunits, mitochondrial apocytochrome b (cob), cytochrome c oxidase, and mitochondrial membrane respiratory chain NADH dehydrogenases (Table 4.16).

Table 4.16: Gene detected in mitochondrial genome assembly of *H. indica*

Mitochondrial genome: scaffold477_17393							
Databases used for annotation: mitfi and Mitos							
Type	**Gene**	**Start**	**End**	**Type**	**Gene**	**Start**	**End**
tRNA	*trnA_0(tgc)*	87	141	**rRNA**	*rrnL_0*	493	1134
	trnH(gtg)	1450	1503		*rrnL_1*	1292	1380
	trnC(gca)	2213	2267		*rrnS*	11694	12373
	trnM(cat)	2268	2328	**rep_origin**	*OH_1-a*	1163	1200
	trnG(tcc)	5511	5565		*OH_1-c*	5217	5279
	trnD(gtc)	5566	5620		*OH_1-b*	13410	13526
	trnT(tgt)	6861	6915		*OH_2*	13831	13892
	trnL1(tag)	7690	7746		*OL*	14119	14148
	trnF(gaa)	8787	8841		*OH_0*	14750	14843
	trnQ(ttg)	8923	8977	**Other important genes**	*nad3*	158	493
	trnR(acg)	8978	9031		*cox2*	1516	2211
	trnI(gat)	9032	9095		*cox1*	2327	3865
	trnS1(tct)	9948	10001		*nad4*	5625	6794
	trnL2(taa)	10002	10056		*cox3*	6922	7689
	trnK(ttt)	10059	10119		*cob-b*	7750	8610
	trnY(gta)	11519	11574		*cob-a*	8598	8825
	trnN(gtt)	11583	11638		*nad2*	9096	9941
	trnS2(tga)	11640	11694		*atp6*	10127	10777
	trnE(ttc)	12389	12445		*nad1-b*	10729	10842
	trnW(tca)	12446	12501		*nad1-a*	10841	11485
	trnV(tac)	13113	13166		*nad4l*	12512	12751
	trnP(tgg)	13381	13435		*nad6*	12755	13132
	trnA_1(tgc)	15329	15383		*nad5*	15384	17033

4.2 Elucidation of genes involved in nematode-bacterium symbiosis by transcriptome sequencing of *Heterorhabditis* at biofilm formation stage of symbiosis

4.2.1 RNA-Sequencing of *Heterorhabditis* nematodes at symbiont biofilm formation stages

Photorhabdus bacterial cells adhere to posterior intestinal cells of the nematode and make a biofilm at the early adult stage (~36-40 h after IJ recovery). Symbiont biofilm formation is a critical step in the maternal transmission of the symbiont bacteria to progeny. The early-adult hermaphrodite animals were visualized under the microscope and documented for symbiont bacterial biofilm in the symbiotic (S) nematodes, whereas no biofilm was found in the axenic (A) nematodes at the same developmental stage (Figure 4.9).

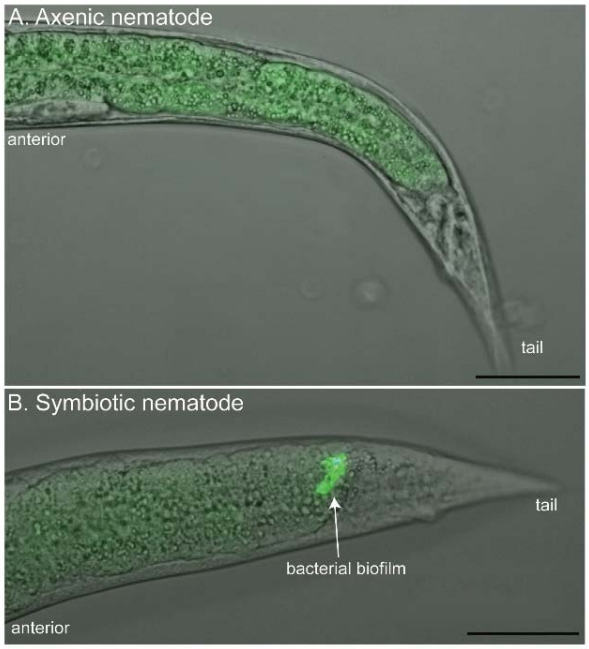


Figure 4.9. Early adult stage of symbiotic and axenic *H. bacteriophora* used for RNA-sequencing. A: Axenic early adult devoid of bacterial biofilm B. Bacterial biofilm in the posterior intestine of the early adult stage of symbiotic nematode.

Approximately 27 to 40 million raw reads were generated per sample for three independent biological replicates each for the two experimental conditions (axenic and symbiotic) (Table 4.17). Read quality filtering by fastp tool resulted in the selection of >94% paired-end HQ reads per sample. A total of 26.28 to 40.58 million HQ reads were obtained for the three replicates of axenic nematodes, whereas 31.91 to 39.60 million HQ reads were obtained for the three replicates of symbiotic nematodes, respectively (Table 4.17). Reads from the three biological replicates showed correlation coefficients from 0.83 to 0.95 and 0.77 to 0.91 for axenic and symbiotic nematodes, respectively (Figure 4.10). Raw sequence data has been deposited in Sequence Read Archive, with accession numbers SRR14162364, SRR14162365, SRR14162366, SRR14162367, SRR14162368, and SRR14162369. The bio-sample accession number is SAMN18646821, and the Bio-project accession number is PRJNA720314.

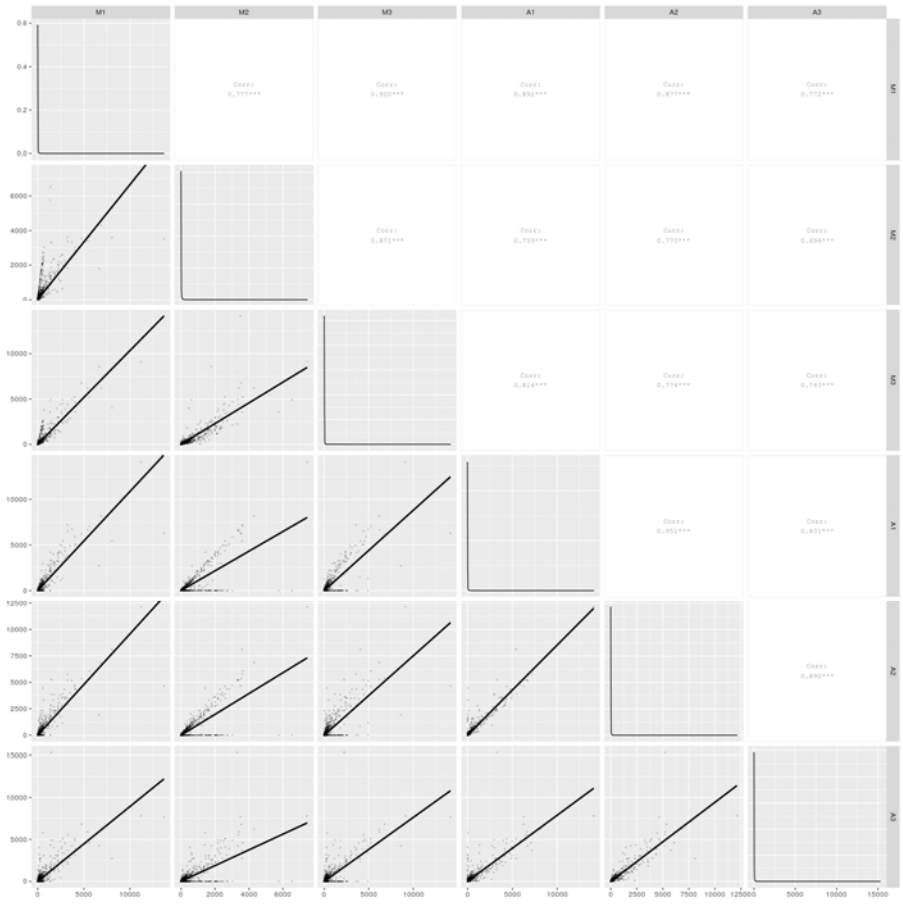


Figure 4.10: Correlation coefficient between various replicates of the RNA-seq samples

Table 4.17: Raw and filtered read statistics for the RNA-Seq experiment

Sample	**Raw reads**			**Filtered reads**			
	Total reads	Total bases (Giga bases)	GC %	Total reads	Total bases (Giga bases)	GC %	% HQ Reads
Axenic (Rep 1)	27,337,042 For:13,668,521 Rev:13,668,521	4.10	44.02	26,288,598 For: 13,144,299 Rev: 13,144,299	3.56	43.32	96.16
Axenic (Rep 2)	29,222,012 For:14,622,006 Rev:14,622,006	4.38	43.99	27,979,572 For: 13,989,786 Rev:13,989,786	3.87	43.33	95.74
Axenic (Rep 3)	41,116,354 For:20,558,177 Rev:20,558,177	6.16	41.45	40,583,640 For: 20,291,820 Rev:20,291,820	6.06	41.44	98.71
Symbiotic (Rep 1)	33,448,662 For: 16,724,331 Rev:16,724,331	5.01	42.75	31,915,618 For: 15,957,809 Rev:15,957,809	4.38	41.87	95.42
Symbiotic (Rep 2)	36,100,070 For: 18,050,035 Rev:18,050,035	5.41	44.26	34,093,176 For: 17,046,588 Rev:17,046,588	4.59	43.41	94.44
Symbiotic (Rep 3)	40,100,056 For: 20,050,028 Rev:20,050,028	6.01	40.99	39,603,638 For: 19,801,819 Rev:19,801,819	5.90	40.96	98.76

4.2.2 Transcriptome assembly, completeness and annotation

The quality-filtered HQ sequence data from all the samples were used to generate a *de novo* reference assembly of 95.7 Mb using Trinity assembler. Assembly was improved by amelioration, removing the duplicates by CD_HIT_EST that led to the assembly of 46,599 transcripts. N50 value of the assembly was 2,681 bp, and the average transcript length was 2,054 bp. The statistical details for the assembly are provided in Table 4.18.

Transcriptome assembly completeness was assessed using BUSCO against eukaryota_odb10.2019-11-20 (255 genes) database, which showed the presence of 95.3% of complete BUSCOs (C), 2% were fragmented (F) and 2.7% were missing (M). A comparison to nematoda_odb10.2019-11-20 (3,131 genes) database found 93.6% BUSCOs as complete (C), 1.6% as fragmented (F) and 4.8% as missing (M) (Table 4.18). Approximately 97% of the reads across the samples mapped backed to our transcriptome assembly.

Table 4.18: Assembly statistics and completeness assessment (by BUSCO against Eukaryota and Nematoda databases) of the *de-novo H. bacteriophora* transcriptome assembly

Assembly statistics		
Number of transcripts	46,599	
Total bases (bp)	95,741,839 (95.7 Mb)	
Minimum sequence length (bp)	501	
Maximum sequence length (bp)	19,436	
Average sequence length (bp)	2,054.59	
N50 value (bp)	2,681	
(G + C)s	38.11 %	
Assembly completeness assessment using BUSCO		
Against eukaryota_odb10.2019-11-20 (Total BUSCO groups = 255)		
Complete BUSCOs (C)	243	95.30 %
Complete and single-copy BUSCOs (S)	75	29.40 %
Complete and duplicated BUSCOs (D)	168	65.90 %
Fragmented BUSCOs (F)	5	2 %
Missing BUSCOs (M)	7	2.70 %
Against nematoda_odb10.2019-11-20 (Total BUSCO groups = 3,131)		
Complete BUSCOs (C)	2,930	93.60 %
Complete and single-copy BUSCOs (S)	1,109	35.40 %
Complete and duplicated BUSCOs (D)	1,821	58.20 %
Fragmented BUSCOs (F)	50	1.60 %
Missing BUSCOs (M)	151	4.80 %

Of the total 46,599 transcripts, 33,762 (72.45%) could be annotated using blastx against Uniprot90 / SwissProt database and Trinotate tool. Ref-Seq analysis revealed that nematode genes were the top hits for 17,952 of the transcripts. Functional characterization of transcripts using Gene Ontology (GO) resulted in 33,760 annotated transcripts. The top ten GO terms enriched in the symbiotic *H. bacteriophora* transcriptome under each category of molecular function, biological processes and cellular components are provided in figure 4.11. Positive regulation of transcription by RNA polymerase II (GO:0045944), multicellular organism development (GO:0007275) and cell differentiation (GO:0030154) were the top three GO terms under biological process category with 936, 921 and 725 transcripts, respectively. Nucleus (GO:0005634), cytoplasm (GO:0005737) and integral component of membrane (GO:0016021) were the most enriched under cellular component category with 7,493, 7,220 and 6,352 transcripts under these terms, respectively. Under the molecular function category, ATP binding (GO:0005524), metal ion binding (GO:0046872), and DNA binding (GO:0003677) were the top three terms with 4,538, 4,397 and 1,689 transcripts, respectively. Functional annotation by KAAS database resulted in the annotation of 19,864 transcripts grouping into 297 pathways, out of which thermogenesis (ko04714), endocytosis (ko04144), and protein processing in endoplasmic reticulum (ko04141) pathways were the most abundant (Figure 4.11).

4.2.3 Differentially expressed transcripts between symbiotic and axenic nematodes

The relative expression of transcripts in symbiotic and axenic early adults of *Heterorhabditis* based on FPKM values is provided in the figure 4.12. The transcripts that were differentially expressed in the symbiotic nematodes as compared to axenic nematodes were identified by using DESeq package. A total of 754 transcripts (1.62% of the total) were differentially expressed in symbiotic nematodes as compared to axenic nematodes. The number of down-regulated transcripts was 547, whereas 207 transcripts were up-regulated in the symbiotic nematodes (figure 4.12). These 754 transcripts are henceforth mentioned as “DEGs”. A list of the top ten up-and down-regulated transcripts is provided in Table 4.19. In addition, 12,151 transcripts (26.05% of total) were exclusive to symbiotic nematodes, whereas 20 transcripts were found only in axenic nematodes.

The Web Gene Ontology Annotation (WEGO) analysis of DEGs divided these genes into three principal GO categories: biological process, cellular component, and molecular function. A total of 37 and 36 terms were enriched in down-regulated and upregulated transcripts respectively (Figure 4.13). In both the cases, cell, cell-part and membrane represented the top three enriched terms under the cellular component category. Binding, catalytic activity and molecular transducer activity and transporter activity were the most enriched terms under the molecular function category. Under the biological processes category, cellular process, biological regulation, and metabolic process terms were the most enriched in down-regulated transcripts, whereas, cellular process, multicellular organismal process, biological regulation, and metabolic process were most enriched terms in the up-regulated transcripts.

The WEGO analysis of the symbiotic nematode-specific transcripts showed enrichment of 52 terms (Figure 4.14). In the cellular component category, cell, cell part and membrane were the top three enriched terms. In the molecular function category, binding, catalytic activity and transporter activity were the top enriched terms, whereas, in the biological process category, cellular process, metabolic process and biological regulation were the three most enriched terms. Pathway analysis of DEGs using the KEGG automated annotation server (KAAS) revealed that ribosome, calcium signalling and neuroactive ligand-receptor interaction were the top three enriched pathways (Figure 4.15A). However, in the symbiotic nematode-specific transcripts, endocytosis, cAMP signalling pathway, and focal adhesion were the top three enriched pathways (Figure 4.15B).

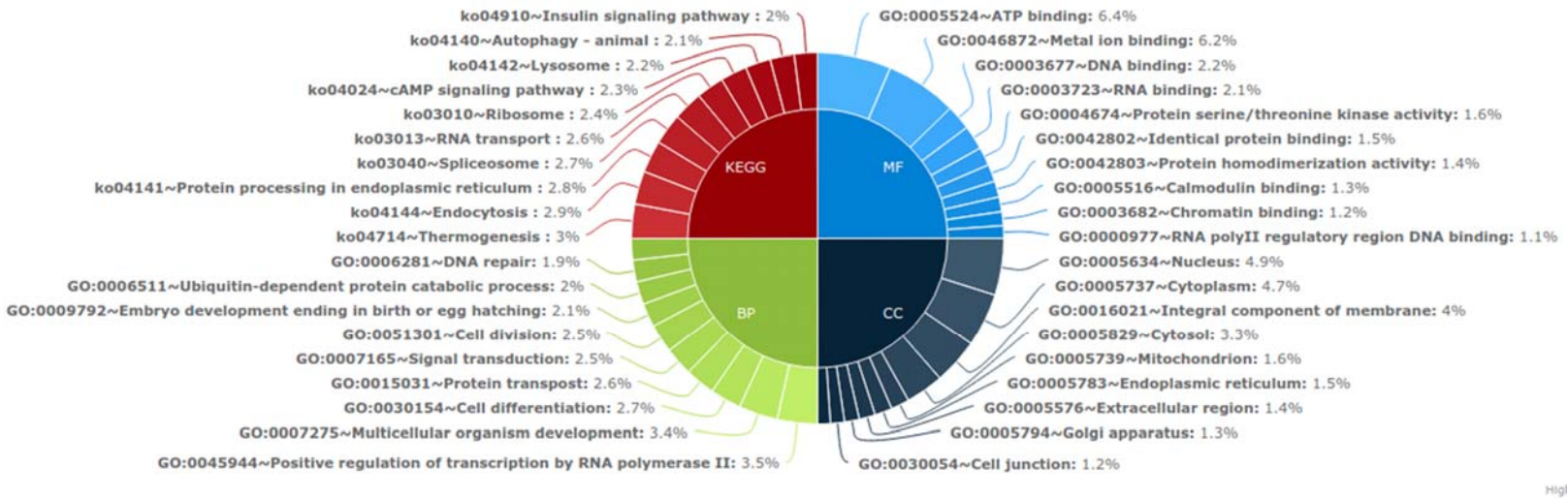


Figure 4.11: Donut chart representing the Gene Ontology (GO) and KAAS (KEGG Automatic Annotation Server) functional annotations of all the transcripts pooled from the treatment (symbiotic nematodes) and control (axenic nematodes) groups. The top 10 enriched gene ontologies (GO) terms under each category of cellular component, biological process and molecular function (the inner circle) and top 10 ten enriched KEGG pathways are represented. The relative area under each GO term/KEGG ID indicates the number of transcripts mapping to that GO term/KEGG ID.

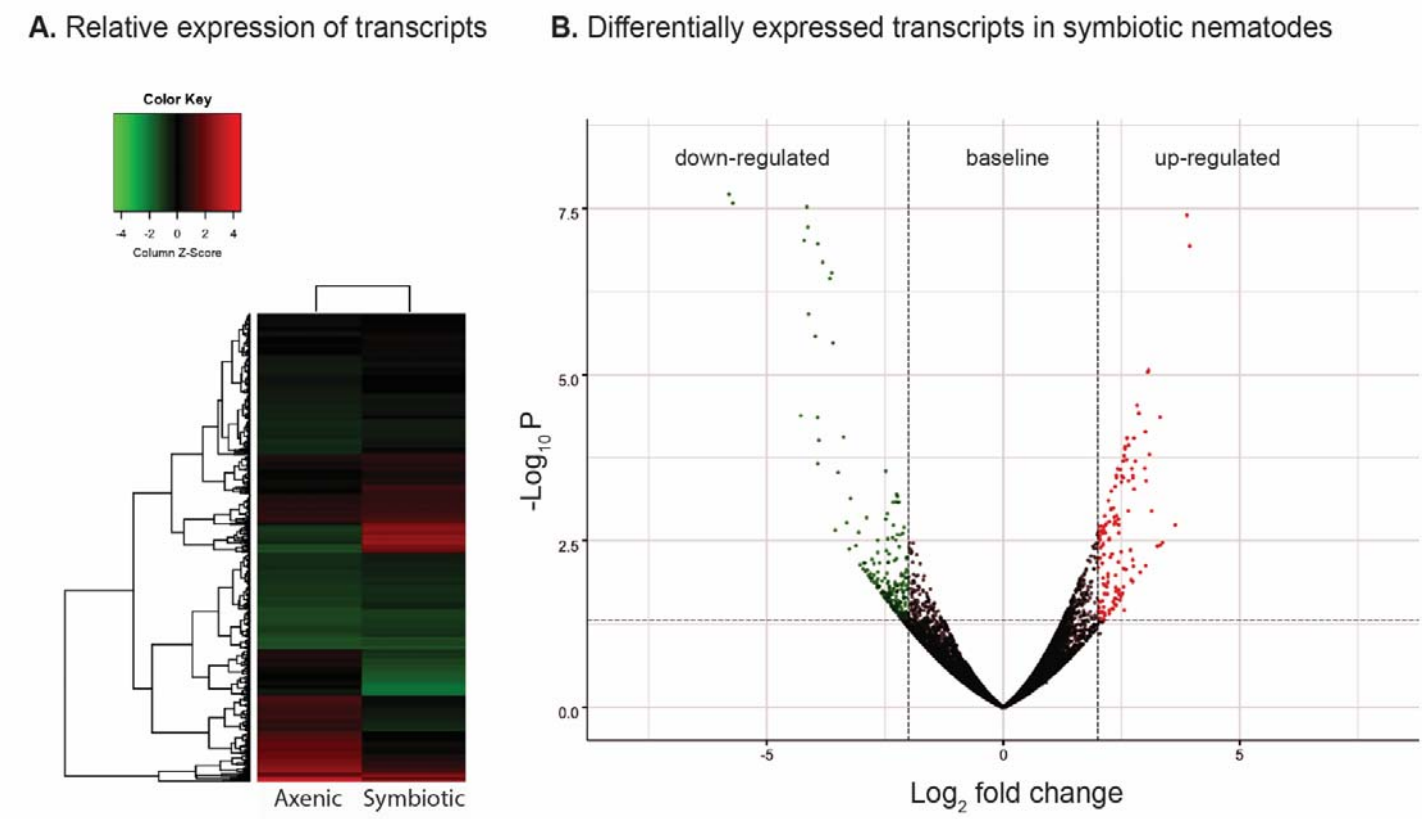


Figure 4.12: A. Heat map showing the relative expression of transcripts in symbiotic and axenic early adults of *Heterorhabditis*. The map was generated by using the FPKM values, and all the three replicates (R1-R3) are presented. **B.** Volcano plots representing differentially expressed transcripts (log2 fold change) in symbiotic nematodes as compared to axenic nematodes. Out of 754 differentially expressed transcripts, 547 transcripts were down-regulated, whereas, 207 transcripts were up-regulated in symbiotic nematodes.

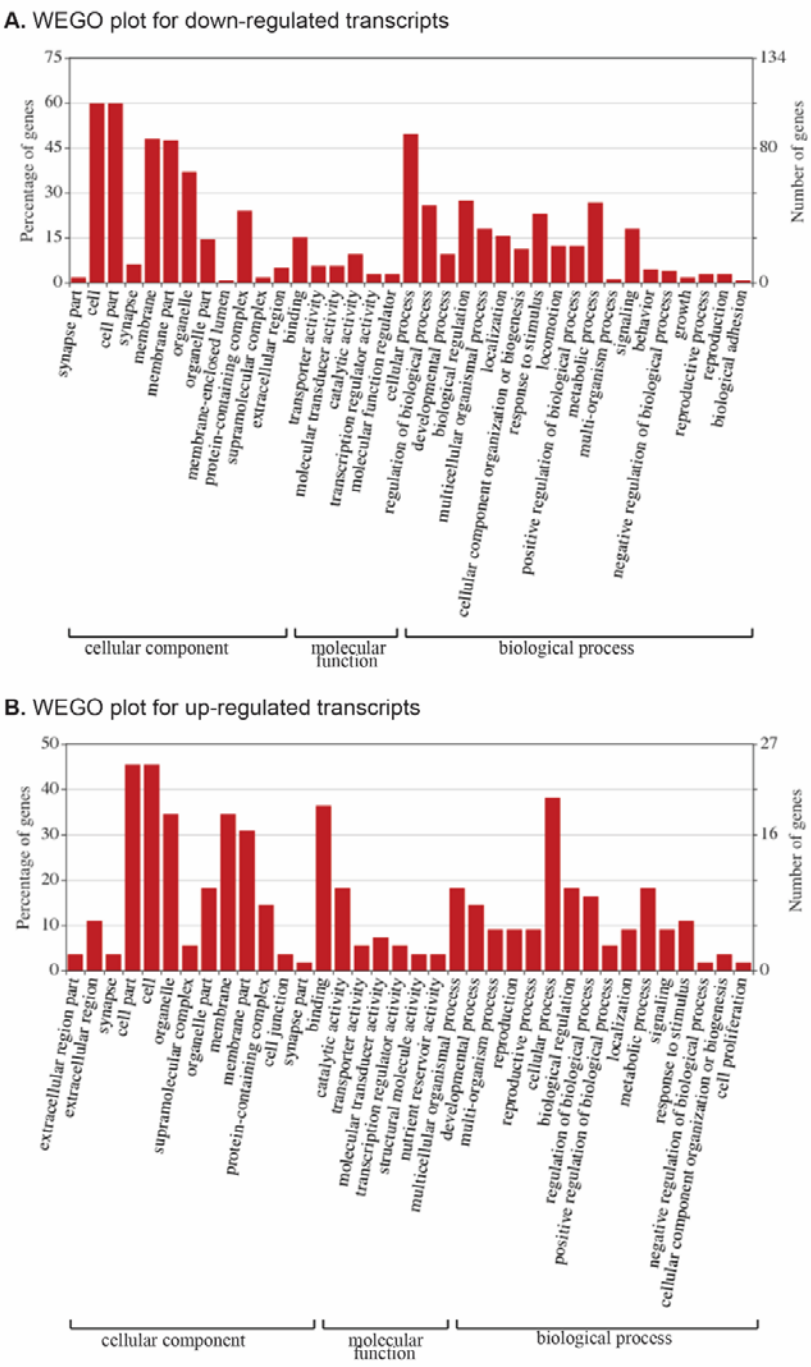


Figure 4.13. Gene Ontology (GO) annotation of differentially expressed transcripts identified in RNA-seq experiment. **A.** up-regulated transcripts B. down-regulated transcripts. WEGO (Web Gene Ontology Annotation Plot) tool was used for visualizing, comparing and plotting GO annotation results.

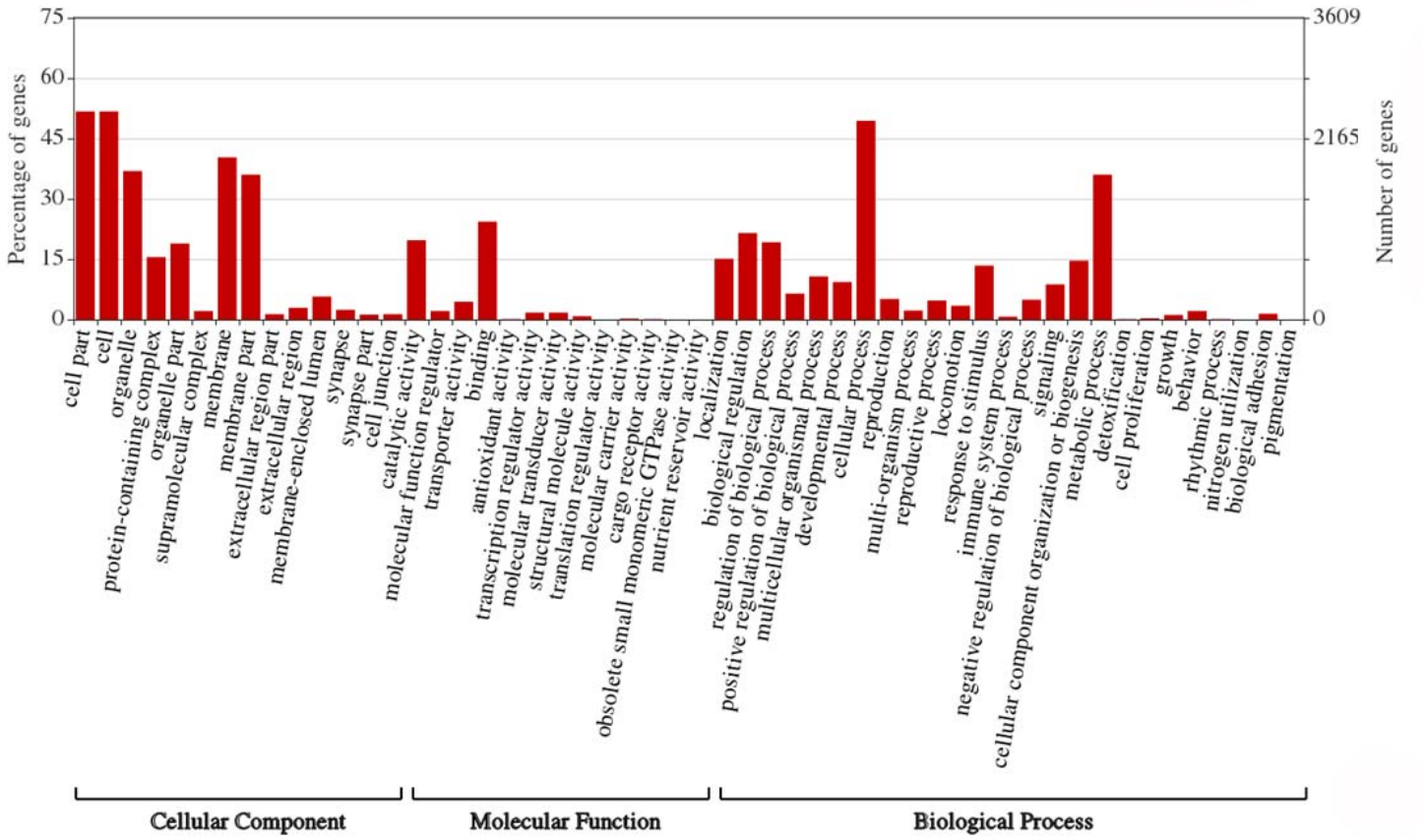


Figure 4.14. Gene Ontology (GO) annotation of symbiotic-nematode-specific transcripts. WEGO (Web Gene Ontology Annotation Plot) tool was used for visualizing, comparing and plotting GO annotation results.

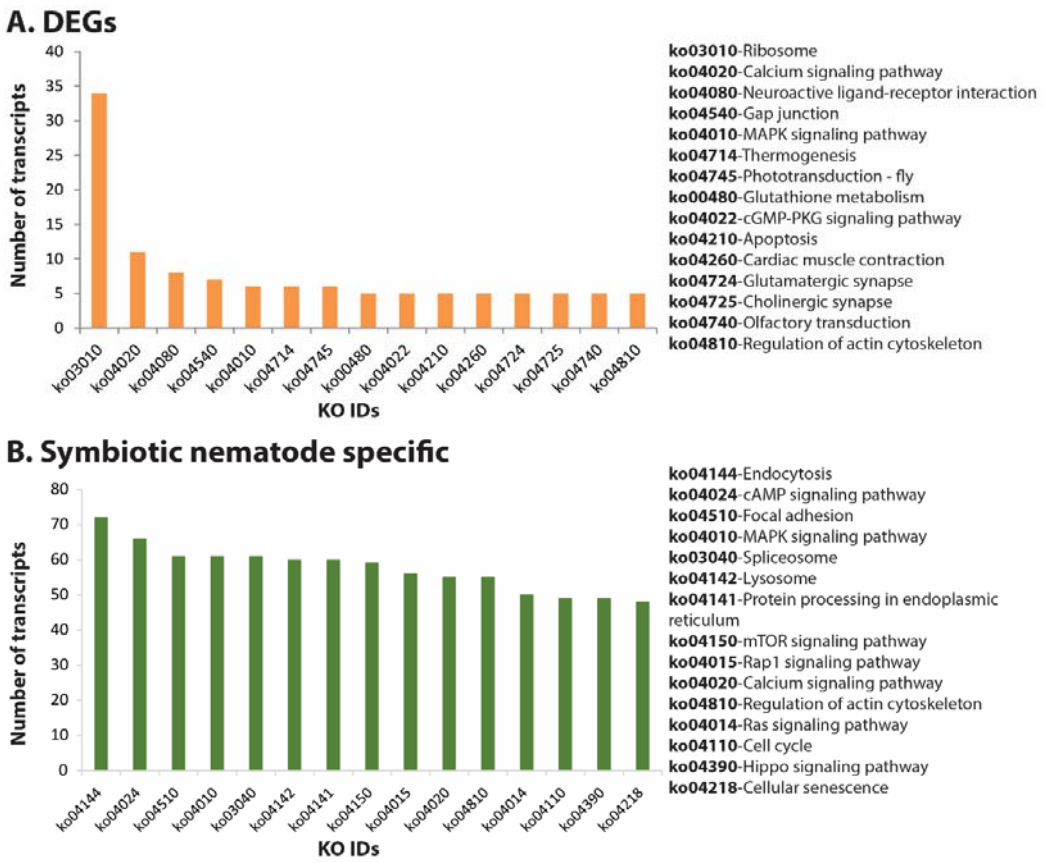


Figure 4.15. KEGG pathway enrichment analysis of transcripts identified in the RNA-seq experiments. **A.** KEGG pathway annotations of differentially expressed transcripts. **B.** KEGG pathway annotations of symbiotic-nematode transcripts. KEGG Automatic Annotation Server (KAAS) was used for pathway annotation analyses.

4.2.4 Genes putatively important for bacterial symbiosis

We hypothesized that in addition to the differentially expressed genes, the genes unique to symbiotic nematodes may be relevant for nematode-bacterial symbiosis. Since the symbiotic nematodes are interacting with the bacteria, we looked for transcripts potentially involved in interaction with bacteria based on annotation, such as GO terms related to responses to bacteria, immune system and defense responses (Figure 4.16). A total of 314 transcripts were identified and mapped to 167 genes based on annotations. A total of 65 genes grouped into 20 GO terms related to response to bacteria. Seventy-three genes grouped under the terms related to the immune system, where the term 'innate immune response' (GO:0045087) was most enriched. All the other terms such as 'defense response' (GO:0006952), 'defense response to fungus'(GO:0050832), 'defense response to virus' (GO:0051607) and 'behavioral defense response' (GO:0002209) were grouped into 'miscellaneous category' containing 69 genes (Figure 4.16). Out of 167 genes, 78 genes assigned to various biological events highly relevant from the perspective of animal-bacterial interaction (Table 4.20). Several transcripts belonging to the canonical nematode immune signalling pathways were identified in the symbiotic *H. bacteriophora* transcriptome. From the transforming growth factor β (TGF-β)/DBL-1 pathway- *sma-6*, *sma-3* and *sma-4*; from p38 mitogen-activated protein kinase (MAPK) pathway- *tir-1*, *nsy-1*, *sek-1* and *pmk-1*; from insulin-like receptor (ILR or DAF-2/DAF-16) pathway- *age-1* and *daf-16*; from extracellular signal-regulated kinase (ERK) pathway- *lin-45* and *mek-2* transcripts, and from c-Jun N-terminal kinase (JNK) pathway- *mlk-1, mek-1* and *kgb-1* were detected in the symbiotic nematode's transcriptome (Figure: 4.17). *abl-1*, *bsk*, and *vhp-1*, which are implicated in stress response pathways were also found in symbiotic nematodes.

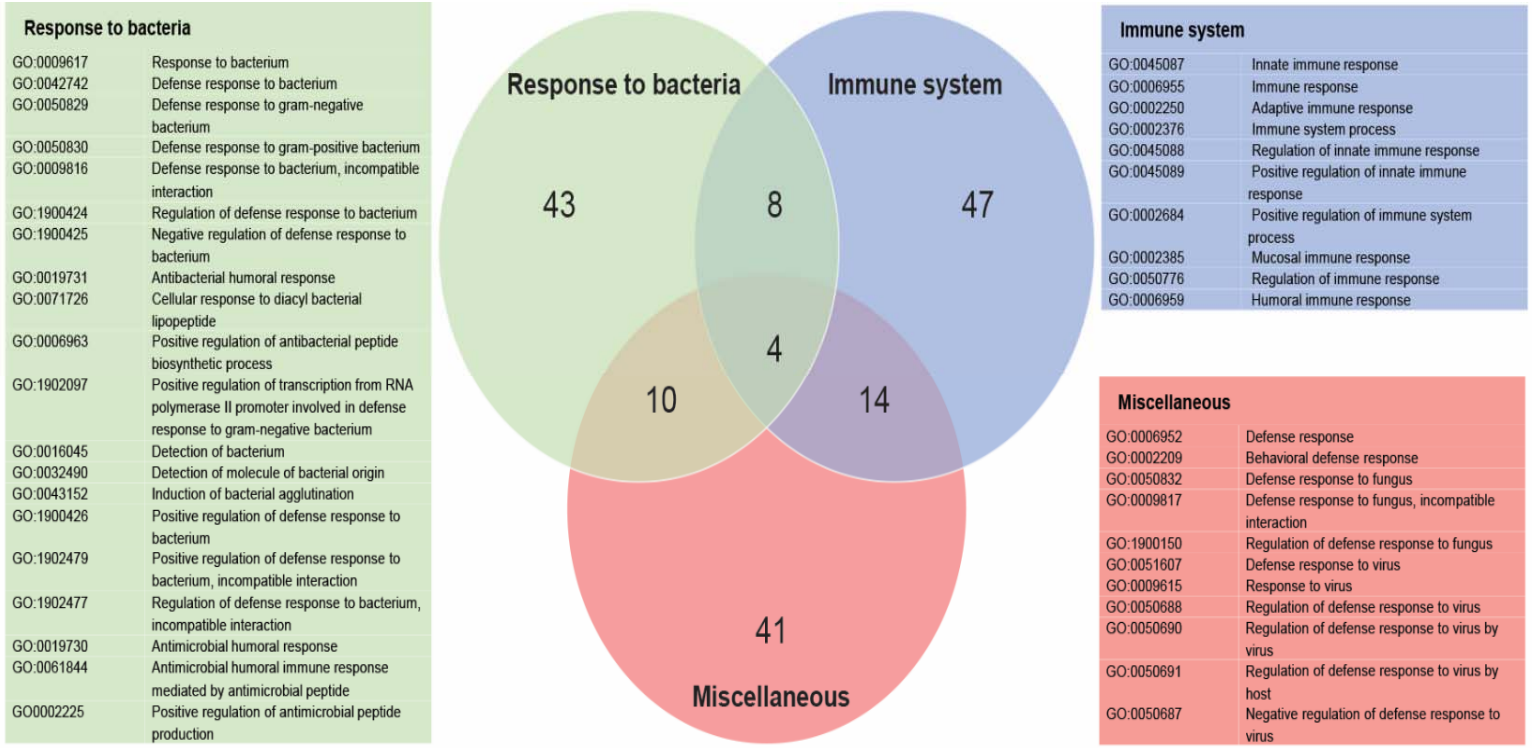


Figure 4.16. Venn diagram showing specific genes pulled out from symbiotic-nematode specific transcriptome (based on GO annotations) putatively involved in responses to bacteria, immune and defense responses.

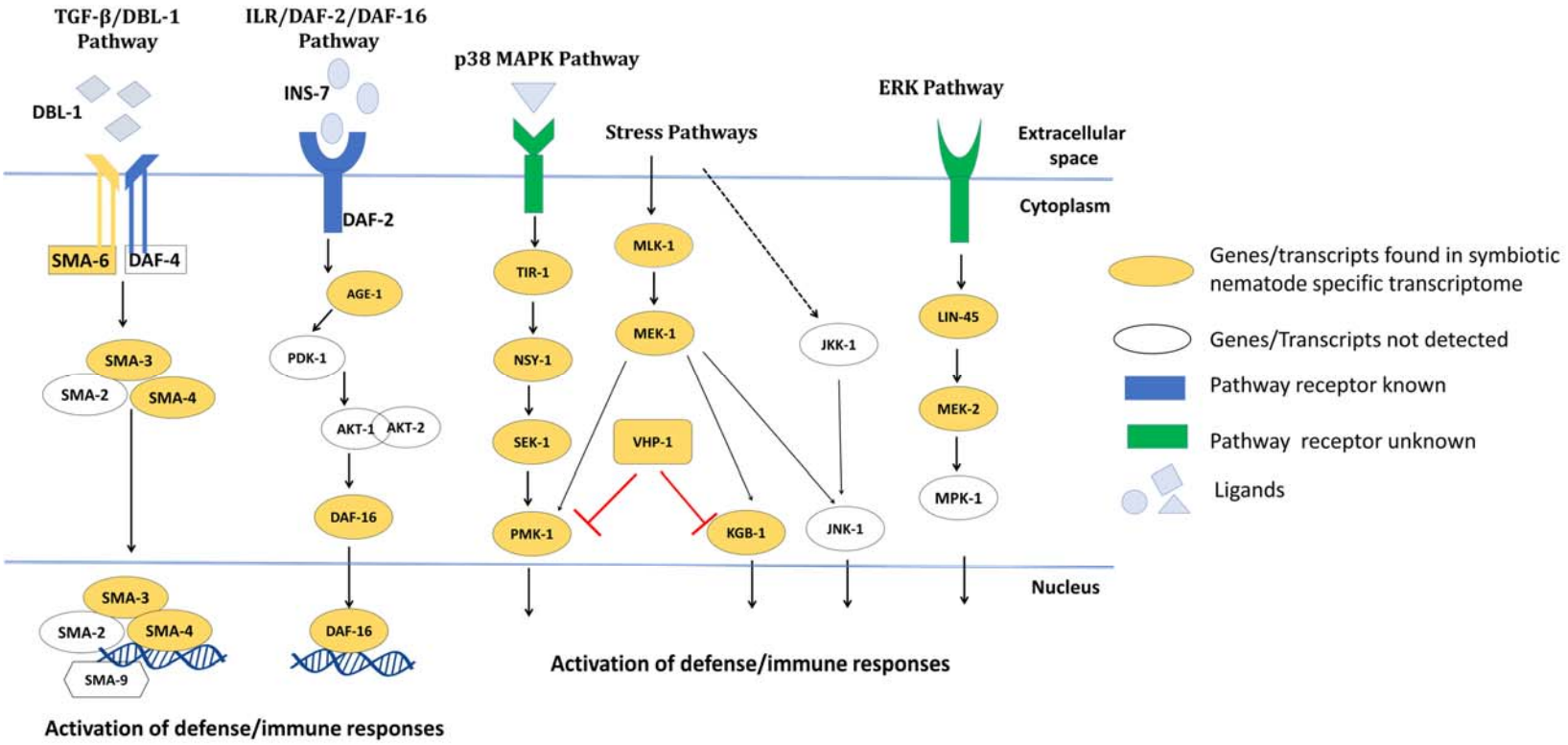


Figure 4.17: Orthologs of canonical nematode immune signalling pathway genes identified in the symbiotic *H. bacteriophora* transcriptome. Gene orthologs detected in the symbiotic nematode transcriptome data are highlighted in yellow color.

4.2.5 Validation of gene expression patterns by qRT-PCR

Expression of 27 randomly selected genes was validated by qRT-PCR. The expression pattern of 23 of the selected genes (*clec-87, dpy-5, col-19, mex-5, plx-2, WDR26, LRR20, IGSF9B, RPLP0, COII, ilys-3, age-1, sma-6, ced-3, ced-4, LBP, DMBT1, daf-16, tir-1, sma-3, daf-4, nsy-1* and *gsnl-1*) was congruent with the RNA-seq data (Figure 6). However, four genes - *ced-1*, *sma-4*, *daf-7* and *sek-1* which were found upregulated in RNA-seq, were found as baseline in qRT-PCR analysis (Figure 4.18).

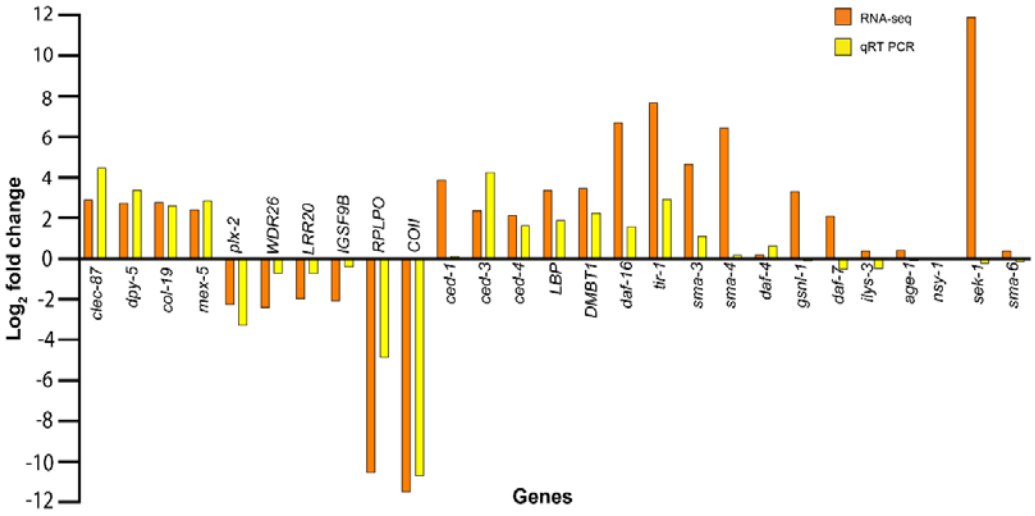


Figure 4.18. qRT PCR validation of expression patterns of the various differentially expressed and unique transcripts identified in RNA-seq experiment. Most of the tested transcripts conformed to their RNA-seq expression patterns.

Table 4.19: Most up- and down-regulated genes (with annotation) in symbiotic nematodes as compared to axenic nematodes

S.No.	Transcript ID	log2 Fold Change	*P*-value	RefSeq ID	UniProt ID: Protein name
	Down-regulated transcripts				
1. .	TRINITY_DN6295_c0_g10_i1--LEN=742	-11.50	1.12E-16	YP_817455.1	P24882: Cytochrome c oxidase subunit 2 (*Ascaris suum*)
2.	TRINITY_DN6151_c0_g3_i1--LEN=591	-10.73	6.41E-19	XP_002634177.1	A0A016RXA6: 60S ribosomal protein L18a (*Ancylostoma ceylanicum*)
3.	TRINITY_DN7450_c0_g2_i3--LEN=1246	-10.66	1.29E-15	XP_013309182.1	A0A0K0D9G2: Ribos_L4_asso_C domain-containing protein (*Angiostrongylus cantonensis*)
4.	TRINITY_DN3931_c0_g3_i1--LEN=584	-10.64	5.98E-20	XP_013299897.1	A0A0N4YQE1: 60S ribosomal protein L21 (*Nippostrongylus brasiliensis*)
5.	TRINITY_DN2790_c0_g2_i1--LEN=688	-10.61	5.24E-17	XP_013296741.1	A0A016UFB8: 60S ribosomal protein L18 (*Ancylostoma ceylanicum*)
6.	TRINITY_DN3587_c0_g1_i1--LEN=1046	-10.55	2.98E-15	XP_002639809.1	A0A016UDG4: 60S acidic ribosomal protein P0 (Ancylostoma ceylanicum)
7.	TRINITY_DN4706_c0_g2_i1--LEN=651	-10.54	1.46E-19	XP_002640009.1	A0A368GUV1: Ribosomal protein L19 (*Ancylostoma caninum*)
8.	TRINITY_DN4889_c0_g2_i6--LEN=1478	-10.54	2.87E-15	XP_013290581.1	A0A0K0DRD7: Uncharacterized protein (Angiostrongylus cantonensis)
9.	TRINITY_DN7104_c0_g6_i1--LEN=722	-10.51	2.04E-13	XP_013295799.1	A0A368G6S8: Elongation factor 1-alpha (*Ancylostoma caninum*)

10.	TRINITY_DN7267_c0_g4_i2--LEN=700	-10.51	5.93E-18	XP_013293533.1	A0A016TFX1: 60S ribosomal protein L6 (*Ancylostoma ceylanicum*)
	Up-regulated transcripts				
1.	TRINITY_DN1989_c0_g1_i3--LEN=902	3.94	1.15E-07	-	A0A368G1T3: Uncharacterized protein (*Ancylostoma caninum*)
2.	TRINITY_DN821_c0_g1_i1--LEN=940	3.93	4.74E-06	NP_504525.1	A0A0N4WJA0: Col_cuticle_N domain-containing protein (*Haemonchus placei*)
3.	TRINITY_DN4902_c0_g1_i3--LEN=1067	3.88	4.01E-08	-	A0A016T2Q5: Uncharacterized protein (*Ancylostoma ceylanicum*)
4.	TRINITY_DN3752_c0_g2_i1--LEN=593	3.63	0.001	-	B2FNK1: Uncharacterized protein (*Stenotrophomonas maltophilia*)
5.	TRINITY_DN6615_c0_g4_i2--LEN=732	3.51	7.8E-07	-	A0A368HAN4: Uncharacterized protein (*Ancylostoma caninum*)
6.	TRINITY_DN3811_c0_g1_i1--LEN=576	3.36	0.003	-	B2FLU5: Conserved hypothetical UptF protein (*Stenotrophomonas maltophilia*)
7.	TRINITY_DN4551_c0_g2_i3--LEN=723	3.32	0.003	-	Q21253: Gelsolin-like protein 1 (*Caenorhabditis elegans*)
8.	TRINITY_DN1989_c0_g1_i2--LEN=761	3.31	4.425E-05	-	A0A016SRK3: Uncharacterized protein (*Ancylostoma ceylanicum*)
9.	TRINITY_DN2337_c0_g1_i1--LEN=509	3.25	0.003	-	D8UFG0: Chloride channel protein (*Volvox carteri* f. *nagariensis*)
10.	TRINITY_DN8481_c0_g1_i1--LEN=787	3.23	0.0009	NP_001123166.1	G0N1Q9: Uncharacterized protein (*Caenorhabditis brenneri*)

Table 4.20: List of *Heterorhabditis* nematode transcripts relevant for bacterial symbiosis identified on the basis of Gene Ontology annotation

S. No.	Gene/Protein annotation (UniProt ID) [*]	Transcripts ID (Top match)	Expression value (FPKM>1)	S. No.	Gene/Protein annotation (UniProt ID)	Transcripts ID (Top match)	Expression value (FPKM>1)
	Lipid binding proteins (Immune surveillance and/or effectors)				**Apoptosis, Autophagy and Endocytosis**		
1	BPI (P17453) [1]	TRINITY_DN3736_c0_g2_i1--LEN=968	3.15	40	*ced-1* (A8XMW6) [2]	TRINITY_DN7259_c0_g1_i17--LEN=3809	4.08
2	LBP (Q2TBI0) [1]	TRINITY_DN5483_c0_g1_i3--LEN=1784	3.37	41	*ced-3* (P42573) [2]	TRINITY_DN5119_c0_g3_i1--LEN=872	2.59
3	DMBT1 (Q9UGM3) [1]	TRINITY_DN7168_c2_g5_i1--LEN=832	3.47	42	*ced-4* (Q60Z52) [4]	TRINITY_DN7359_c0_g1_i14--LEN=873	3.86
4	CD36 (P26201) [1]	TRINITY_DN3374_c0_g1_i2--LEN=1346	10.7	43	RNF216 (Q9NWF9) [1]	TRINITY_DN7732_c0_g6_i1--LEN=516	2.7
	Carbohydrate binding proteins (Immune surveillance and/or effectors)			44	RAB14 (Q5ZKU5) [1]	TRINITY_DN4406_c0_g1_i1--LEN=558	2.08
5	*lec-8* (Q09610) [1]	TRINITY_DN3367_c0_g2_i1--LEN=663	13.3	45	TIAL1 (Q01085) [1]	TRINITY_DN4648_c0_g1_i4--LEN=1490	4.53
6	*irg-7* (A0A131MBU3) [9]	TRINITY_DN5710_c0_g2_i4--LEN=793	85.08	46	ACD11 (O64587) [2]	TRINITY_DN7077_c0_g2_i7--LEN=2155	12.93
7	CLEC4A (Q9UMR7) [1]	TRINITY_DN7631_c0_g3_i8--LEN=569	3.27	47	BCL3 (P20749) [2]	TRINITY_DN7233_c0_g3_i1--LEN=1494	2.92
8	LGALS9 (Q3MHZ8) [1]	TRINITY_DN6177_c1_g3_i9--LEN=1440	6.51	48	SKP2 (Q13309) [1]	TRINITY_DN8007_c0_g2_i1--LEN=698	2.96

9	LY75 (O60449) [2]	TRINITY_DN7631_c0_g3_i1--LEN=1056	19.79	49	*atg-16.2* (Q09406) [3]	TRINITY_DN8112_c0_g1_i13--LEN=4799	7.43
10	CHID1 (Q5EAB4) [2]	TRINITY_DN1976_c0_g1_i1--LEN=1263	4.71	50	*lgg-2* (Q23536) [1]	TRINITY_DN7243_c0_g2_i6--LEN=1207	10.39
	Kinases, Phosphatases and other protein adaptors (Signal transduction)			51	SYT11 (Q9BT88) [1]	TRINITY_DN7833_c1_g1_i23--LEN=3591	13.08
11	*abl-1* (P03949) [2]	TRINITY_DN6235_c0_g1_i1--LEN=4462	8.17	52	SQSTM1 (Q13501) [3]	TRINITY_DN5917_c0_g2_i1--LEN=1881	13.71
12	*kgb-1* (O44408) [2]	TRINITY_DN5181_c1_g1_i16--LEN=1829	12.88	53	RUBCN (Q92622) [1]	TRINITY_DN6053_c0_g1_i1--LEN=2907	8.39
13	*mlk-1* (A0A0K3AV08) [5]	TRINITY_DN7254_c0_g1_i31--LEN=3797	4.82	54	ILRUN (Q5F3N9) [1]	TRINITY_DN5070_c0_g2_i1—LEN=1357	7.37
14	*sek-1* (G5EDF7) [1]	TRINITY_DN6553_c0_g1_i10--LEN=1921	11.9	55	C3 (Q2UVX4) [1]	TRINITY_DN6795_c1_g2_i1--LEN=1263	2.42
15	*bsk* (P92208) [1]	TRINITY_DN4022_c0_g1_i2--LEN=1669	6.28	56	faf (A0A0B4K7S0) [1]	TRINITY_DN8152_c2_g1_i3--LEN=635	2.36
16	*sma-6* (Q09488) [3]	TRINITY_DN8152_c1_g1_i4--LEN=1147	11.88	57	WASL (O00401) [1]	TRINITY_DN6405_c0_g2_i3--LEN=633	3.8
17	*dkf-2* (O45818) [1]	TRINITY_DN7345_c0_g1_i3--LEN=1337	4.65	58	DENND1B (Q6P3S1) [2]	TRINITY_DN5933_c0_g2_i8--LEN=2315	4.82
18	*dapk-1* (O44997) [1]	TRINITY_DN8019_c0_g2_i13--LEN=5058	6.85		**Lysozymes**		
19	*ksr-1* (G5EFD2) [1]	TRINITY_DN7095_c0_g2_i11--LEN=2109	2.23	59	*ilys-2* (O76358) [2]	TRINITY_DN6977_c0_g2_i7--LEN=627	62.75
20	*pkc-3* (A8WUG4) [3]	TRINITY_DN7906_c0_g2_i1--LEN=5269	13.61	60	*ilys-3* (O76357) [1]	TRINITY_DN6977_c0_g2_i9--LEN=1285	51.59

21	MAPKBP1 (O60336) [1]	TRINITY_DN7860_c1_g2_i1--LEN=5202	5.71		**Peroxidases**		
21	*vhp-1* (Q10038) [3]	TRINITY_DN7713_c0_g1_i23--LEN=2827	18.84	61	*bli-3* (O61213) [3]	TRINITY_DN7613_c0_g1_i19--LEN=5208	4.8
23	PPM1D (O15297) [1]	TRINITY_DN6457_c0_g7_i1--LEN=3295	7.97	62	*hpx-2* (P90820) [2]	TRINITY_DN7546_c0_g1_i24--LEN=1757	9.7
24	*tir-1* (Q86DA5) [1]	TRINITY_DN7733_c0_g4_i1--LEN=3597	7.68	63	*skpo-1* (Q20616) [1]	TRINITY_DN6978_c0_g3_i1--LEN=719	2.54
25	*sma-3* (P45896) [3]	TRINITY_DN6489_c0_g1_i2--LEN=1842	5.67	64	*gst-5* (Q09596) [3]	TRINITY_DN6328_c1_g2_i2--LEN=692	186.58
26	*daf-7* (P92172) [3]	TRINITY_DN7805_c0_g1_i21--LEN=3469	7.33	65	*gstk-1* (Q09652) [1]	TRINITY_DN7642_c1_g2_i1--LEN=716	6.77
27	SMAD3 (P84022) [1]	TRINITY_DN7761_c0_g1_i13--LEN=2875	20.57	66	GLO-3 (Q24JJ8) [1]	TRINITY_DN4173_c0_g2_i1--LEN=1815	2.57
	Transcription factors				**Transmembrane transporters**		
28	*daf-16* (O16850) [2]	TRINITY_DN7792_c0_g1_i15--LEN=2589	11.92	67	*pgp-1* (P34712) [1]	TRINITY_DN7508_c0_g2_i8--LEN=4768	8.8
29	*elt-2* (Q10655) [1]	TRINITY_DN5192_c0_g2_i2--LEN=2006	8.51	68	*pgp-3* (P34713) [4]	TRINITY_DN6803_c0_g2_i2--LEN=4096	6.99
30	*pnr* (P52168) [1]	TRINITY_DN6774_c0_g3_i2--LEN=1137	7.33	69	SLC17A5 (Q9NRA2) [9]	TRINITY_DN4851_c0_g2_i4--LEN=2169	9.83
31	*fos-1* (G5ECG2) [1]	TRINITY_DN7459_c0_g2_i8--LEN=2687	2.47	70	*aqp-10* (Q09369) [5]	TRINITY_DN3980_c0_g2_i4--LEN=1379	17.26
32	TFEB (P19484) [1]	TRINITY_DN8141_c1_g1_i9--LEN=3647	2.32	71	AQP3 (Q08DE6) [3]	TRINITY_DN4537_c0_g1_i1--LEN=1340	7.7

33	PAX5 (Q02548) [1]	TRINITY_DN1903_c0_g2_i1--LEN=865	6.58		**Mucosa associated protein**		
	Phospholipases (lipolytic)			72	MALT1 (Q9UDY8) [1]	TRINITY_DN5749_c0_g1_i29--LEN=3385	4.71
34	PLA2G1B (P00593) [1]	TRINITY_DN2774_c0_g1_i1--LEN=665	17.14	73	*Cad99C* (Q9VAF5) [1]	TRINITY_DN5774_c0_g2_i2--LEN=2550	3.3
35	PLA2G6 (O60733) [4]	TRINITY_DN4942_c0_g2_i3--LEN=3202	3.4	74	MR1 (C1ITJ8) [1]	TRINITY_DN6683_c0_g1_i17--LEN=5161	2.55
36	*plc-1* (G5EFI8) [2]	TRINITY_DN8254_c0_g2_i32--LEN=5949	3.22		Antimicrobial proteins		
	Metallopeptidases (proteolytic)			75	*acantho1* (Q8I948) [1]	TRINITY_DN5834_c0_g2_i1--LEN=3199	23.65
37	*spg-7* (Q9N3T5) [1]	TRINITY_DN6871_c0_g1_i2--LEN=2590	16.2	76	Antimicrobial protein Ace-AMP1 (Q41258) [2]	TRINITY_DN4773_c1_g1_i5--LEN=1481	5.57
38	*zmp-2 (*O44836) [1]	TRINITY_DN8041_c0_g1_i2--LEN=2285	11.6	77	Venom serine protease inhibitor (A0A2R4SV19) [5]	TRINITY_DN4538_c0_g1_i2--LEN=1065	89.02
39	ADAM8 (P78325) [1]	TRINITY_DN5445_c0_g2_i1--LEN=1466	3.47	78	*psidin* (Q17DK2) [2]	TRINITY_DN8064_c0_g2_i7--LEN=2622	7.08

4.3 Standardization of RNAi-soaking methodology for gene function validation in *Heterorhabditis*

4.3.1 Core RNAi pathway components in *Heterorhabditis bacteriophora* relative to *Caenorhabditis elegans*

The two of the predicted proteome of *H. bacteriophora* [*H. bacteriophora* strain TTO1 (BioProject PRJNA13977, Bai et al, 2013) and *H. bacteriophora* strain G2a1223 (McLean et al, 2018)] were analyzed to identify the RNAi machinery components. A total of 59 and 60 orthologues of RNAi pathway genes were detected in HbTTO1 and HbG2a1223 respectively (Table 4.21). Orthologs of *rsd-2, rsd-6, sid-2, eri-3, lin-15b, ekl-5, mes-3, mut-16,* and *rde-2* were not detected in either of the examined proteomes, even with a very flexible criteria of e value 0 and bit score >40. Furthermore, *pash-1* was not found in predicted proteome of *H. bacteriophora* TTO1.

Table 4.21: RNAi pathway genes detected in *H. bacteriophora* genome in comparison to *C. elegans*

S.No.	Gene name	Query ID	Subject ID HbTTO1	Bit score	Subject ID HbG2a1223	Bit score
	Small RNA Biosynthesis					
1.	*dcr-1*	K12H4.8.1	Hba_20705	940	g1407.t1	2023
2.	*drh-1*	F15B10.2a.1	Hba_20705	68.9	g7737.t1	77.4
3.	*drh-3*	D2005.5.1	Hba_07956	107	g7737.t1	271
4.	*drsh-1*	F26E4.10a.1	Hba_01080	205	g8917.t1	721
5.	*pash-1*	T22A3.5a.1	-	-	g9560.t1	290
6.	*rde-4*	T20G5.11.1	Hba_14605	48.9	g6603.t1	89
7.	*xpo-1*	ZK742.1a.1	Hba_11041	501	g6048.t1	1577
8.	*xpo-2*	Y48G1A.5a.1	Hba_15432	660	g3302.t1	1090
9.	*xpo-3*	C49H3.10a.1	Hba_15445	693	g3316.t1	612
	dsRNA uptake and spreading					
10.	*rsd-3*	C34E11.1.1	Hba_08667	202	g10488.t1	202
11.	*sid-1*	C04F5.1.1	Hba_07262	108	g12689.t1	63.2
	siRNA secondary amplification					
12.	*ego-1*	F26A3.3.1	Hba_03302	174	g13725.t1	996
13.	*rrf-1*	F26A3.8a.1	Hba_03302	161	g9244.t1	855
14.	*rrf-3*	F10B5.7.1	Hba_03302	76.6	g9244.t1	381
15.	*smg-2*	Y48G8AL.6.1	Hba_16583	386	g7388.t1	1139
16.	*smg-5*	W02D3.8.1	Hba_00596	63.9	g8802.t1	103
17.	*smg-6*	Y54F10AL.2a.1	Hba_20514	328	g1239.t1	360
	RISC-argonautes					
18.	*alg-1*	F48F7.1a.1	Hba_15995	136	g10773.t1	1704
19.	*alg-2*	T07D3.7a.1	Hba_15995	132	g10773.t1	1589

20.	*alg-3*	T22B3.2a.1	Hba_15995	103	g10773.t1	453
21.	*alg-4*	ZK757.3a.1	Hba_15995	103	g10773.t1	452
22.		C04F12.1.1	Hba_07017	72	g4466.t1	178
23.		C16C10.3.1	Hba_21004	63.5	g980.t1	248
24.	*csr-1*	F20D12.1a.1	Hba_07017	63.2	g11026.t1	181
25.	*ergo-1*	R09A1.1a.1	Hba_15995	63.5	g10773.t1	189
26.		F58G1.1.1	Hba_07017	74.7	g14175.t1	534
27.	*nrde-3*	R04A9.2.1	Hba_21004	51.6	g980.t1	223
28.	*ppw-1*	C18E3.7c.1	Hba_19411	75.1	g4466.t1	348
29.	*ppw-2*	Y110A7A.18.1	Hba_07017	72	g14175.t1	510
30.	*prg-1*	D2030.6.1	Hba_15995	786	g5317.t1	394
31.		R06C7.1.1	Hba_07017	77.4	g14175.t1	586
32.	*rde-1*	K08H10.7.1	Hba_17351	91.3	g2058.t1	300
33.	*sago-1*	K12B6.1.1	Hba_19410	58.9	g4466.t1	330
34.	*sago-2*	F56A6.1a.1	Hba_19411	77.4	g4466.t1	350
35.		T22H9.3.1	Hba_21004	64.3	g980.t1	238
36.		T23D8.7.1	Hba_15995	58.9	g10773.t1	573
37.		Y49F6A.1.1	Hba_21004	63.2	g980.t1	255
38.		ZK1248.7.1	Hba_07017	77	g14175.t1	590
	RISC components					
39.	*ain-1*	C06G1.4.1	Hba_18316	231	g4761.t1	300
40.	*ain-2*	B0041.2a.1	Hba_18316	68.9	g4761.t1	93.6
41.	*tsn-1*	F10G7.2.1	Hba_19639	355	g3035.t1	1153
42.	*vig-1*	F56D12.5a.1	Hba_14279	144	g5790.t1	145
	RNAi inhibitors					
43.	*adr-1*	H15N14.1c.1	Hba_17025	471	g3217.t1	570
44.	*adr-2*	T20H4.4.1	Hba_17025	45.4	g3217.t1	82.4
45.	*eri-1*	T07A9.5a.1	Hba_20423	97.4	g1849.t1	127
46.	*eri-5*	Y38F2AR.1a.1	Hba_15753	112	g5492.t1	140
47.	*eri-6/7*	C41D11.7.1	Hba_11497	101	g7388.t1	124
48.	*xrn-1*	Y39G8C.1.1	Hba_14852	340	g5533.t1	939
49.	*xrn-2*	Y48B6A.3.1	Hba_14852	212	g9741.t1	1206
	Nuclear effectors					
50.	*cid-1*	K10D2.3.1	Hba_14774	251	g7774.t1	342
51.	*ekl-1*	F22D6.6.1	Hba_15753	179	g5492.t1	314
52.	*ekl-4*	Y105E8A.17a.1	Hba_18987	79.3	g3660.t1	494
53.	*ekl-6*	T16G12.5.1	Hba_17757	160	g2835.t1	159
54.	*gfl-1*	M04B2.3.1	Hba_14500	304	g4665.t1	304
55.	*mes-2*	R06A4.7.1	Hba_14023	88.6	g5811.t1	244
56.	*mes-6*	C09G4.5.1	Hba_15773	55.8	g5512.t1	153
57.	*mut-2*	K04F10.6a.1	Hba_14774	68.9	g1991.t1	127
58.	*mut-7*	ZK1098.8.1	Hba_09387	130	g10200.t2	173
59.	*rha-1*	T07D4.3a.1	Hba_06990	817	g10416.t1	1377
60.	*zfp-1*	F54F2.2a.1	Hba_20827	348	g1532.t1	356

4.3.2 RNAi by soaking of the post-embryonic stage of *H. bacteriophora*

RNAi by soaking of the post-embryonic stage in dsRNA solution was proved in *H. bacteriophora* (Ciche and Sternberg, 2007). As recommended eggs in the pretzel stage of embryonic development were used. When eggs were used for RNAi by soaking, we observed poor, unsynchronized hatch and higher mortality of hatched juveniles. Optimization of soaking protocols suggested that *Heterorhabditis* eggs can hatch efficiently and synchronously if they are provided with proper aeration. Eppendorf tubes are not suitable for hatching but 24 well plates with flat to round bottom providing more surface area are efficient. Despite the proper egg hatching, higher mortality of hatched juveniles was observed in the soaking solution. It was an impediment in drawing the correlation between the treatment and phenotypic effects. We observed that the L1s coming out of eggs were very delicate and not suitable for prolonged soaking.

4.3.3 Screening of *Heterorhabditis* developmental stages for dsRNA uptake using FITC as a marker of uptake

Infective juveniles (IJs): The infective juveniles (IJs) of the *Heterorhabditis* spp. are developmentally arrested non-feeding third stage juveniles. It is the only free-living stage in the nematode life cycle. IJ stage is similar to the dauer stage of the *C. elegans*, and the infective L3 stage of many animal-parasitic nematodes. The IJs showed no fluorescence despite 16-24 h of incubation in FITC solution. There was slight FITC uptake in the buccal cavity when IJs were soaked in the presence of a neuro-stimulant, but not beyond the oesophagus-intestinal valve (Figure 4.19). To confirm the route of entry of FITC molecules into the IJs body i.e., to check for any trans-cuticular absorption, cephalic openings were occluded using cyanoacrylate tissue adhesive and then soaked in FITC. No fluorescence was observed in such IJs.

Desiccation prior to soaking was tested to see if it enhances FITC uptake in IJs. Nematodes were desiccated using saturated potassium sulfate solution (reducing relative humidity to 97% at 20 °C) and then rehydrated with the soaking buffer containing 1 mg/ml FITC+ neuro-stimulant. When desiccated IJs were rehydrated, less than 50% of nematodes recovered back to normal activity. Rehydrated IJs showed FITC absorption within 6-7s hours of rehydration, when immersed in FITC. However, only 50-65 % of rehydrated IJs showed FITC uptake (Figure 4.20).

Post-IJ-Recovery stages (PIJR): Nematodes come out of the dauer stage and continue their development. PIJR stages are feeding stages. We considered 3 PIJR stages i.e., L4 (10-12 h post-recovery) and early adult (36-40 h post-recovery) and adult hermaphrodite (108-116 h post-recovery). Enhanced uptake of FITC was observed in post IJ recovery stages both in the presence as well as in the absence of neuro-stimulant (Figure 4.21). Strong fluorescence was observed in the intestinal tract of L4 and early adult stages. Hatching of juveniles was observed when adult hermaphrodites containing eggs were soaked. So FITC did not affect the hatching physiology of this nematode. Juveniles hatching out of the mother nematode also showed strong FITC uptake. However, the viability of hatched juveniles was affected within 16 h of soaking (Figure 4.22).

In order to note the variation between species in the incorporation of FITC, the same experiments were conducted on *H. indica* IJ and PIJR developmental stages. *H. indica* IJS showed no fluorescence even after an extended period of soaking in FITC for more than 24h. Post IJ recovery stage (12 h post-recovery stage) showed fluorescence (Figure 4.23). We compared the FITC uptake in the IJ stage of another EPN *Steinernema*. There was strong uptake of FITC by the infective juveniles of *Steinernema*. *Steinernema* IJs exhibited trans cuticular absorption of FITC which was further enhanced by octopamine (Figure 4.24).

Based on the above observations, the PIJR stage (10-12 h post-recovery stage) of *H. bacteriophora* was considered for further RNAi-soaking experiments, as this stage can uptake FITC strongly, and more uniformly as compared to IJs and is sturdier as compared to adult/L1s.

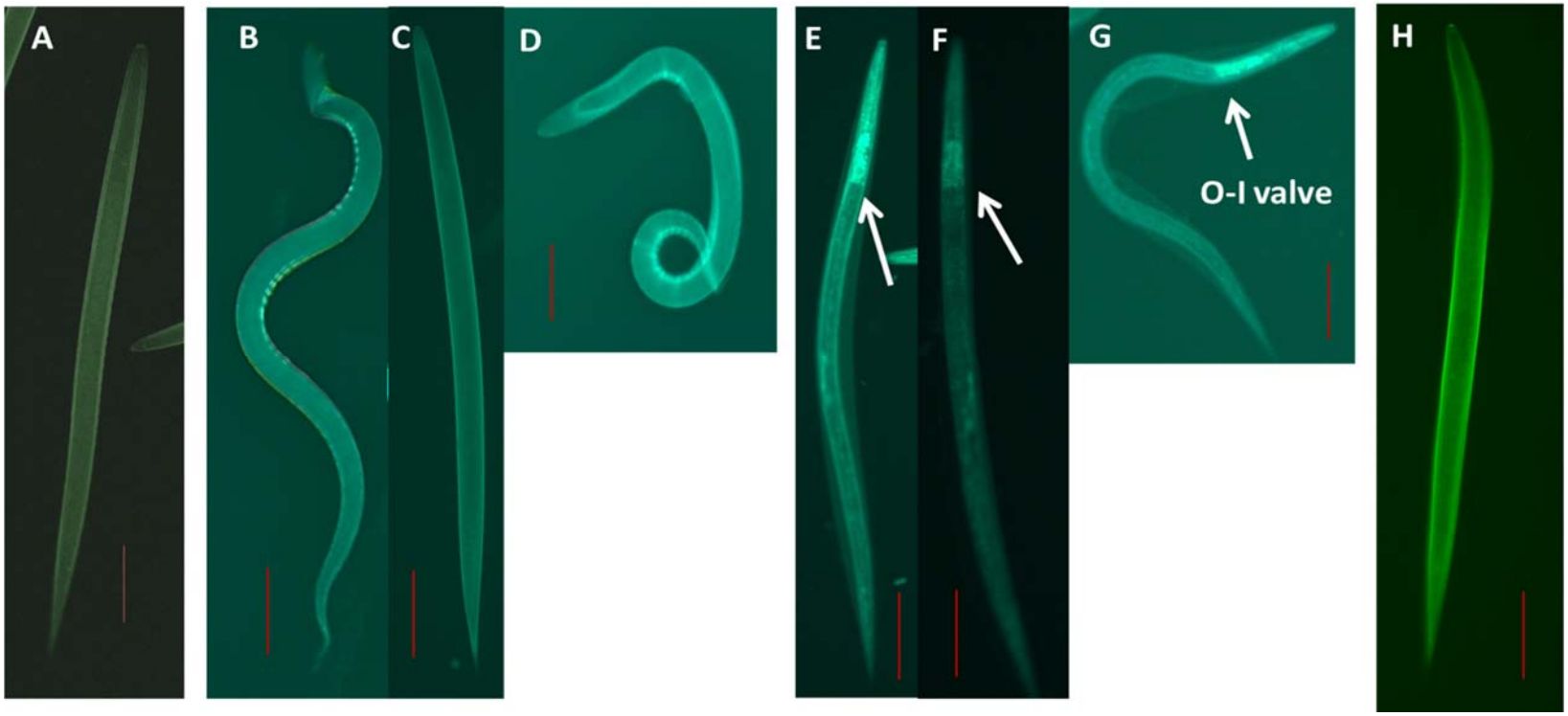


Figure 4.19: Screening of FITC uptake in the infective juveniles of *H. bacteriophora.* A: Control (nematode soaked in M9 buffer); B-D: IJs soaked in FITC solution (1mg/ml) without neuro-stimulant showed no uptake of FITC; E-G: IJs soaked in FITC solution with neuro-stimulant (1mg/ml FITC + 50 mM octopamine) showed slight uptake of FITC only up to oesophagus-intestinal (O-I) valve; H: IJ with occluded cephalic openings showed no FITC uptake.

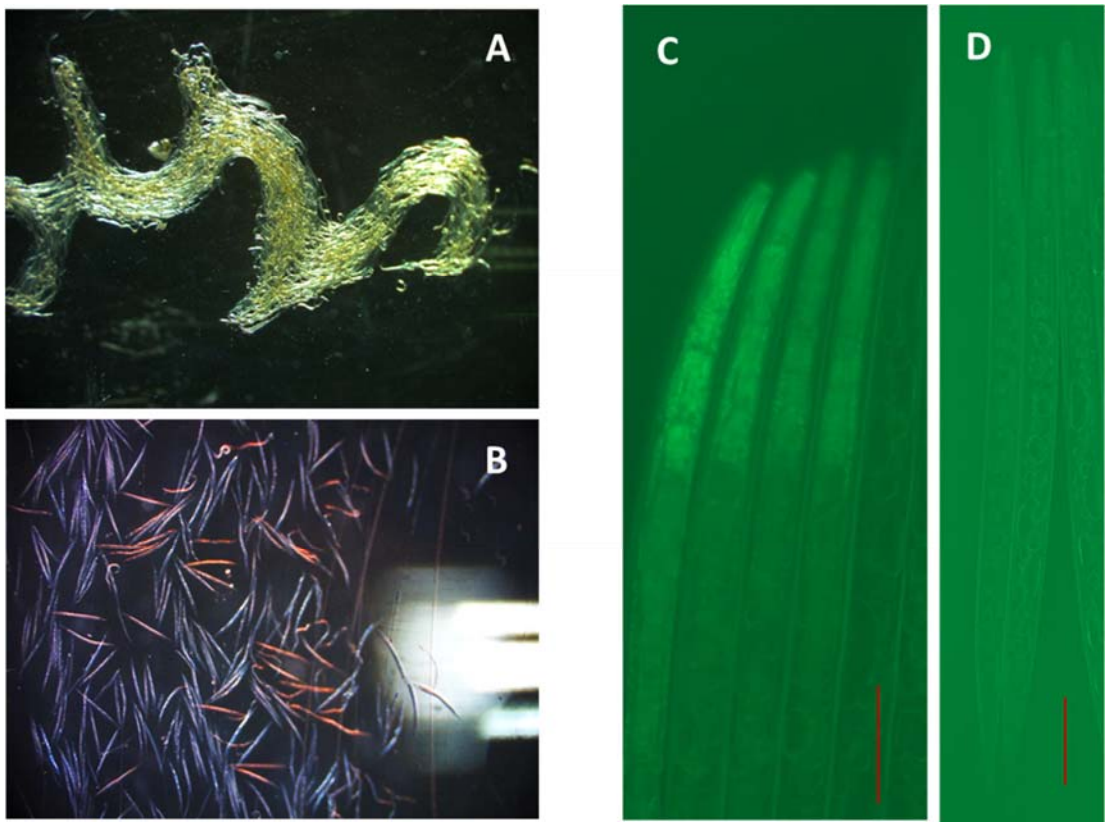


Figure 4.20: Monitoring of FITC uptake in *H. bacteriophora* IJs using desiccation prior to soaking protocol. **A:** Desiccation of *H. bacteriophora* IJs before soaking; B: Rehydrating the IJs in the soaking solution containing FITC; C: Rehydrating IJs showed gradual uptake of FITC; D: Control

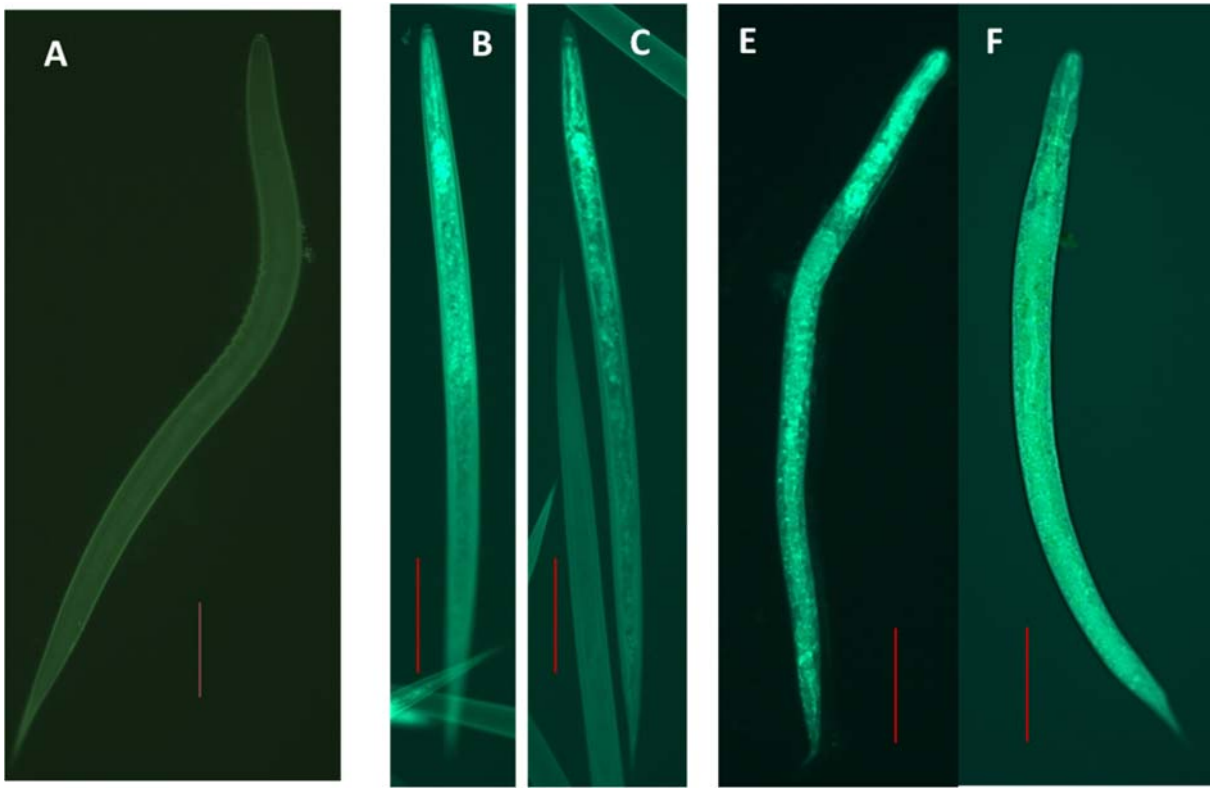


Figure 4.21: Screening of FITC uptake in post-IJ recovery stages of *H. bacteriophora*. A: Control; B-C: L4s (10-12 h post recovery) showed strong uptake of FITC; E-F: Early adults (36-40 h post recovery) showed strong uptake of FITC.

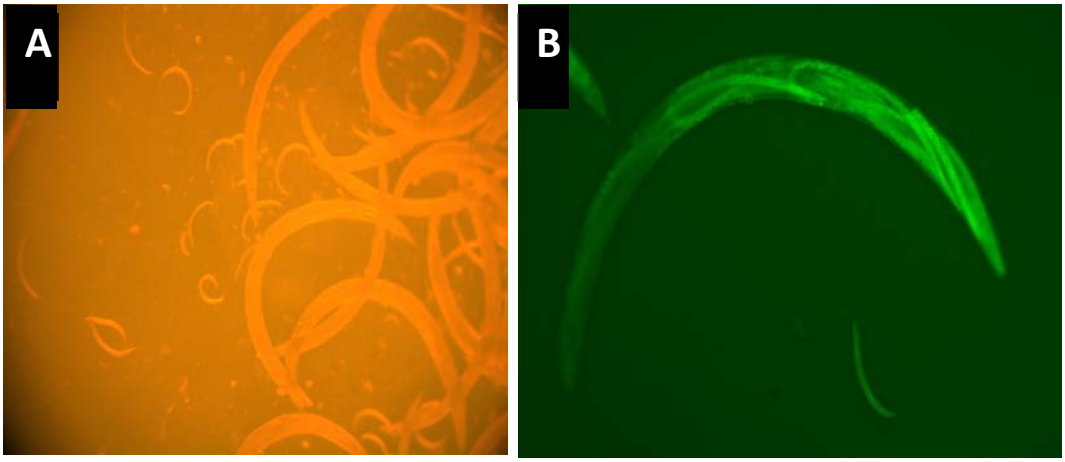


Figure 4.22: Screening of FITC uptake in adult hermaphrodites of *H. bacteriophora.* Hatching of juveniles was observed when adult hermaphrodites containing eggs were soaked. The viability of hatched juveniles were affected.

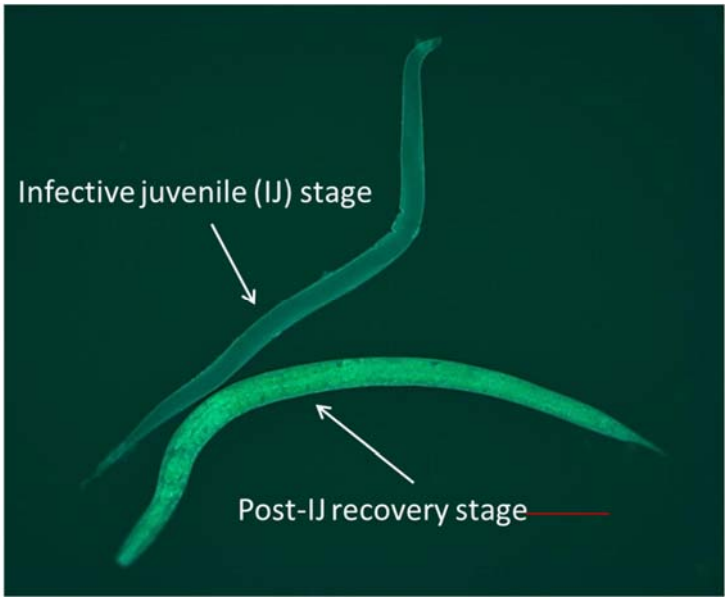


Figure 4.23: Screening of FITC uptake in IJ and PIJR stages of *H. indica*. IJs showed no uptake of FITC, post-IJ recovery stage showed strong uptake of FITC.

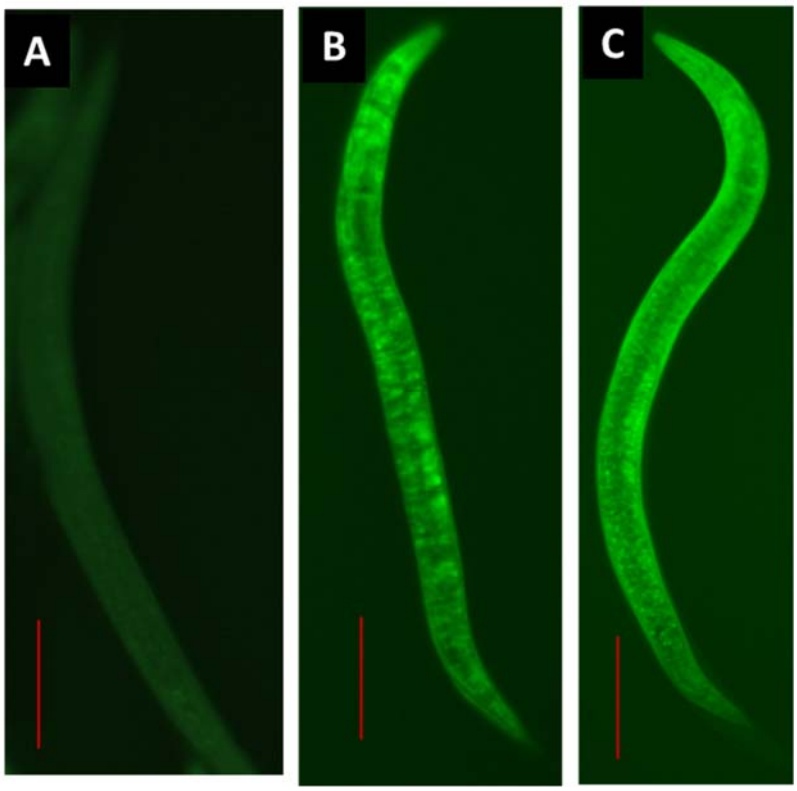


Figure 4.24: Screening of FITC uptake in the infective juveniles of *S. abbasi*. A: Control; B: *S. abbasi* IJ soaked in FITC without neuro-stimulant; C: *S. abbasi* IJ soaked in FITC with neuro stimulant. *Steinernema* IJs showed strong transcuticular absorption of FITC.

4.3.4 Investigating RNAi-soaking in PIJR stage: phenotypic analysis

Genes *Hb-dpy-7*, *Hb-dpy-13* and *Hb-cct-2* which have proven RNAi phenotype in *H. bacteriophora* nematode were selected for validation of dsRNA soaking in PIJR stage. Nematodes were soaked for 24 h in dsRNA solution and then transferred to the lawns of symbiont bacteria and observed every 24 h for 5 days after soaking.

Worms treated with dsRNA of *Hb-dpy-13* showed a significant reduction in body length as compared to control worms and 76-86 % penetrance was observed in three replicates (Table 4.22). Average penetrance was around 57 % in ds-*Hb-dpy-7* treated worms. Phenotyping of worms treated with dsRNA of dumpy genes versus wild type hermaphrodite as observed through stereo zoom microscope, 48 h post soaking is shown in figure 4.25. Wild-type hermaphrodites were approximately 2.2 mm in length and the treated worms were approximately 0.7 mm and fragile in nature. ds-*Hb-cct-2* RNA treated worms were phenotyped 55-60 h after soaking. A significant decrease in larval hatching was observed in the case of treated worms. There were very few to no

eggs in the observed worms (average 70 % penetrance). Control worms showed normal gonad development and egg-laying in 4-5 days (Figure 4.26).

Figure 4.25: Phenotyping of *H. bacteriophora* PIJR nematodes soaked in dsRNA of *Hb-dpy-13* and *Hb-dpy-7* genes. Worms treated with dsRNA of dumpy genes showed reduced body length as compared to wild-type worms.

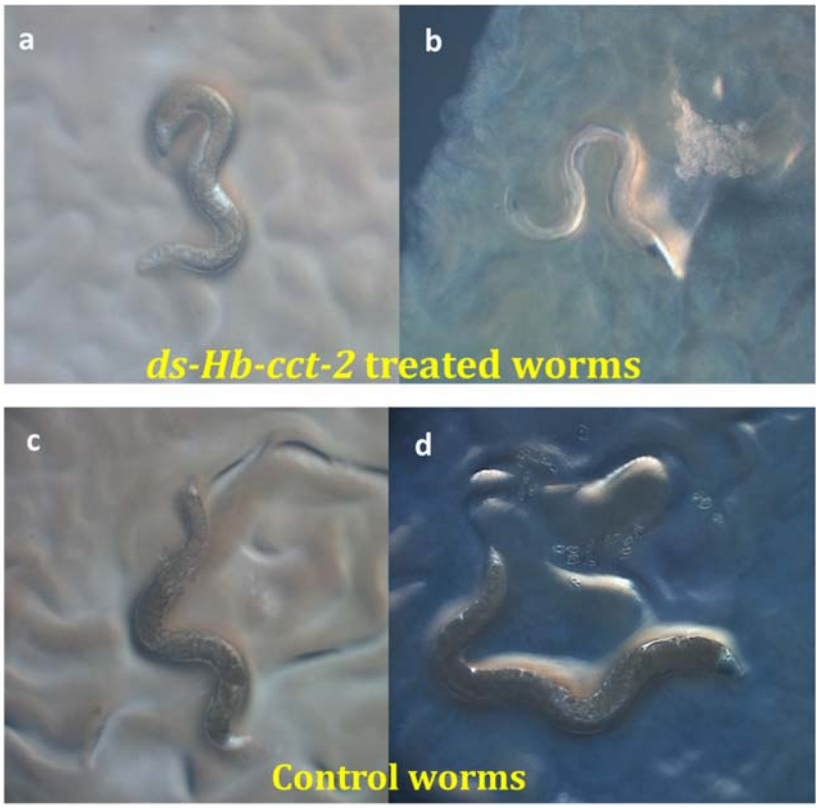


Figure 4.26: Phenotyping of *H. bacteriophora* PIJR nematodes soaked in dsRNA of *Hb-cct-2* gene. Gonad development was affected in the worms treated with *ds-Hb-cct-2* gene. Normal gonad development and proper egg laying were observed in control worms.

4.3.5 Investigating RNAi-soaking in PIJR stage: gene expression analysis by using qRT-PCR

The expression of selected genes was quantitated in dsRNA treated versus nontreated worms by RT-qPCR. In all the cases gene expression was significantly ($P<0.05$) down-regulated in dsRNA treated worms. Expression of *Hb-dpy-7*, *Hb-dpy-13* and *Hb-cct-2* was downregulated in dsRNA treated worms by -4.86, -6.76, and -7.46 folds respectively (Figure 4.27).

Table 4.22: RNAi phenotyping in *H. bacteriophora* worms, 48-60 h after soaking of post-IJ recovery stages in dsRNA solution

	Trail I			Trail II			Trail III					
	Number of worms exhibiting distinguishable phenotype	Number of worms analysed	% Penetrance	Number of worms exhibiting distinguishable phenotype	Number of worms analysed	% Penetrance	Number of worms exhibiting distinguishable phenotype	Number of worms analysed	% Penetrance	Average Number of worms exhibiting distinguishable phenotype	Average Number of worms analysed	% Penetrance
Hb-dpy-7	14	25	56	10	18	55.55	13	22	59.09	37	65	56.92
Hb-dpy-13	19	22	86.36	22	27	81.48	23	30	76.66	64	79	81.01
Hb-cct-2	15	20	75	13	20	65	11	15	73.33	39	55	70.90

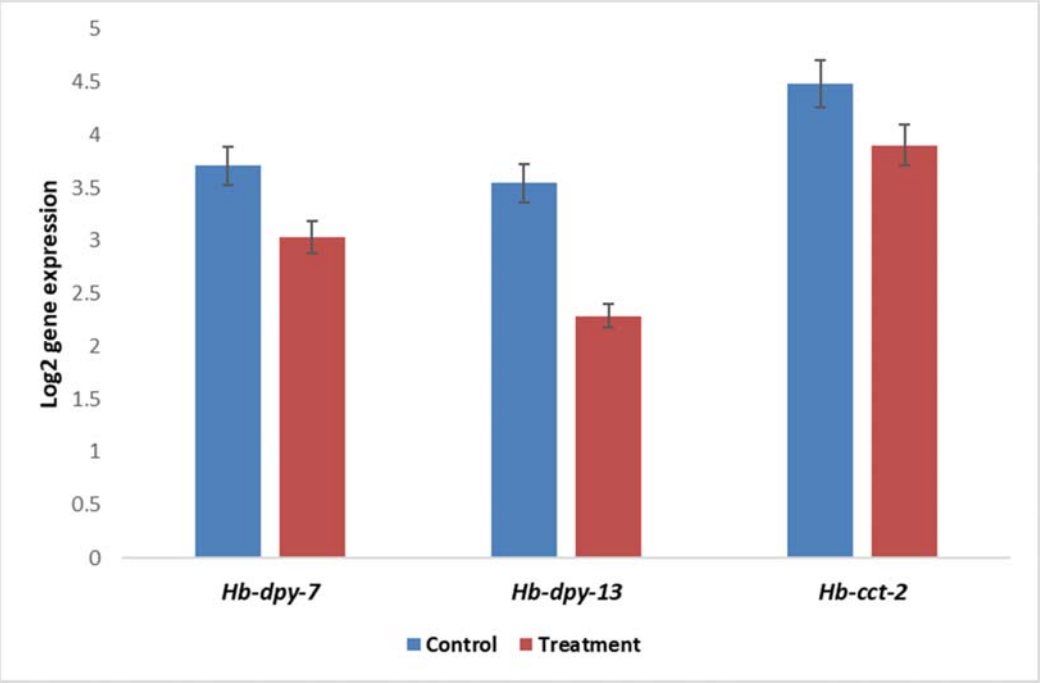


Figure 4.27: Target gene expression quantification in dsRNA treated versus nontreated *H. bacteriophora* worms by RT-qPCR. Expression of target genes were significantly ($P<0.05$) down-regulated in dsRNA treated worms

5.1 The draft genome of *Heterorhabditis indica*

Entomopathogenic nematodes (EPNs) are a distinct category of insect-parasitic nematodes which have evolved mutualistic association with entomopathogenic bacteria. EPNs are capable of killing the insect hosts rapidly with the aid of bacterial partners and they pass on their associated bacteria to offspring (Dillman et al., 2012). Nematodes belonging to Steinernematidae and Heterorhabitidae exhibit mutualistic association with insect-pathogenic *Xenorhabdus* and *Photorhabdus* respectively and are extensively studied EPN families. Twenty-one *Heterorhabditis* and one hundred *Steinernema* species have been described from various parts of the world (Bhat et al., 2020). Recently some species of *Oscheius* belonging to Rhabditidae family have been shown to use the symbiotic relationship with insect-pathogenic bacteria like *Serratia* to parasitize insect hosts and are categorized under EPN category. Around eleven species out of 45 reported species of *Oscheius* are described as EPNs (Al-Zaidawi et al., 2019; Kumar et al., 2019; Zhang et al., 2019). *Heterorhabditis* and *Oscheius* nematodes belong to the clade 9 and *Steinernema* belongs to clade 10 (Holterman et al., 2006). EPNs have been successfully used as biological control agents to manage important insect pests, especially soil-dwelling and/or plant-boring insect pests. Some of these nematodes are produced commercially in several countries and utilized in crop protection programs (Lacey et al., 2015; Lacey and Georgis, 2012). These EPNs are also an excellent genetically tractable model for studying mutualism and parasitism, as well as ecological adaptations and models for biological screening of anthelmintics (Barrón-Bravo et al., 2021; Campos–Herrera et al., 2012). However, their full potential as bio-control agents and as a model system remains under-exploited.

Omics information such as genome and transcriptome data and genome interrogation and editing techniques such as RNAi and CRISPR-Cas9 are potent methods to explore EPN biology and lay the groundwork for improving their bio-control traits (Lu et al., 2016). Genomic information about EPNs is scant in the public domain. Prior to this work, whole-genome information was available for only one heterorhabditid (*H. bacteriophora*), seven steinernematid (*S. carpocapsae, S. scapterisci, S. feltiae, S. glaseri, S. monticolum, S. diaprepesi, S. khuongi)* and one *Oscheius* (*Oscheius* sp. TEL-2014) nematode in the public domains such as NCBI and

Wormbase ParaSite (Bai et al., 2013; Baniya and DiGennaro, 2021; Dillman et al., 2015; Lephoto et al., 2016; Rougon-Cardoso et al., 2016; Serra et al., 2019). The size of the published EPN genomes vary from 77 Mb (*H. bacteriophora*) to 121.6 Mb (*S. feltiae* strain NW), and the number of protein-coding gene sequences from 21,250 (*H. bacteriophora*) to 36,007 (*S. monticolum*) (Baniya and DiGennaro, 2021; Lu et al., 2016). Mitochondrial genomes or mitogenome information is available for *H. bacteriophora* (18,128 bp), *S. carpocapsae* Breton strain (13,925 bp), *S. carpocapsae* strain All (13,924 bp), *S. glaseri* strain Sds102 (15,182), *S. kushidai* strain Hamakita (13,851 bp) and *S. litorale* strain IbKt142 (21,403 bp) (Kikuchi et al., 2016; Montiel et al., 2006; Regeai et al., 2021).

The whole genome of an inbred Indian strain of *Heterorhabditis*, *H. indica* Hms1-i20 was solved in the present study. The genome was assembled by combining the high-quality Illumina sequence data from genomic libraries of 300 bp, 600 bp and a 5 kb mate-pair library. The draft genome of *H. indica* Hms1-i20 is 91.26 Mb in size which is ~ 14 Mb larger in size as compared to *H. bacteriophora* TTO1-M13e genome. *H. indica* genome was assembled in 3,538 scaffolds, whereas *H. bacteriophora* genome was assembled in 1,263 scaffolds. Nevertheless, N50 metric was higher in case of *H. indica* genome (587.27 kb) as compared to N50 of 312.33 kb in case of *H. bacteriophora* TTO1. Genome completeness assessment by BUSCO and CEGMA showed that the genome is more than 90% complete, providing confidence in further gene prediction and annotation. The numbers of protein-coding gene models predicted for *H. indica* assembly are 10,494 which is lesser than proteins predicted in draft genome of *H. bacteriophora*. As compared to *H. bacteriophora* annotated gene models, a higher percentage of genes were annotated in *H. indica*, where 91.44 % of the predicted genes could be annotated. The percentage of orphan genes was relatively low, with only 898 (8.5%) of the predicted genes finding no exact matches in the databases utilized for annotation. Additionally, the mitochondrial genome of *H. indica* was assembled which is 17,393 bp in size. This is the first mitogenome information for *H. indica* and will be a useful resource for evolutionary and population studies in *Heterorhabditis* nematodes.

5.1.1 Comparative analysis between *Heterorhabditis indica, Heterorhabditis bacteriophora, Steinernema carpocapsae, Caenorhabditis elegans* and *Oscheius tipulae*

The genomes of four additional nematodes, namely *H. bacteriophora* (PRJNA13977), *S. carpocapsae* (PRJNA202318), *C. elegans* (PRJNA13758), and *O. tipulae* (PRJNA644888) were chosen for comparison with the genome of *H. indica*. *H. bacteriophora* and *S. carpocapsae* were chosen because they are both entomopathogenic nematodes, and *C. elegans* is a well-established model and a close relative of *Heterorhabditis* worms. The genome of an entomopathogenic *Oscheius* sp. TEL-2014 is highly fragmented, with just 44% assembly completeness (Lephoto et al., 2016). Therefore, as a representative of the *Oscheius* clade, a chromosomal-scale genome of *O. tipulae* (PRJNA644888) was included in the investigation (Gonzalez de la Rosa et al., 2020). Comparative analysis of *H. indica* genome in comparison to four other nematode genomes revealed that *H. indica* shared the highest number of orthologous groups with *H. bacteriophora*, followed by *C. elegans* and *S. carpocapsae*. It shared the lowest number of orthologous groups with *O. tipulae*. Comparative protein domain analysis revealed that various protein domains were present in higher abundance in *C. elegans, S. carpocapase*, and *O. tipulae* in comparison to *H. indica*, whereas the domain counts in *H. bacteriophora* was lesser. This disparity could be primarily attributed to specific biology and/or to different genome size. Additionally, sequencing platforms, assembly, and annotation methodologies, completeness of assemblies impact the number of protein domains predicted. The chromosome level genome sequence of *C. elegans* is nearly 100% complete with no gaps as compared to any other nematodes and serve as reference assembly for gene prediction and annotation in other nematodes. As a result, in the comparison analysis, the number of protein domains found in *C. elegans* was higher (The *C. elegans* sequencing consortium, 1998; Yoshimura et al., 2019).

5.1.2 Candidate gene and gene families characterized in *Heterorhabditis indica* genome

Genome data may provide information about candidate genes that underpin the parasitic and mutualistic biology of EPNs (Lu et al., 2016). The genome of *H. indica* was further characterized, revealing a slew of genes that could be involved in key

biological processes. Proteins containing transmembrane domains play critical roles in organisms, such as receptors, transporters, and in signaling, nutrition absorption, respiration, energy production, and immune response, etc. Approximately 24% of the predicted *H. indica* proteins contained one or more transmembrane domains. G protein-coupled receptors (GPCRs) belong to a family of seven transmembrane proteins. Several GPCRs are important sensory receptors. It is very well established that nematode GPCRs are an integral part of chemoreception. The abundance and diversity of GPCRs could be correlated to the niche inhabited by a nematode (Robertson & Thomas, 2006; Srinivasan et al., 2013). In the case of entomopathogenic nematodes, GPCRs are implicated in the chemoreception of insect odors and host-seeking in addition to sensing other environmental cues. There are more than 600 predicted GPCRs in different *Steinernema* genomes and 82 in the *H. bacteriophora* genome (Bai et al., 2013; Lu et al., 2016). A comprehensive analysis of heptathetical transmembrane proteins identified 56 GPCRs in the case of *H. indica*. These putative olfactory receptors could be investigated further to understand the host-seeking abilities and host specificity of these EPNs.

Secreted proteins play a crucial role in organism growth, reproduction, interactions with mutualistic bacterial, and pathogen defense, among other things (Bai et al., 2013). Several nematode-secreted proteins have been identified as effector molecules that aid parasitic abilities (Chang et al., 2019; Rehman et al., 2016). In the case of EPNs, the repertoire of secreted proteins may have a role in parasitic interactions with insects, influencing host range and/or specificity. EPN-proteases and protease inhibitors are a class of secreted proteins that have been linked to invasion, immunological regulation or evasion in the host, killing, and degradation of host insect tissues (Lu et al., 2016). Characterization of *H. indica* secretome based on the genome information identified 370 putative secreted proteins plus 36 proteins localizing to mitochondria, endoplasmic reticulum, nucleus and glycophosphatidylinositol (GPI) anchored proteins. A total of 38 peptidases and 31 peptidase inhibitors were identified. As compared to *H. bacteriophora* it is relatively a higher number, where only 19 peptidases and 9 peptidase inhibitors were reported (Bai et al., 2013). In contrast to heterorhabditids, steinernematids have a larger number of predicted proteases and protease inhibitors with signal peptides. Genome analysis demonstrated the expansions of *Steinernema*-specific serine and metalloproteases as well as several families of

protease inhibitors in steinernematids (Dillman et al., 2015). Another noteworthy nematode-specific gene family implicated in nematode parasitism of plants, animals, and insects is the Fatty Acid and Retinol Binding (FAR) proteins (Yuan et al., 2021). FAR proteins have a role in immune evasion and/or immune suppression by combating the lipid-based defense mechanisms of hosts (Parks et al., 2021). Three FAR proteins were reported in *H. bacteriophora* (Lu et al., 2016). We identified two FAR proteins in the secretome of *H. indica*. However, this number is significantly less when compared to the number of FAR proteins detected in *Steinernema* nematodes. The FAR protein family appears to be expanded in steinernematid genomes, with 40-50 FAR genes found in different species of *Steinernema* (Chang et al., 2019; Lu et al., 2016). It was postulated that *Heterorhabditis* nematodes depend on *Photorhabdus* symbionts for secreted effectors (Bai et al., 2013; Waterfield et al., 2009). Certain studies have shown that EPNs like *H. bacteriophora* rely on their symbiotic bacteria to overcome the host immunity and kill the insect host, whereas EPNs like *S. carpocapsae* are lethal even without their bacteria (Han and Ehlers, 2000). This in part explains the limited repertoire of secreted proteins observed in the genome of *Heterorhabditis* when compared to Steinernematids. Further mechanistic investigations are required to understand the role of specific secreted products in facilitating parasitic and/or other interactions. Investigating the secretome of EPNs can help to gain a better understanding of parasite evolution and improve biocontrol features.

The pathogen-host interaction analysis on *H. indica* genome identified more than 140 genes putatively involved in interaction with insects. Interestingly, several genes involved in interaction with vertebrates and plants were also identified in *H. indica* genomes. These nematodes are found in the soil/rhizosphere of plants where they hunt for herbivore insects, as well as interact with other soil-associated fauna and flora. Plants that have been harmed by insects have been demonstrated to attract *H. megidis* nematodes (Rasmann et al., 2005). However, little is known about the molecular level interaction of EPNs with their surrounding biotic system. Genome data, combined with particular analysis, form the basis for further molecular characterization of genes and processes involved in nematode interactions with other organisms in their environment. EPNs are also utilized as a model system to study nematode infections in higher vertebrate animals (Dillman et al., 2015; Vadnal et al., 2017), and a thorough examination of the list of probable *Heterorhabditis* genes implicated in vertebrate

interaction could aid in the understanding of evolution and mechanisms of animal parasitism.

5.1.3 Putative horizontal gene transfer events, transposon and other regulatory elements identified in *Heterorhabditis indica* genome

Recent studies using genome data have demonstrated that the evolution of eukaryote organisms has been remarkably influenced by the acquisition of a number of genes by lateral or horizontal gene transfer (LGT or HGT) across different domains of life. HGT is a significant driving factor in genome evolution. In the case of plant-parasitic nematodes, it is very well established that these nematodes have acquired some genes from soil microorganisms through HGT and it helped them to evolve plant-parasitic lifestyle (Danchin and Rosso, 2012; Haegeman et al., 2011). Such analysis has not been done in the case of entomopathogenic nematodes. Here we present the first-ever HGT analysis in the case of EPNs. *Heterorhabditis* seems to have acquired some gens from Proteobacteria and Fungi like *Rhizopus* and *Beauveria*. Even the HGT score supported *Beauveria* as a top donor, interestingly which is an entomopathogenic fungus. Further phylogenetic and expression investigation may validate these putative HGT events, traits conferred by the acquired genes.

Transposons are repetitive DNA sequences that have the capability to move from one location to another in the genome. Thus, they are considered an important contributor for gene and genome evolution (Kozlowski et al., 2021). In this study, we reported the abundance, diversity, the distribution of different families of transposable elements in *H. indica* genome. Approximately 15 % of *H. indica* genome scaffolds were harboring one or another class of transposable elements. A total of 1,548 transposable elements belonging to 11 different classes were detected in *H. indica* genome. In *H. bacteriophora* genome, an abundance of mariner transposons with 138 predicted in Mariner transposase was reported (Bai et al., 2013b). In our analysis, we could identify 131 loci belonging to Tc1-Mariner/TIR/DNA Transposon class. However, transposon belonging to the classes such as Zator/TIR/DNA Transposon, hAT/TIR/DNA Transposon and Gypsy/LTR/Retrotransposon were mostly enriched in *H. indica* genome. Identified TE can be utilized in future studies to understand the role of mobile elements in the nematode genome evolution and nematode genetic variability studies. Transposons can be exploited as an important genetic tool for genetic manipulation or transformation studies (Kozlowski et al., 2021; Muñoz-López & García-Pérez, 2010).

Noncoding RNAs (ncRNAs) act as cellular regulators without encoding proteins. ncRNAs regulate gene expression at the transcriptional and post-transcriptional levels. Certain ncRNAs appear to control epigenetic processes as well (Li & Liu, 2019). A total of 631 ncRNA were identified on 249 scaffolds of *H. indica* genome. Out of them, 269 were tRNA genes. Moreover, 7 t-RNA-Sec loci encoding selenocysteine (Sec) (21st proteinogenic amino acid) were identified. In the case of *H. bacteriophora,* 254 transfer RNA (tRNA) genes and 1 tRNA pseudogene and 0 tRNA-Selenocysteine genes were predicted (Bai et al., 2013s). t-RNA-Sec genes were also reported in *S. carpocasae, C. elegans, Pristionchus pacificus,* and *Ascaris suum* (Rougon-Cardoso et al., 2016). The identification of t-RNA-Sec genes in certain nematodes, but not all, may be explained by a lineage-specific study of selenocysteine machinery (Otero et al., 2014). In *Steinernema* species, the correlation between non-coding RNA regulatory networks and variation in host-finding behaviors was demonstrated (Warnock et al., 2021). Along with genomic data, further expression studies can help to decipher the regulatory role of detected ncRNAs and epigenetic mechanisms in *Heterorhabditis* nematodes.

5.1.4 Identification of molecular markers for population genetics of *Heterorhabditis* nematodes

Lastly, nematode population genetics and the identification and application of relevant markers for this purpose are very important in understanding the origins, expansion, and intraspecific diversity of nematodes. Microsatellites or simple sequence repeats (SSRs) are tandem repeat sequences of 1–6 bp that serve as informative genetic markers for population genetics (Stock and Reid, 2004). Earlier 3,794 microsatellite loci were predicted in 506 contigs of *H. bacteriophora* draft genome (Bai et al., 2013). Some polymorphic microsatellite markers were developed and used to distinguish populations of *H. bacteriophora* (Bai et al., 2009). A total of 2,954 microsatellite loci were predicted in 459 scaffolds of the current *H. indica* draft genome. The microsatellites identified in this study will be useful for the genetic diversity and phylogeographic studies of *Heterorhabditis* isolated from various geographical areas of the world. This genomic resource will facilitate functional and comparative genomic studies and genetic exploration in *Heterorhabditis* nematodes. Many potential genes and gene families were discovered during the study that could be exploited to improve biocontrol traits in these nematodes. Genome data can be used to decipher genome-

scale phylogeny, evolution, epigenetics, other regulatory networks and mechanistic investigation of EPN biology.

5.2 Comparative transcriptomics of symbiotic and axenic early adults of *Heterorhabditis* nematodes to decipher the nematode factors involved in symbiosis with *Photorhabdus* bacteria.

Complex multicellular animals and plants exhibit strong interdependencies with their associated microbes (Gilbert et al., 2012). Our understanding of animal-microbe symbiosis, and the tremendous impact the bacterial symbionts have had on animal evolution, and various aspects of animal biology have grown tremendously in recent years (McFall-Ngai, 2008). However, not much is known about the animal host factors that help maintain the intricate balance with their bacterial symbionts. As compared to the animals associated with multiple genera/species of bacteria, the nematode-bacterial symbiotic relationship of *Heterorhabditis* nematodes and *Photorhabdus* bacteria offers a powerful model system to study the intricacies of animal-microbe symbiosis in a mono-specific symbiotic association (Ciche, 2007). Here we used the *Heterorhabditis*-*Photorhabdus* system to identify the nematode factors involved in symbiosis with the bacteria.

We identified 754 differentially expressed transcripts in the symbiotic nematodes as compared to axenic worms. Analysis of DEGs indicated that the symbiotic bacteria alter the gene expression in the host animals. The ribosomal pathway was the most affected, and the top down-regulated genes were related to ribosomal proteins, confirming a down-regulation of translation in the nematode host. Expression of *COII* encoding 'Cytochrome c oxidase subunit 2' and *MT-CYB* encoding 'Cytochrome b' were also significantly down-regulated in symbiotic nematodes. Bacteria are known to affect ribosomal and mitochondrial pathways (Cohen and Troemel, 2015). As seen in *C. elegans*, nematodes can sense perturbations in the core processes such as blockade of mitochondrial, transcription and translation-related pathways, ATP synthesis, proteasomes in intestine or hypodermis, and induce defence responses such as avoidance behaviour and defence genes expression (Dunbar et al., 2012; McEwan et al., 2012; Melo and Ruvkun, 2012). The observed changes in expression levels of core housekeeping genes in our study may be indicative of a trade-

off in which elevated immunity is prioritized over other secondary functions, as seen in other models (for example *C. elegans*-microbe interactions).

5.2.1 A repertoire of immune recognizers and effectors identified in the symbiotic *Heterorhabditis* nematode specific transcripts

Additionally, 12,151 transcripts unique to symbiotic nematodes were identified and we ascertained that these transcripts (with FPKM value >1) are not contaminants. As opposed to the prevalent approach of analyzing only the differentially expressed genes between two conditions (here symbiotic vs. axenic nematodes), we also analysed these 12,151 transcripts and found several transcripts relevant to the phenotype. Symbiotic nematodes are cohabiting with bacterial symbionts, whereas axenic nematodes are not - which is a drastic difference for the nematode. It is possible that some transcripts are exclusively transcribed under symbiotic conditions. This is in conformation with several research publications on tissue/cell/condition-specific transcripts (Kim et al., 2014; Rödelsperger et al., 2020; Sun et al., 2021; Zhu et al., 2016). In addition, unique transcripts in symbiotic nematodes might be a result of alternative splicing (AS) under specific physiological conditions (here presence of bacterial symbionts). Nematodes are known to have AS in the range of 20-30%, typically with 2-3 transcripts per AS locus (Abubucker et al., 2014). However, a deeper analysis of the AS in our experimental conditions is beyond the scope of this study.

Comprehensive annotation of the differentially expressed and unique transcripts indicated that several of them are involved in nematode-bacterial interactions, notably in nematode immunity. Host immunity has a vital role in the host-microbe interactions as well as animal growth and development (Boehm et al., 2012; Gerardo et al., 2020; Ryu et al., 2010; Tan and Shapira, 2011). Additionally, the immune system also coordinates biochemical and cellular responses to changes in the molecular landscape in hosts and helps maintain a healthy balance between the host and symbionts (Brown et al., 2013; Eberl, 2010; Hooper et al., 2012). We anticipated four possibilities during symbiont attachment and biofilm formation in the host nematode intestine. First, the nematode´s immune system is insensitive to the presence of the symbiont due to the active suppression exerted by the bacteria resulting from years of co-evolution. Second, the immune system of the nematode is active, but only at the local level near or in infected cells, or in response to biofilm formation. Third, the bacteria use a cloaking mechanism to remain invisible to the host immune response. The fourth possibility

represents a combination of all the above. The transcriptional response of symbiotic nematodes observed in our study suggests that the nematode immune system responds to the presence of symbiotic bacteria indicating a combination of all the above-mentioned scenarios.

Interactions between hosts and microbes are largely mediated through microbe-derived ligands, i.e., microbe-associated molecular patterns (MAMPs) and pattern recognition receptors (PRRs) of hosts (Koropatnick et al., 2004). Bacterial factors such as extracellular polysaccharides, lipopolysaccharides, and fimbrial structures are key to symbiosis by *Photorhabdus* bacteria (Easom et al., 2010; Somvanshi et al., 2010). Our results show that the symbiotic nematodes express orthologues of proteins which are known to have pattern recognition receptor activity. For example, orthologs of mammalian proteins such as lipopolysaccharide-binding protein (LBP), bactericidal permeability-increasing protein (BPI), deleted in malignant brain tumors 1 protein (DMBT) / Glycoprotein 340, membrane associated glycoprotein CD36, MHC class I-related protein 1 (MR1) and orthologs of *C. elegans* proteins such as probable galaptin LEC-8, DAF-7 and IRG-7 (an infection response protein) were detected in the transcriptome of symbiotic nematodes. Additionally, several mammalian lectin orthologs with carbohydrate-binding activity (for example, CLEC4A (C-type lectin domain family 4 member A), LGALS9 (Galectin-9), and LY75/CLEC13B (Lymphocyte antigen 75) were found expressing in symbiotic nematodes. All these proteins exhibit a range of properties such as lipid binding, carbohydrate-binding, and glycolipid binding, and they play roles in immune surveillance and/or as immunological effectors (Beamer et al., 1998; End et al., 2009; Engelmann and Pujol, 2010; Means et al., 2009; Nicholas and Hodgkin, 2004; Pees et al., 2021; Yunger et al., 2017).

Several transcripts expressed in symbiotic nematodes mapped to canonical nematode immune pathways such as MAPK cascades (PMK-1 p38 MAPK, ERK, and JNK pathways), DAF-2/ILR pathway, and TGF-β/DBL-1. Additionally, enrichment of cAMP, mTOR, Rap1, Calcium and Ras signalling pathways was also observed in symbiotic nematodes. A number of recognized kinases, phosphatases, and other proteins (*abl-1, vhp-1, MAPKBP1, PPM1D*) are known to facilitate pathway interactions and/or pathway regulation to achieve immunological balance (Burton et al., 2006; Kim et al., 2004; Kim and Ewbank, 2018; Marsh et al., 2011; Mizuno et al., 2004). Furthermore, transcription factors act downstream of immune pathways to

regulate gene expression (Dierking et al., 2016; Engelmann and Pujol, 2010). Transcription factors (TF) such as forkhead TF (*daf-16*), GATA TFs (*elt-2*, *pnr*), *fos-1*, STAT TFs (*sta-1, sta-2*) and TFEB which regulate defense response to bacteria by regulating the expression of antibacterial/antimicrobial effectors were also identified in symbiotic nematodes. Transcription factors ELT-2 and TFEB are known to act in a tissue-specific manner and regulate defence responses during intestinal infection of nematodes with bacteria (Shapira et al., 2006; Visvikis et al., 2014).

A variety of immune effectors and mechanisms such as autophagy, cell death, endocytosis, production of ROS, lysozymes, proteolytic and lipolytic enzymes and antimicrobial proteins have been implicated in controlling microbial load and activity in host-microbial interactions and play role in defence against infections (Broderick, 2016; Dierking et al., 2016; Engelmann and Pujol, 2010; Nyholm and Graf, 2012). In our study, several genes associated with cell death/apoptosis, autophagy and lysozymes (for example: *ced-1*, *ced-3*, *ced-4*, *atg16.2*, *lgg-2*, *ilys-2* and *ilys-3*) were detected exclusively in symbiont-associated nematodes, suggesting these genes could be important in regulating levels of symbiont population and reduce the deleterious effects of infections on host fitness as seen in other studies (Aballay and Ausubel, 2001; Curt et al., 2014; Dunbar et al., 2012; Haskins et al., 2008; Jia et al., 2009; Zou et al., 2014). Moreover, endocytosis (ko04144) and focal adhesion (ko04510) were the most enriched pathways in symbiotic nematodes. Symbionts that establish a biofilm on the maternal intestine gradually invade the rectal gland cells and become intracellular. Internalization of symbionts has been linked to the endocytosis phenomena (Ciche, 2007; Duperron, 2017). Focal adhesions are integrin containing multi-protein structures. They are associated with the plasma membrane and serve as mechanical linkages between internal actin bundles and the external substrate. Endocytosis along with focal adhesion associated proteins could be exploited by symbionts for attachment and sequestration into host rectal gland cells. This kind of molecular interaction has been studied in *Staphylococcus aureus*, which uses the integrin pathway and focal adhesion kinases to attach to host cells and for internalization (Agerer et al., 2005; Josse et al., 2017). Several peroxidase encoding genes such as *bli-3*, *hpx-2*, *skpo-1*, *gst-5*, *gstk-1* and *glo-3* were detected in symbiotic nematodes. Peroxidases catalyse the generation of ROS (reactive oxygen species), which are well-known microbicidal effectors and signalling molecules (Chávez et al., 2007, 2009; van der Hoeven et al.,

2012). Numerous studies in invertebrates have implicated the role of ROS in the regulation of gut microbiota (Douglas, 2014; Nyholm and Graf, 2012). *Photorhabdus* symbionts are known to utilize pentose phosphate pathway to avoid oxidative stress and hence persist in their animal hosts (An and Grewal, 2010), and the expression of many peroxidases and reductases in symbiotic nematodes correlate with this observation. It may be summarised that the transcriptional response of symbiotic *Heterorhabditis* nematodes resembles the response against any pathogenic bacteria since various transcripts involved in immune responses at several levels (from pathogen sensing to production of antimicrobials) are active in symbiotic nematodes as compared to axenic nematodes. However, we observed a limited repertoire of markers of immune recognition, immune-activation as well as effectors compared to nematode pathogenic bacterial interactions. These responses observed in symbiotic nematodes might represent non-systemic and evolved localized responses to maintain mutualistic bacteria at non-threatening levels.

5.2.2 *Heterorhabditis*-*Photorhabdus* association is a 'mixed' symbiotic system

In the context of symbiont evolution, animal-microbe symbiotic systems have been classified as open, closed and mixed (Perreau and Moran, 2022). In the open systems (e.g., bobtail squids and *Aliivibrio fischeri*), microbes colonise their hosts repeatedly from external environmental niches and both the partners manipulate each other at molecular level to achieve symbiosis. The closed systems (e.g., aphids and *Buchnera aphidocola*) are characterised by strictly maternal transmission of symbionts enforcing strict clonality and a reduction of host immune responses against symbionts (Gerardo et al., 2020). In mixed systems (e.g., common fruit fly and *Wolbachia pipientis*), vertical transmission is the rule, with the occasional horizontal transmission of symbionts (Perreau and Moran, 2022). Based on the nematode-bacteria symbiotic life cycle (Ciche and Sternberg, 2007; Waterfield et al., 2009), species associations and evolutionary history (Maneesakorn et al., 2011; Sajnaga and Kazimierczak, 2020), and a large number of molecular perturbations observed in our results - it may be suggested that the *Heterorhabditis* nematode-*Photorhabdus* bacteria symbiosis is a mixed symbiotic system, and exhibits characteristics of an evolving symbiosis. Based on our transcriptome data, we propose a model (Figure 5.1) indicating how the nematode-bacterial symbiosis is achieved and maintained in early adults. Subject to further functional validations, our findings suggest that *Heterorhabditis* nematode immune

system plays a pivotal role in the maintenance of symbiosis with its *Photorhabdus* bacterial partner. It's a two-way process - the bacteria influence core physiological functions of the host, whereas the nematode actively senses the symbionts and keeps them in check through active regulation of its immune system. To our knowledge, this is the first transcriptomic investigation of animal factors involved in the maintenance and regulation of microbial symbiosis.

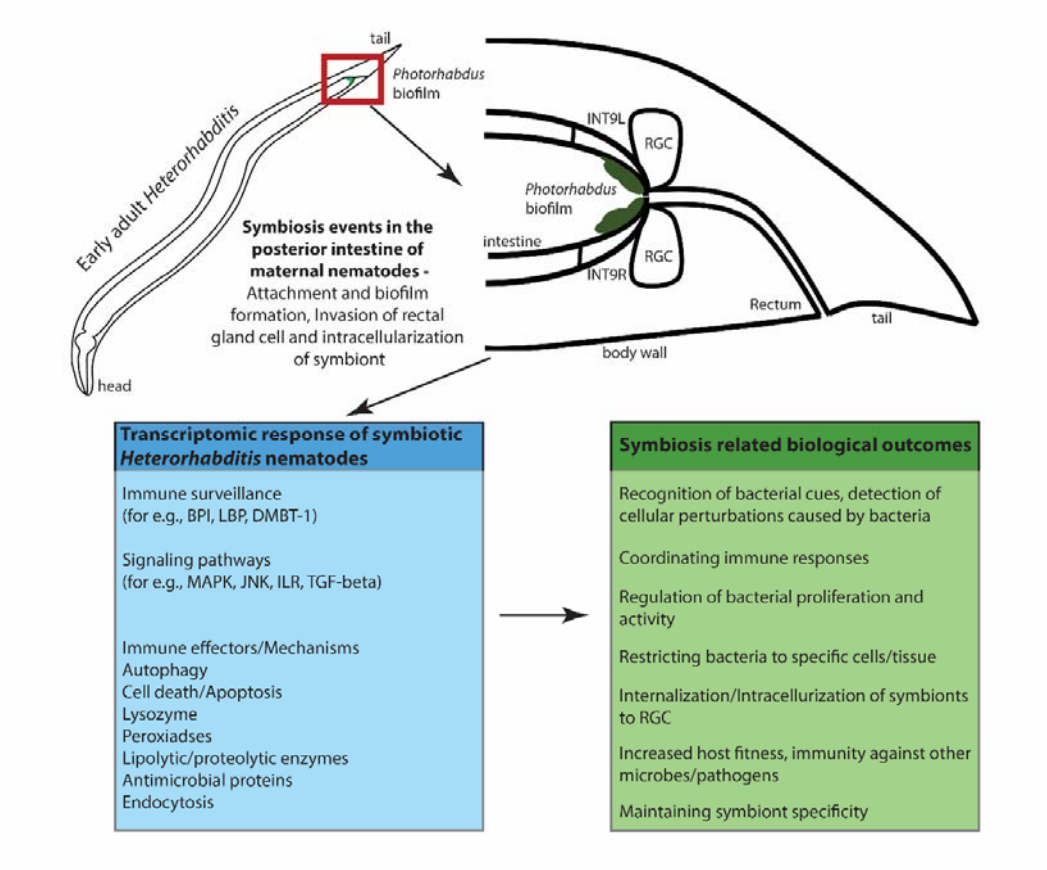


Figure 5.1. A model presenting symbiotic events and the transcriptomic responses in early adult stage of *Heterorhabditis* maternal nematodes, and their interpretation in relevance to symbiosis with *Photorhabdus* bacteria. Colonization processes involve two-way crosstalk between host and symbiont. From the nematode host side, based on gene expression profiling, expression of immune system components which are involved in the recognition of bacteria, bacteria-derived molecules and detection of any cellular perturbations caused by bacteria are observed. It leads to activation of signalling pathways and results in an array of host immune and defence responses such as autophagy, apoptosis and production of antimicrobial proteins. These effectors and mechanisms regulate symbiont bacterial numbers and help achieve successful symbiosis. INT9L/INT9R – posterior nematode intestinal cells; RGC- rectal gland cells.

5.3 Standardization of RNAi gene silencing in *Heterorhabditis* nematodes

Genomic and transcriptomic data of *Heterorhabditis* sp. are being generated in order to understand the biology of the nematode. Molecular tools are required to translate genomic resources into further functional studies. The lack of optimized methods to carry out functional genomics restricts the further validation and utilization of this data. RNA interference (RNAi) is a technique for studying the gene function in a direct, targeted manner by knocking down the specific mRNAs and analyzing the phenotypic effects (Dalzell et al., 2011). RNAi-based gene silencing was discovered in *C. elegans* (Fire et al., 1998). Soaking nematodes in a dsRNA-containing solution is a convenient and widely used method for achieving RNAi in nematodes. Soaking of post-embryonic developmental stage in dsRNA solution and microinjection-based delivery of dsRNA into the gonad of adult hermaphrodites for RNAi induced gene silencing has been reported in *H. bacteriophora* (Ciche and Sternberg, 2007; Ratnappan et al., 2016).

When *Heterorhabditis* eggs were used for RNAi by soaking, we observed poor, unsynchronized hatch and higher mortality of hatched juveniles. It impeded in drawing the correlation between the treatment and phenotypic effects. L1s coming out of eggs were very delicate and not suitable for prolonged soaking and this affected the reproducibility of this method. Several studies have reported the difficulties in the induction of RNAi or inconsistencies in RNAi datasets for several nematodes (Knox et al., 2007; Maule et al., 2011; Viney and Thompson, 2008). Some studies showed significant differences in the RNAi mechanism even among closely related species (Nuez and Félix, 2012). Inefficient or inconsistent RNAi was observed in many animal-parasitic nematodes like *Haemonchus* and *Heligmosomoides* and free-living nematodes like *Rhabditis* sp, *Mesorhabditis* sp., *Oscheius* sp and *Acrobeloides* sp (Geldhof et al., 2006; Lendner et al., 2008; Wheeler et al., 2012). Effects of RNAi are known to vary among genes, species, methods used to introduce dsRNA, developmental stages, and even experiments (Félix, 2008; Geldhof et al., 2007; Knox et al., 2007).

RNAi effector complements in *Heterorhabditis* nematodes were described in comparison to *C. elegans* as a first step. In *C. elegans,* 77 genes are characterized to be involved in core RNAi machinery which is categorized into six functional groups namely small RNA biosynthesis, dsRNA uptake and spreading, siRNA secondary amplification, AGOs, and RISC, RNAi inhibitors, and nuclear effectors (Dalzell et al.,

2011). A total of 59 and 60 gene orthologs were detected in *H. bacteriophora* and *H. indica* respectively. Although *C. elegans* encodes significantly more RNAi pathway effectors than *Heterorhabditis* nematodes, *Heterorhabditis* worms seem to have the majority of the components of the RNAi machinery. Secondly, dsRNA soaking efficiency in different developmental stages of *Heterorhabditis* was tested using the fluorescent dye fluorescein isothiocyanate (FITC) as a marker of uptake. This study showed that L1s hatching from eggs were highly sensitive and not suitable for prolonged soaking. The *Heterorhabditis* infective juveniles (IJ) showed no FITC absorption. Addition of neuro-stimulant octopamine, and desiccation prior to soaking induced uptake of FITC only up to the oesophagus-intestinal valve region. The IJs are non-feeding, developmentally arrested dauers with an additional cuticula sheath. These factors might be restricting the uptake of exogenous molecules in the case of *Heterorhabitis* IJs. However, infective juveniles of *S. abbasi* showed strong uptake of FITC even without neuro-stimulant. Transcuticular absorption of FITC was observed in *Steinernema* IJs. This in part explains why RNAi was successful in IJ stage of *Steinernema* (Morris et al., 2017), but challenging in the IJ stage of *Heterorhabditis*.

Post-IJ recovery (PIJR) stages of *Heterorhabditis* showed efficient uptake of FITC which was further enhanced by octopamine. *Heterorhabditis* nematodes shed their extra cuticle and start feeding during the PIJR stage, which could be contributing to the significant uptake of FITC from the soaking solution. Correlation between feeding/non-feeding stages of *Heterorhabdtis* and FITC uptake suggests that it is being actively ingested in the intestinal tract of the PIJR stage, and not just passively absorbed through the cuticle. This study suggests that the PIJR stage could be the most amenable stage for the dsRNA delivery through the soaking method for *Heterorabditis* nematodes. Efficacy of RNAi by soaking PIJR stage in dsRNA solution of three test genes was tested. *Hb-dpy-13* and *Hb-dpy-7* are dumpy genes which encode nematode cuticular proteins and *Hb-cct-2* gene regulates gonad development in worms (Ratnappan et al., 2016). Worms treated with dsRNA of dumpy genes and *Hb-cct-2* gene showed reduced body length and affected gonad development respectively. A significant and consistent decrease in the target gene expression was observed in dsRNA-treated worms. These experiments were repeated 3 times and were found to work consistently. These findings provide a proof of concept for successful RNAi using the PIJR stage for soaking in dsRNA solution.

Based on the observations in the present study it is suggested that developmental stage, feeding strategy, cuticle influence the amenability of RNAi in the case of nematodes. According to Lendner et al. (2008), factors such as sequence specificity, gene expression level, stage-specific expression, stability of the target protein as well as strength and structure of the cuticle are likely to influence RNAi in nematodes (Lendner et al., 2008). Dalzell et. al. (2011) reported that differences in RNAi effector protein complements cannot explain the variances in RNAi competencies in different nematode species. They suggested that differences in the expression levels or localization of RNAi effector proteins across life stages and/or species, as well as other factors such as culture conditions, uncharacterized morphological differences (For example cuticle permeability to nucleic acids), and other genetic diversity in the nematode populations, could influence the success of RNAi (Dalzell et al., 2011; Maule et al., 2011). The site of gene expression influences susceptibility to RNAi by soaking, as demonstrated in the case of *Haemonchus*, which had previously been reported to be insensitive to RNAi (Samarasinghe et al., 2011). They observed a consistent silencing of genes expressed in the intestine, excretory cell, or amphids by RNAi and suggested that genes expressed in sites accessible to the environment are more likely to be susceptible to RNAi by soaking (Samarasinghe et al., 2011). RNAi soaking protocols can be further optimized for other developmental stages of *Heterorhabditis* nematodes such as the IJ stage by taking into consideration of target gene, expression site, dsRNA concentration, and exposure time. For example, increased dsRNA concentration (5mg/ml dsRNA+50mM serotonin, as opposed to 1mg/ml dsRNA used in most cases) and an extended period of incubation up to 48 h in dsRNA produced a strong knockdown of target genes in IJs in another EPN *S. carpocapsae* (Morris et al., 2017). Therefore, based on our study and previous reports, we believe that optimization of RNAi protocols depending on nematode species, developmental stage, and target genes is crucial to use it as a valid and feasible functional genomics tool.

In conclusion, the deliverables of this thesis work include a high-quality genome sequence for the entomopathogenic nematode *H. indica* and transcriptomic data of symbiosis relevant early adult stage of *Heterorhabditis.* A model based on RNA-seq data was proposed to explain potent nematode factors involved in symbiosis with *Photorhabdus* bacteria. Lastly, a standardized protocol for RNAi-based functional validation of genes in *Heterorhabditis* was reported. Genome and transcriptome

resources may help to understand and to design tools to improve biocontrol traits of EPNs. These resources can be used for the investigation of regulatory elements, epigenetics, research on evolution and functional specialization of entomopathogenic nematodes. The omics data along with standard molecular tools may promote further utilization of *Heterorhabditis* nematodes as model system to understand ubiquitous animal-microbe symbiosis and host-parasite interactions.

Symbiotic associations between animals and microbes are ubiquitous in nature. However, the understanding of animal factors involved in symbiosis with bacterial partners is limited. The entomopathogenic nematode (EPN) *Heterorhabditis* lives in monospecific symbiotic association with a gamma proteobacteria *Photorhabdus* and offers an excellent genetically tractable model to study animal-microbe symbiosis. *Heterorhabditis* nematodes and their bacterial partners are highly virulent insect pathogens and employed as biocontrol agents against insect pests. Symbiotic bacteria have a major impact on nutrition, growth, reproduction, and virulence traits of EPNs. Symbiotic association between this nematode and bacteria is highly specific and *Photorhabdus* symbionts are maternally transmitted to the next generation offspring of *Heterorhabditis*. Several molecular determinants of bacteria involved in this symbiotic association have been identified. However, information on nematode factors that govern the symbiosis are lacking. Additionally, genomic information of *Heterorhabditis* nematodes available in public domain is very scant, where only one genome assembly of *H. bacteriophora* is available.

Here, we present the first draft genome assembly for *H. indica*, which is widely present in the warmer and tropical climatic regions and is one of the most commercialized EPNs. An Indian strain of *H. indica* was inbred for 20 generations (designated *H. indica* Hms1-i20) to obtain genetically homogenous nematodes for genome sequencing. The draft genome of *H. indica* was sequenced using three genomic libraries of 300 bp, 600 bp and 5 kb sizes by Illumina HiSeq platform. The size of the draft genome assembly was 91.26 Mb, comprising 3,538 scaffolds. Genome assembly is 90.5% complete as assessed using BUSCO (Benchmarking Universal Single-Copy Orthologs) and 10,494 protein-coding genes were predicted in the present genome. Identification of orthologous genes present in *H. indica* Hms1-i20 genome compared to four other nematode genomes revealed that *H. indica* shared 6,574 groups with *H. bacteriophora*, 6,635 with *C. elegans*, 6,228 with *S. carpocapsae* and 6,669 with *O. tipulae*. A total of 4,824 orthogroups were common between these five nematodes analysed. Protein domain and secretome characterization identified 2,525 transmembrane domain proteins and 370 putative secreted proteins. Subsequent analyses identified 56 GPCRs, 30 peptidases, 31 peptidase inhibitors and 2 Fatty-acid

retinol binding proteins in *H. indica* proteome and such proteins may facilitate nematode interactions with insect hosts or bacterial symbionts. Pathogen-host interaction analysis identified a total of 143 putative *H. indica* proteins involved in interaction with insect hosts. Many potential genes and gene families were discovered during the study that could be exploited to improve biocontrol traits in these nematodes.

Additionally, 2,549 microsatellite loci, 1,548 transposable elements, 631 non-coding RNA loci, 21 likely cases of horizontal gene transfer (HGT) were identified. Transposons and events like horizontal gene transfer contribute to the evolution of the genome. ncRNA are important cellular regulators which govern epigenetic mechanisms. The microsatellites identified in this study will be useful for the genetic diversity and phylogeographic studies of *Heterorhabditis* isolated from various geographical areas of the world. The mitochondrial genome of *H. indica* was assembled separately which is 17,393 bp in size and 46 genes were detected in mitochondrial genome assembly. *H. indica* genome data can be used to decipher genome-scale phylogeny, evolution, epigenetics, other regulatory networks and mechanistic investigation of EPN biology. This genomic resource will facilitate functional and comparative genomic studies and genetic exploration in *Heterorhabditis* nematodes.

RNA-sequencing was used to investigate nematode factors involved in symbiosis at the early-adult stage of *H. bacteriophora*. *Photorhabdus* symbiont attach and form biofilm in the posterior intestinal cells of early adults (36-40 h post IJ recovery) and this step is very crucial in the specific transmission of symbionts to the next generation. Comparative transcriptomics analysis of symbiotic and axenic early adults identified 46,599 transcripts. A total of 754 differentially expressed transcripts were identified in symbiotic nematodes, where 207 transcripts were up-regulated and 547 transcripts were down-regulated. Additionally, 12,151 transcripts exclusively expressed in symbiotic nematodes. Pathway analysis showed that ribosome, calcium signaling pathway, neuroactive ligand-receptor interaction pathways are differentially expressed. Endocytosis, cAMP signalling and focal adhesion were the top enriched pathways in symbiotic nematodes. A large number of transcripts involved in nematode immune/defence responses against bacteria such as lysozymes, peroxidases, autophagy, apoptosis and antimicrobial proteins were identified in symbiotic nematode specific transcriptome. Subject to functional validation of the identified transcripts, our findings suggest that nematode immune system plays a pivotal role in maintenance of symbiosis

with its bacterial partner. It's a two-way process - the bacteria influence core physiological functions of the host, whereas the nematode actively senses the symbionts and keeps them in check through active regulation of its immune system. Moreover, screening for dsRNA uptake in different developmental stages of *Heterorhabditis* and subsequent RNAi experiments provided the proof of concept for successful RNAi by soaking post-IJ recovery (PIJR) stage in dsRNA solution. This standardized method can be further adapted for gene function validation in *Heterorhabditis* nematodes.

In conclusion, this study generates a high-quality genome sequence for the entomopathogenic nematode *H. indica* and transcriptomic data of symbiosis relevant early adult stage of *Heterorhabditis.* RNA-seq data provides a comprehensive view of potent nematode factors involved in symbiosis with *Photorhabdus* bacteria. To the best of our knowledge, this is the first global investigation of animal factors involved in the modulation of microbial symbiosis. Lastly, a standardized protocols for RNAi based functional validation of genes in *Heterorhabditis* was reported. Genome and transcriptome resource may help to understand and to design tools to improve biocontrol traits of EPNs. These resources can be used for the investigation of evolution and functional specialization of entomopathogenic nematodes. The omics data along with standard molecular tools may promote further utilization of *Heterorhabditis* nematodes as model system to understand ubiquitous animal-microbe symbiosis and host-parasite interactions.

A molecular investigation of *Heterorhabditis* nematode factors involved in symbiosis with *Photorhabdus* bacteria

ABSTRACT

The specific association between the entomopathogenic nematode (EPN) *Heterorhabditis* and entomopathogenic bacteria (EPB) *Photorhabdus* offers a powerful model to study animal-bacterial symbiosis. Several molecular determinants of bacteria involved in the symbiotic association have been identified. However, information on nematode factors that govern the symbiosis is lacking. Additionally, genomic information of *Heterorhabditis* nematodes available in the public domain is very scant. Here, we present the first draft genome sequence of *H. indica*, which is widely present in the warmer and tropical climatic regions and is one of the most commercialized EPNs. The draft genome of the *H. indica* Hms1-i20 was sequenced using three genomic libraries of 300 bp, 600 bp and 5 kb sizes by Illumina HiSeq platform. The size of the draft genome assembly was 91.26 Mb, comprising 3,538 scaffolds. 10,494 protein-coding genes were predicted in the present genome. Comparative analysis of *H. indica* genome in comparison to four other nematode genomes revealed that *H. indica* shared 6,574 orthologous groups with *H. bacteriophora*, 6,635 with *C. elegans*, 6,228 with *S. carpocapsae* and 6,669 with *O. tipulae*. Protein domain and secretome characterization identified 2,525 transmembrane domain proteins and 370 putative secreted proteins. Subsequent analyses identified 56 GPCRs, 38 peptidases, 31 peptidase inhibitors and 2 Fatty-acid retinol binding proteins in *H. indica* proteome and such proteins may facilitate nematode interactions with insect hosts or bacterial symbionts. Additionally, 2,549 microsatellite loci, 1,548 transposable elements, 631 non-coding RNA loci, 21 likely cases of horizontal gene transfer (HGT) were identified. The mitochondrial genome of *H. indica* was assembled separately which is 17,393 bp in size and 46 genes were detected in mitochondrial genome. RNA-sequencing was employed to investigate nematode factors involved in symbiosis at the early-adult stage of *H. bacteriophora*. A total of 754 differentially expressed transcripts were identified in symbiotic nematodes. Additionally, 12,151 transcripts were unique to symbiotic nematodes. Endocytosis, cAMP signalling and focal adhesion were the top enriched pathways in symbiotic nematodes, and a large number of transcripts involved in nematode immune/defence responses against bacteria were identified. Subject to functional validation of the

identified transcripts, our findings suggest that the nematode immune system plays a pivotal role in maintenance of symbiosis with its bacterial partner. To the best of our knowledge, this is the first global investigation of animal factors involved in the modulation of microbial symbiosis. Moreover, screening for dsRNA uptake in different developmental stages of *Heterorhabditis* and subsequent RNAi experiments provided the proof of concept for successful RNAi by soaking post-IJ recovery stage in dsRNA solution. This method can be adopted for gene function validation in *Heterorhabditis* nematodes in future.

हेटेरोरैबडाइटिस सूत्रकृमि के *फोटोरैबड्स* जीवाणु के साथ सहजीविता में लिप्त कारकों का आण्विक अनुसंधान

सार

कीटरोगजनक सुत्रकृमि *हेटेरोरैबडाइटिस* एवं सहजीवी जीवाणु *फोटोरैबड्स* का विशिष्ट संबंध पशुओं और जीवाणु के बीच सहजीविता को समझने का एक सशक्त मॉडल है। जीवाणुओं के ऐसे कई कारकों के बारे में पता है जो इन सहजीवी संबंधो में महत्वपूर्ण भूमिका निभाते हैं, परन्तु ऐसे किसी भी सूत्रकृमि कारकों के बारे में कोई जानकारी उपलब्ध नहीं है। इसके अलावा इन *हेटेरोरैबडाइटिस* सूत्रकृमि के बारे में जीनोमिक जानकारी की उपलब्धता भी सीमित है। यहाँ हम *है. इंडिका* की ड्राफ्ट जीनोम सीक्वेंस प्रस्तुत करते है जो कि गर्म एवं उष्ण पर्यावरणीय क्षेत्रों में पाया जाने वाला व्यवसयिक दृष्टि से एक महतवपूर्ण सूत्रकृमि है। *है. इंडिका* ऐच.एम्.एस-आई२० की ड्राफ्ट जीनोम सीक्वेंस को ३००बीपी, ६००बीपी और ५केबी की तीन जीनोमिक लाइब्रेरियो को इल्लुमिना हाईसेक प्लेटफार्म का उपयोग करके सीक्वेंस किया गया। यह जीनोम ३५३८ स्कैफोल्ड एवं ९१.२६ एम्.बी. आकार का था। इसमें १०,४९४ प्रोटीन कोडिंग जीन पाय गए। चार अन्य सूत्रकृमि के जीनोम से तुलना करने पर पाया गया कि *है. बैक्टेरियोफोरा* के साथ ६५७४, *सीनोरैबडाइटिस एलेगंस* के साथ ६६३५, *स्टीनरनेमा कार्पोकेप्सी* के साथ ६२२८ एवं *ओशियस टिपुलि* के साथ ६६६९ ऑर्थोलोगोस जीनों में समानता है। प्रोटीन डोमेन और सीक्रेटोम के अध्ययन में २५२५ ट्रांस- मेम्ब्रेन प्रोटीन और सीक्रेटेड प्रोटीन पाय गए। इसके बाद ५६ जी पी सी आर, ३० पेप्टाइडेस, ३१ पेप्टाइडेस इन्हीबिटर एवं २ फैटी एसिड रेटिनॉल बॉन्डिंग प्रोटीन भी पाय गए जो कि इन सूत्रकृमियों के कीटों या सहजीवी जीवाणओं के साथ संबंधों में संलग्न हो सकते हैं। साथ ही २५४९ माइक्रो सॅटॅलाइट लोकस, ६३१ नॉन कोडिंग आर.एन. ए और हॉरिजॉन्टल जीन अंतरण के २१ केस पहचाने गए. एच. इंडिका का माइटोकोन्ड्रीयल जीनोम १७३९३ बीपी का था जिसमे ४६ जीन पाए गए। सहजीवी जीवाणु के साथ संबंधो में लिप्त कारकों का पता लगाने के लिए *है. बैक्टेरियोफोरा* की अग्र-व्यसक अवस्था का आर.ऍन.ए सीक्वेंसिंग के द्वारा अनुसंधान किया गया। सहजीवी सूत्रकृमियों में कुल ७५४ जीन तुलनात्मक रुप से अलग एक्सप्रेस होते पाय गए। इसके अलावा १२१५१ ट्रांसक्रिप्ट सहजीवी सूत्रकृमियों में विशिष्ट रुप से प्रकट पाय गए जो कि असहजीवी सूत्रकृमियों में नहीं थे। एंडोसाइटोसिस, सीऐऍमपी सिग्नलिंग और फोकल ऐडहेसन सहजीवी सूत्रकृमियों में सबसे प्रमुख पाथवे थी। जीवाणु प्रतिरक्षा में सम्मलित कई ट्रांस्क्रिप्टों को भी पहचाना गया। हमारे अध्ययन से ऐसा प्रतीत होता है कि सुत्रकृमि प्रतिरक्षा प्रणाली सहजीवी जीवाणुओं के साथ संबंधों को बनाय रखने में एक महत्वपूर्ण भूमिका निभाती है। हमारी जानकारी के अनुसार पशुओं के सहजीवी जीवाणुओं के संबंध के ऊपर प्रणाली स्तर का यह पहला अनुसंधान है। इसके अलावा डी.एस. आरनए में सोक करने के बाद पाया गया कि पोस्ट आई जे रिकवरी अवस्था इस सूत्रकृमि की सबसे उपयुक्त अवस्था है, जिसको कि आगे उपयोग में लाया जा सकता है। हमारी जानकारी के अनुसार पशुओं के साथ सहजीवी जीवाणुओं के संबंध के ऊपर प्रणाली स्तर का यह पहला अनुसंधान है।

BIBLIOGRAPHY

Aballay, A., and Ausubel, F. M. (2001). Programmed cell death mediated by *ced-3* and *ced-4* protects *Caenorhabditis elegans* from *Salmonella typhimurium*-mediated killing. *Proceedings of the National Academy of Sciences*, *98*(5), 2735–2739.

Abd-Elgawad, M. M. (2021). *Photorhabdus* spp.: An Overview of the Beneficial Aspects of Mutualistic Bacteria of Insecticidal Nematodes. *Plants, 10*(8), 1660.

Abubucker, S., McNulty, S. N., Rosa, B. A., and Mitreva, M. (2014). Identification and characterization of alternative splicing in parasitic nematode transcriptomes. *Parasites and vectors, 7*(1), 1-12.

Adhikari, B. N., Lin, C.-Y., Bai, X., Ciche, T. A., Grewal, P. S., Dillman, A. R., Chaston, J. M., Shapiro-Ilan, D. I., Bilgrami, A. L., Gaugler, R., Sternberg, P. W., and Adams, B. J. (2009). Transcriptional profiling of trait deterioration in the insect pathogenic nematode *Heterorhabditis bacteriophora*. *BMC Genomics*, *10*(1), 609.

Agerer, F., Lux, S., Michel, A., Rohde, M., Ohlsen, K., & Hauck, C. R. (2005). Cellular invasion by *Staphylococcus aureus* reveals a functional link between focal adhesion kinase and cortactin in integrin-mediated internalisation. *Journal of Cell Science*, *118*(10), 2189–2200.

Al-Zaidawi, J. B., Karimi, J., and Moghadam, E. M. (2019). Molecular characterizations of the entomopathogenic nematodes, *Heterorhabditis bacteriophora* and *Oscheius myriophilus* from Iraq. *Egyptian Journal of Biological Pest Control*, *29*(1), 1-9.

An, R., and Grewal, P. S. (2010). Molecular Mechanisms of Persistence of Mutualistic Bacteria *Photorhabdus* in the Entomopathogenic Nematode Host. *PLoS One*, *5*(10), e13154.

Anbesse, S., Sumaya, N. H., Dörfler, A. V., Strauch, O., and Ehlers, R. U. (2013). Selective breeding for desiccation tolerance in liquid culture provides genetically stable inbred lines of the entomopathogenic nematode *Heterorhabditis bacteriophora*. *Applied Microbiology and Biotechnology*, *97*(2), 731-739.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology*, *11*(10), R106.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online: https://www.bibsonomy.org/bibtex/f230a919c34360709aa298734d63dca3

Bager, R., Roghanian, M., Gerdes, K., and Clarke, D. J. (2016). Alarmone (p)ppGpp regulates the transition from pathogenicity to mutualism in *Photorhabdus luminescens*. *Molecular Microbiology*, *100*(4), 735–747.

Bai, X., Adams, B. J., Ciche, T. A., Clifton, S., Gaugler, R., Hogenhout, S. A., Spieth, J., Sternberg, P. W., Wilson, R. K., and Grewal, P. S. (2009). Transcriptomic analysis of the entomopathogenic nematode *Heterorhabditis bacteriophora* TTO1. *BMC Genomics*, *10*(1), 205.

Bai, X., Adams, B. J., Ciche, T. A., Clifton, S., Gaugler, R., Kim, K., Spieth, J., Sternberg, P. W., Wilson, R. K., and Grewal, P. S. (2013). A Lover and a Fighter: The Genome Sequence of an Entomopathogenic Nematode *Heterorhabditis bacteriophora*. *PLoS One*, *8*(7), e69618.

Bai, X., Grewal, P. S., Hogenhout, S. A., Adams, B. J., Ciche, T. A., Gaugler, R., and Sternberg, P. W. (2007). Expressed sequence tag analysis of gene representation in insect parasitic nematode *Heterorhabditis bacteriophora*. *The Journal of Parasitology*, *93*(6), 1343–1349.

Bai, X., Saeb, A. T. M., Michel, A., and Grewal, P. S. (2009). Isolation and characterization of microsatellite loci in the entomopathogenic nematode *Heterorhabditis bacteriophora*. *Molecular Ecology Resources*, *9*(1), 207–209.

Baniya, A., and DiGennaro, P. (2021). Genome announcement of *Steinernema khuongi* and its associated symbiont from Florida. *G3*, *11*(4), jkab053.

Bankar, K. G., Todur, V. N., Shukla, R. N., and Vasudevan, M. (2015). Ameliorated de novo transcriptome assembly using Illumina paired end sequence data with Trinity Assembler. *Genomics Data*, *5*, 352–359.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A.

(2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, *19*(5), 455–477.

Barrón-Bravo, O., Montiel-Maya, I., Cruz-Avalos, A., Avila-Ramos, F., Ochoa, J. M., and Angel-Sahagún, C. (2021). Entomopathogenic Nematodes: Biological Model of Studies with Anthelmintics. *https://www.intechopen.com/chapters/78234*

Beamer, L. J., Carroll, S. F., and Eisenberg, D. (1998). The BPI/LBP family of proteins: A structural analysis of conserved regions. *Protein Science*, *7*(4), 906–914.

Begum, K., Mohl, J. E., Ayivor, F., Perez, E. E., and Leung, M.-Y. (2020). GPCR-PEnDB: A database of protein sequences and derived features to facilitate prediction and classification of G protein-coupled receptors. *Database*, *2020*, baaa087.

Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). MISA-web: A web server for microsatellite prediction. *Bioinformatics*, *33*(16), 2583–2585.

Bennett, H. P. J., and Clarke, D. J. (2005). The pbgPE operon in *Photorhabdus luminescens* is required for pathogenicity and symbiosis. *Journal of Bacteriology*, *187*(1), 77–84.

Bernsel, A., Viklund, H., Hennerdal, A., and Elofsson, A. (2009). TOPCONS: Consensus prediction of membrane protein topology. *Nucleic Acids Research*, *37*(Web Server issue), W465–W468.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., and Stadler, P. F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*(2), 313–319.

Besemer, J., and Borodovsky, M. (2005). GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Research*, *33*, W451–W454.

Bhasin, M., and Raghava, G. P. S. (2004). GPCRpred: An SVM-based method for prediction of families and subfamilies of G-protein coupled receptors. *Nucleic Acids Research*, *32*, W383-389.

Bhat, A. H., Askary, T. H., Ahmad, M. J., Suman, Aasha, and Chaubey, A. K. (2019). Description of *Heterorhabditis bacteriophora* (Nematoda: Heterorhabditidae)

isolated from hilly areas of Kashmir Valley. *Egyptian Journal of Biological Pest Control*, *29*(1), 1-7.

Bhat, A. H., Chaubey, A. K., and Askary, T. H. (2020). Global distribution of entomopathogenic nematodes, *Steinernema* and *Heterorhabditis*. *Egyptian Journal of Biological Pest Control*, *30*(1), 1-15.

Bhat, C. G., Chauhan, K., Phani, V., Papolu, P. K., Rao, U., and Somvanshi, V. S. (2019). Expression of *Heterorhabditis bacteriophora* C-type lectins, *Hb-clec-1* and *Hb-clec-78*, in context of symbiosis with *Photorhabdus* bacteria. *Symbiosis*, *77*(1), 49–58.

Bintrim, S. B., and Ensign, J. C. (1998). Insertional Inactivation of Genes Encoding the Crystalline Inclusion Proteins of *Photorhabdus luminescens* Results in Mutants with Pleiotropic Phenotypes. *Journal of Bacteriology*, *180*(5), 1261–1269.

Blaxter, M., and Koutsovoulos, G. (2015). The evolution of parasitism in Nematoda. *Parasitology*, *142*(1), S26-39.

Boehm, A.-M., Khalturin, K., Anton-Erxleben, F., Hemmrich, G., Klostermeier, U. C., Lopez-Quintero, J. A., Oberg, H.-H., Puchert, M., Rosenstiel, P., Wittlieb, J., and Bosch, T. C. G. (2012). FoxO is a critical regulator of stem cell maintenance in immortal *Hydra*. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(48), 19697–19702.

Boemare, N. E., Akhurst, R. J., and Mourant, R. G. (1993). DNA relatedness between *Xenorhabdus* spp.(Enterobacteriaceae), symbiotic bacteria of entomopathogenic nematodes, and a proposal to transfer *Xenorhabdus luminescens* to a new genus, *Photorhabdus* gen. nov. *International Journal of Systematic and Evolutionary Microbiology, 43*(2), 249-255.

Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*, *27*(4), 578–579.

Bowen, D. J., and Ensign, J. C. (2001). Isolation and characterization of intracellular protein inclusions produced by the entomopathogenic bacterium *Photorhabdus luminescens*. *Applied and Environmental Microbiology*, *67*(10), 4834–4841.

Brameier, M., Krings, A., and MacCallum, R. M. (2007). NucPred Predicting nuclear localization of proteins. *Bioinformatics*, *23*(9), 1159–1160.

Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525–527.

Broderick, N. A. (2016). Friend, foe or food? Recognition and the role of antimicrobial peptides in gut immunity and *Drosophila* –microbe interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1695), 20150295.

Brown, E. M., Sadarangani, M., and Finlay, B. B. (2013). The role of the immune system in governing host-microbe interactions in the intestine. *Nature Immunology*, *14*(7), 660–667.

Bulgheresi, S. (2016). All the microbiology nematodes can teach us. *FEMS Microbiology Ecology*, *92*(2), fiw007.

Burnell, A., and Stock, S. P. (2000). *Heterorhabditis*, *Steinernema* and their bacterial symbionts—Lethal pathogens of insects. *Nematology*, *2*(1), 31–42.

Burton, E. A., Pendergast, A. M., and Aballay, A. (2006). The *Caenorhabditis elegans* ABL-1 tyrosine kinase is required for *Shigella flexneri* pathogenesis. *Applied and Environmental Microbiology*, *72*(7), 5043–5051.

Campos–Herrera, R., Barbercheck, M., Hoy, C. W., and Stock, S. P. (2012). Entomopathogenic Nematodes as a Model System for Advancing the Frontiers of Ecology. *Journal of Nematology*, *44*(2), 162–176.

Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez Alvarado, A., and Yandell, M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, *18*(1), 188–196.

Chalabaev, S., Turlin, E., Bay, S., Ganneau, C., Brito-Fravallo, E., Charles, J.-F., Danchin, A., and Biville, F. (2008). Cinnamic acid, an autoinducer of its own biosynthesis, is processed via Hca enzymes in *Photorhabdus luminescens*. *Applied and Environmental Microbiology*, *74*(6), 1717–1725.

Chang, D. Z., Serra, L., Lu, D., Mortazavi, A., and Dillman, A. R. (2019). A core set of venom proteins is released by entomopathogenic nematodes in the genus *Steinernema*. *PLoS Pathogens*, *15*(5), e1007626.

Chaston, J., and Goodrich-Blair, H. (2010). Common trends in mutualism revealed by model associations between invertebrates and bacteria. *FEMS Microbiology Reviews*, *34*(1), 41–58.

Chaston, J. M., Suen, G., Tucker, S. L., Andersen, A. W., Bhasin, A., Bode, E., Bode, H. B., Brachmann, A. O., Cowles, C. E., Cowles, K. N., Darby, C., de Léon, L., Drace, K., Du, Z., Givaudan, A., Herbert Tran, E. E., Jewell, K. A., Knack, J. J., Krasomil-Osterfeld, K. C., and Goodrich-Blair, H. (2011). The entomopathogenic bacterial endosymbionts *Xenorhabdus* and *Photorhabdus*: Convergent lifestyles from divergent genomes. *PloS One*, *6*(11), e27909.

Chávez, V., Mohri-Shiomi, A., & Garsin, D. A. (2009). Ce-Duox1/BLI-3 Generates Reactive Oxygen Species as a Protective Innate Immune Mechanism in *Caenorhabditis elegans*. *Infection and Immunity*, *77*(11), 4983–4989.

Chávez, V., Mohri-Shiomi, A., Maadani, A., Vega, L. A., & Garsin, D. A. (2007). Oxidative Stress Enzymes Are Required for DAF-16-Mediated Immunity Due to Generation of Reactive Oxygen Species by *Caenorhabditis elegans*. *Genetics*, *176*(3), 1567–1577.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890.

Ciche, T. (2007). The biology and genome of *Heterorhabditis bacteriophora*, In WormBook, ed. The *C. elegans* Research Community, WormBook, Doi/10.1895/wormbook.1.135.1, http://www.wormbook.org.

Ciche, T. A., Bintrim, S. B., Horswill, A. R., & Ensign, J. C. (2001). A Phosphopantetheinyl transferase homolog is essential for *Photorhabdus luminescens* to support growth and reproduction of the entomopathogenic nematode *Heterorhabditis bacteriophora*. *Journal of Bacteriology*, *183*(10), 3117–3126.

Ciche, T. A., Kim, K., Kaufmann-Daszczuk, B., Nguyen, K. C. Q., and Hall, D. H. (2008). Cell Invasion and Matricide during *Photorhabdus luminescens* Transmission by *Heterorhabditis bacteriophora* Nematodes. *Applied and Environmental Microbiology*, *74*(8), 2275–2287.

Ciche, T. A., and Sternberg, P. W. (2007). Postembryonic RNAi in *Heterorhabditis bacteriophora*: A nematode insect parasite and host for insect pathogenic symbionts. *BMC Developmental Biology*, *7*(1), 101.

Clarke, D. J. (2020). *Photorhabdus*: A tale of contrasting interactions. *Microbiology*, *166*(4), 335–348.

Cohen, L. B., and Troemel, E. R. (2015). Microbial pathogenesis and host defense in the nematode *C. elegans*. *Current Opinion in Microbiology*, *23*, 94–101.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics, 21*(18), 3674–3676.

Curt, A., Zhang, J., Minnerly, J., and Jia, K. (2014). Intestinal autophagy activity is essential for host defense against *Salmonella typhimurium* infection in *Caenorhabditis elegans*. *Developmental and Comparative Immunology*, *45*(2), 214–218.

Dalzell, J. J., McVeigh, P., Warnock, N. D., Mitreva, M., Bird, D. M., Abad, P., Fleming, C. C., Day, T. A., Mousley, A., Marks, N. J., and Maule, A. G. (2011). RNAi effector diversity in nematodes. *PLoS Neglected Tropical Diseases*, *5*(6), e1176.

Danchin, E., and Rosso, M.-N. (2012). Lateral gene transfers have polished animal genomes: Lessons from nematodes. *Frontiers in Cellular and Infection Microbiology*, *2,* 27.

De Ley, P. A. (2006). A quick tour of nematode diversity and the backbone of nematode phylogeny. In WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.41.1, http://www.wormbook.org.

Dierking, K., Yang, W., and Schulenburg, H. (2016). Antimicrobial effectors in the nematode *Caenorhabditis elegans*: An outgroup to the Arthropoda. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1695), 20150299.

Dillman, A. R., Chaston, J. M., Adams, B. J., Ciche, T. A., Goodrich-Blair, H., Stock, S. P., and Sternberg, P. W. (2012). An Entomopathogenic Nematode by Any Other Name. *PLoS Pathogens*, *8*(3), e1002527.

Dillman, A. R., Macchietto, M., Porter, C. F., Rogers, A., Williams, B., Antoshechkin, I., Lee, M.-M., Goodwin, Z., Lu, X., Lewis, E. E., Goodrich-Blair, H., Stock, S. P., Adams, B. J., Sternberg, P. W., and Mortazavi, A. (2015). Comparative genomics of *Steinernema* reveals deeply conserved gene regulatory networks. *Genome Biology*, *16*(1), 200.

Doolittle, W. F., and Booth, A. (2017). It's the song, not the singer: An exploration of holobiosis and evolutionary theory. *Biology & Philosophy*, *32*(1), 5–24.

Douglas, A. E. (2014). The molecular basis of bacterial-insect symbiosis. *Journal of Molecular Biology*, *426*(23), 3830–3837.

Duchaud, E., Rusniok, C., Frangeul, L., Buchrieser, C., Givaudan, A., Taourit, S., Bocs, S., Boursaux-Eude, C., Chandler, M., Charles, J.-F., Dassa, E., Derose, R., Derzelle, S., Freyssinet, G., Gaudriault, S., Médigue, C., Lanois, A., Powell, K., Siguier, P., and Kunst, F. (2003). The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens. Nature Biotechnology*, *21*(11), 1307–1313.

Dunbar, T. L., Yan, Z., Balla, K. M., Smelkinson, M. G., and Troemel, E. R. (2012). *C. elegans* Detects Pathogen-Induced Translational Inhibition to Activate Immune Signaling. *Cell Host & Microbe*, *11*(4), 375–386.

Duperron, S. (2017). 5—Symbiosis and Evolution. In S. Duperron (Ed.), *Microbial Symbioses* (pp. 79–107). Elsevier, ISBN 9781785482205, https://doi.org/10.1016/B978-1-78548-220-5.50005-2.

Easom, C. A., and Clarke, D. J. (2012). HdfR is a regulator in *Photorhabdus luminescens* that modulates metabolism and symbiosis with the nematode *Heterorhabditis*. *Environmental Microbiology*, *14*(4), 953–966.

Easom, C. A., Joyce, S. A., and Clarke, D. J. (2010). Identification of genes involved in the mutualistic colonization of the nematode *Heterorhabditis bacteriophora* by the bacterium *Photorhabdus luminescens*. *BMC Microbiology*, *10*(1), 45.

Eberl, G. (2010). A new vision of immunity: Homeostasis of the superorganism. *Mucosal Immunology*, *3*(5), 450–460.

End, C., Bikker, F., Renner, M., Bergmann, G., Lyer, S., Blaich, S., Hudler, M., Helmke, B., Gassler, N., Autschbach, F., Ligtenberg, A. J. M., Benner, A.,

Holmskov, U., Schirmacher, P., Nieuw Amerongen, A. V., Rosenstiel, P., Sina, C., Franke, A., Hafner, M., and Mollenhauer, J. (2009). DMBT1 functions as pattern-recognition molecule for poly-sulfated and poly-phosphorylated ligands. *European Journal of Immunology*, *39*(3), 833–842.

Engel, Y., Windhorst, C., Lu, X., Goodrich-Blair, H., and Bode, H. B. (2017). The Global Regulators Lrp, LeuO, and HexA Control Secondary Metabolism in Entomopathogenic Bacteria. *Frontiers in Microbiology*, *8*, 209.

Engelmann, I., and Pujol, N. (2010). Innate immunity in C. elegans. *Advances in Experimental Medicine and Biology*, *708*, 105–121.

Félix, M.-A. (2008). RNA interference in nematodes and the chance that favored Sydney Brenner. *Journal of Biology*, *7*(9), 34.

Fielenbach, N., and Antebi, A. (2008). *C. elegans* dauer formation and the molecular basis of plasticity. *Genes & Development*, *22*(16), 2149–2165.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, *391*(6669), 806–811.

Fischer-Le Saux, M., Viallard, V., Brunel, B., Normand, P., and Boemare, N. E. (1999). Polyphasic classification of the genus *Photorhabdus* and proposal of new taxa: *P. luminescens* subsp. *luminescens* subsp. nov., *P. luminescens* subsp. *akhurstii* subsp. nov., *P. luminescens* subsp. *laumondii* subsp. nov., *P. temperata* sp. nov., *P. temperata* subsp. *temperata* subsp. nov. and *P. asymbiotica* sp. nov. *International Journal of Systematic Bacteriology*, *49 Pt 4*, 1645–1656.

Frantz, A. L., Rogier, E. W., Weber, C. R., Shen, L., Cohen, D. A., Fenton, L. A., Bruno, M. E. C., and Kaetzel, C. S. (2012). Targeted deletion of MyD88 in intestinal epithelial cells results in compromised antibacterial immunity associated with downregulation of polymeric immunoglobulin receptor, mucin-2, and antibacterial peptides. *Mucosal Immunology*, *5*(5), 501–512.

Froger, A., and Hall, J. E. (2007). Transformation of plasmid DNA into *E. coli* using the heat shock method. *Journal of Visualized Experiments: JoVE*, *6*, 253.

Fukasawa, Y., Tsuji, J., Fu, S.-C., Tomii, K., Horton, P., and Imai, K. (2015). MitoFates: Improved prediction of mitochondrial targeting sequences and their cleavage sites. *Molecular & Cellular Proteomics: MCP*, *14*(4), 1113–1126.

Gattiker, A., Gasteiger, E., and Bairoch, A. (2002). ScanProsite: A reference implementation of a PROSITE scanning tool. *Applied Bioinformatics*, *1*(2), 107–108.

Gaudriault, S., Duchaud, E., Lanois, A., Canoy, A.-S., Bourot, S., DeRose, R., Kunst, F., Boemare, N., and Givaudan, A. (2006). Whole-Genome Comparison between *Photorhabdus* Strains To Identify Genomic Regions Involved in the Specificity of Nematode Interaction. *Journal of Bacteriology*, *188*(2), 809–814.

Gaugler, R. (Ed.). (2002). *Entomopathogenic nematology*. Wallingford, UK: CAB International. https://www.cabi.org/ISC/ebook/20023023437.

Geldhof, P., Murray, L., Couthier, A., Gilleard, J. S., McLauchlan, G., Knox, D. P., and Britton, C. (2006). Testing the efficacy of RNA interference in *Haemonchus contortus*. *International Journal for Parasitology*, *36*(7), 801–810.

Geldhof, P., Visser, A., Clark, D., Saunders, G., Britton, C., Gilleard, J., Berriman, M., and Knox, D. (2007). RNA interference in parasitic helminths: Current situation, potential pitfalls and future prospects. *Parasitology*, *134*(Pt 5), 609–619.

Gerardo, N. M., Hoang, K. L., and Stoy, K. S. (2020). Evolution of animal immunity in the light of beneficial symbioses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1808), 20190601.

Gilbert, S. F., Sapp, J., and Tauber, A. I. (2012). A symbiotic view of life: We have never been individuals. *The Quarterly Review of Biology*, *87*(4), 325–341.

Gonzalez de la Rosa, P. M., Thomson, M., Trivedi, U., Tracey, A., Tandonnet, S., and Blaxter, M. (2020). A telomere-to-telomere assembly of *Oscheius tipulae* and the evolution of rhabditid nematode chromosomes. *G3, 11*(1), https://doi.org/10.1093/g3journal/jkaa020.

Gozel, U., and Gozel, C. (2016). Entomopathogenic Nematodes in Pest Management. In H. K. Gill, and G. Goyal (Eds.), *Integrated Pest Management (IPM):*

Environmentally Sound Pest Management. IntechOpen. https://doi.org/10.5772/63894

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652.

Haegeman, A., Jones, J. T., and Danchin, E. G. J. (2011). Horizontal gene transfer in nematodes: A catalyst for plant parasitism? *Molecular Plant-Microbe Interactions: MPMI*, *24*(8), 879–887.

Hallem, E. A., Rengarajan, M., Ciche, T. A.,and Sternberg, P. W. (2007). Nematodes, bacteria, and flies: A tripartite model for nematode parasitism. *Current Biology: CB*, *17*(10), 898–904.

Han, R., and Ehlers, R. U. (2000). Pathogenicity, development, and reproduction of *Heterorhabditis bacteriophora* and *Steinernema carpocapsae* under axenic in vivo conditions. *Journal of Invertebrate Pathology*, *75*(1), 55–58.

Hapeshi, A., and Waterfield, N. R. (2017). *Photorhabdus asymbiotica* as an Insect and Human Pathogen. *Current Topics in Microbiology and Immunology*, *402*, 159–177.

Hashmi, S., Hashmi, G., and Gaugler, R. (1995). Genetic transformation of an entomopathogenic nematode by microinjection. *Journal of Invertebrate Pathology*, *66*(3), 293-296.

Hashmi, S., Hashmi, G., Glazer, I., and Gaugler, R. (1998). Thermal response of *Heterorhabditis bacteriophora* transformed with the *Caenorhabditis elegans* hsp70 encoding gene. *The Journal of Experimental Zoology*, *281*(3), 164–170.

Hasi-Imi, S., Hatab, M. A. A., and Gaugler, R. R. (1997). *GFP:* green fluorescent protein a versatile gene marker for entomopathogenic nematodes. *Fundamental and Applied nematology, 20*(4), 323-327.

Haskins, K. A., Russell, J. F., Gaddis, N., Dressman, H. K., and Aballay, A. (2008). Unfolded Protein Response Genes Regulated by CED-1 are Required for *Caenorhabditis elegans* Innate Immunity. *Developmental Cell*, *15*(1), 87–97.

Hay, F. S. (2008). Nematodes-the good, the bad and the ugly. https://www.apsnet.org/edcenter/resources/archive/NewsViews/Pages/Nematodes.aspx

Hodda, M., Peters, L., and Traunspurger, W. (2009). Nematode diversity in terrestrial, freshwater aquatic and marine systems. In M. J. Wilson and T. Kakouli-Duarte (Eds.), *Nematodes as environmental indicators* (pp. 45–93). Wallingford, UK: CAB International. https://www.cabi.org/cabebooks/ebook/20093184806

Holterman, M., van der Wurff, A., van den Elsen, S., van Megen, H., Bongers, T., Holovachov, O., Bakker, J., and Helder, J. (2006). Phylum-Wide Analysis of SSU rDNA Reveals Deep Phylogenetic Relationships among Nematodes and Accelerated Evolution toward Crown Clades. *Molecular Biology and Evolution*, *23*(9), 1792–1800.

Hooper, L. V., Littman, D. R., and Macpherson, A. J. (2012). Interactions between the microbiota and the immune system. *Science, 336*(6086), 1268–1273.

Hu, K., and Webster, J. M. (2000). Antibiotic production in relation to bacterial growth and nematode development in *Photorhabdus–Heterorhabditis* infected *Galleria mellonella* larvae. *FEMS Microbiology Letters*, *189*(2), 219–223.

Huang, D. W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M. W., Lane, H. C., and Lempicki, R. A. (2007). DAVID Bioinformatics Resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Research*, *35*, W169–W175.

Jia, K., Thomas, C., Akbar, M., Sun, Q., Adams-Huet, B., Gilpin, C., and Levine, B. (2009). Autophagy genes protect against *Salmonella typhimurium* infection and mediate insulin signaling-regulated pathogen resistance. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(34), 14564–14569.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., and Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics, 30*(9), 1236–1240.

Josse, J., Laurent, F., and Diot, A. (2017). Staphylococcal Adhesion and Host Cell Invasion: Fibronectin-Binding and Other Mechanisms. *Frontiers in Microbiology*, *8*, 2433.

Joyce, S. A., Brachmann, A. O., Glazer, I., Lango, L., Schwär, G., Clarke, D. J., and Bode, H. B. (2008). Bacterial biosynthesis of a multipotent stilbene. *Angewandte Chemie (International Ed. in English)*, *47*(10), 1942–1945.

Joyce, S. A., and Clarke, D. J. (2003). A *hexA* homologue from *Photorhabdus* regulates pathogenicity, symbiosis and phenotypic variation. *Molecular Microbiology*, *47*(5), 1445–1457.

Joyce, S. A., Lango, L., and Clarke, D. J. (2011). The Regulation of Secondary Metabolism and Mutualism in the Insect Pathogenic Bacterium *Photorhabdus luminescens*. *Advances in Applied Microbiology*, *76*, 1–25.

Jühling, F., Pütz, J., Bernt, M., Donath, A., Middendorf, M., Florentz, C., and Stadler, P. F. (2012). Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. *Nucleic Acids Research*, *40*(7), 2833–2845.

Käll, L., Krogh, A., & Sonnhammer, E. L. L. (2007). Advantages of combined transmembrane topology and signal peptide prediction—The Phobius web server. *Nucleic Acids Research*, *35*, W429-432.

Karp, X. (2018). Working with dauer larvae. In WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.180.1, http://www.wormbook.org.

Kaya, H. K., Aguillera, M. M., Alumai, A., Choo, H. Y., de la Torre, M., Fodor, A., Ganguly, S., Hazır, S., Lakatos, T., Pye, A., Wilson, M., Yamanaka, S., Yang, H., and Ehlers, R.-U. (2006). Status of entomopathogenic nematodes and their

symbiotic bacteria from selected countries or regions of the world. *Biological Control*, *38*(1), 134–155.

Kaya, H. K., and Gaugler, R. (1993). Entomopathogenic Nematodes. *Annual Review of Entomology*, *38*(1), 181–206.

Kaya, H. K., and Patricia Stock, S. (1997). Chapter VI - Techniques in insect nematology. In L. A. Lacey (Ed.), *Manual of Techniques in Insect Pathology* (pp. 281–324). Academic Press. ISBN 9780124325555, https://doi.org/10.1016/B978-012432555-5/50016-6.

Kikuchi, T., Afrin, T., and Yoshida, M. (2016). Complete mitochondrial genomes of four entomopathogenic nematode species of the genus *Steinernema*. *Parasites and Vectors*, *9*(1), 430.

Kim, D. H., and Ewbank, J. J. (2018). Signaling in the innate immune response. In WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.83.2, http://www.wormbook.org.

Kim, D. H., Liberati, N. T., Mizuno, T., Inoue, H., Hisamoto, N., Matsumoto, K., and Ausubel, F. M. (2004). Integration of *Caenorhabditis elegans* MAPK pathways mediating immunity and stress resistance by MEK-1 MAPK kinase and VHP-1 MAPK phosphatase. *Proceedings of the National Academy of Sciences*, *101*(30), 10990–10994.

Kim, M.-S., Pinto, S. M., Getnet, D., Nirujogi, R. S., Manda, S. S., Chaerkady, R., Madugundu, A. K., Kelkar, D. S., Isserlin, R., Jain, S., Thomas, J. K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabuddhe, N. A., Balakrishnan, L., Advani, J., George, B., Renuse, S., … Pandey, A. (2014). A draft map of the human proteome. *Nature*, *509*(7502), 575–581.

Kiontke, K., and Fitch, D. H. A. (2013). Nematodes. *Current Biology*, *23*(19), R862–R864.

Knox, D. P., Geldhof, P., Visser, A., and Britton, C. (2007). RNA interference in parasitic nematodes of animals: A reality check? *Trends in Parasitology*, *23*(3), 105–107.

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics*, *5*(1), 59.

Koropatnick, T. A., Engle, J. T., Apicella, M. A., Stabb, E. V., Goldman, W. E., and McFall-Ngai, M. J. (2004). Microbial factor-mediated development in a host-bacterial mutualism. *Science, 306*(5699), 1186–1188.

Kozlowski, D. K. L., Hassanaly-Goulamhoussen, R., Da Rocha, M., Koutsovoulos, G. D., Bailly-Bechet, M., and Danchin, E. G. J. (2021). Movements of transposable elements contribute to the genomic plasticity and species diversification in an asexually reproducing nematode pest. *Evolutionary Applications*, *14*(7), 1844–1866.

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology*, *305*(3), 567–580.

Kumar, P., Ganguly, S., and Somvanshi, V. S. (2015). Identification of virulent entomopathogenic nematode isolates from a countrywide survey in India. *International Journal of Pest Management*, *61*(2), 135–143.

Kumar, P., Jamal, W., Somvanshi, V. S., Chauhan, K., and Mumtaz, S. (2019). Description of *Oscheius indicus* n. sp. (Rhabditidae: Nematoda) from India. *Journal of Nematology*, *51*, e2019-04.

Lacey, L. A., and Georgis, R. (2012). Entomopathogenic Nematodes for Control of Insect Pests Above and Below Ground with Comments on Commercial Production. *Journal of Nematology*, *44*(2), 218–225.

Lacey, L. A., Grzywacz, D., Shapiro-Ilan, D. I., Frutos, R., Brownbridge, M., and Goettel, M. S. (2015). Insect pathogens as biological control agents: Back to the future. *Journal of Invertebrate Pathology*, *132*, 1–41.

Laetsch, D. R., and Blaxter, M. L. (2017). BlobTools: Interrogation of genome assemblies. *F1000Research*, *6*, 1287.

Lambert, K., and Bekal, S. (2002). The History of Nematodes. Introduction to Plant-Parasitic Nematodes, *Department of Crop Sciences*. https://www.apsnet.org/edcenter/disandpath/nematode/intro/Pages/IntroNematodes.aspx

Langer, A., Moldovan, A., Harmath, C., Joyce, S. A., Clarke, D. J., and Heermann, R. (2017). HexA is a versatile regulator involved in the control of phenotypic heterogeneity of *Photorhabdus luminescens*. *PLoS One*, *12*(4), e0176535.

Lango, L., and Clarke, D. J. (2010). A metabolic switch is involved in lifestyle decisions in *Photorhabdus luminescens*. *Molecular Microbiology*, *77*(6), 1394–1405.

Lendner, M., Doligalska, M., Lucius, R., and Hartmann, S. (2008). Attempts to establish RNA interference in the parasitic nematode *Heligmosomoides polygyrus*. *Molecular and Biochemical Parasitology*, *161*(1), 21–31.

Lephoto, T. E., Mpangase, P. T., Aron, S., and Gray, V. M. (2016). Whole genome sequence of *Oscheius* sp. TEL-2014 entomopathogenic nematodes isolated from South Africa. *Genomics Data*, *7*, 259–261.

Levy, N., Faigenboim, A., Salame, L., Molina, C., Ehlers, R.-U., Glazer, I., and Ment, D. (2020). Characterization of the phenotypic and genotypic tolerance to abiotic stresses of natural populations of *Heterorhabditis bacteriophora*. *Scientific Reports*, *10*(1), 10500.

Li, B., and Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, *12*, 323.

Li, J., and Liu, C. (2019). Coding or Noncoding, the Converging Concepts of RNAs. *Frontiers in Genetics*, *10*, 496.

Li, L., Stoeckert, C. J., and Roos, D. S. (2003). OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Research*, *13*(9), 2178–2189.

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., Li, S., Yang, H., Wang, J., & Wang, J. (2010). De novo assembly of human genomes with massively parallel short read sequencing. *Genome Research*, *20*(2), 265–272.

Li, W., and Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)*, *22*(13), 1658–1659.

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, *25*(4), 402–408.

Lu, D., Baiocchi, T., and Dillman, A. R. (2016). Genomics of Entomopathogenic Nematodes and Implications for Pest Control. *Trends in Parasitology*, *32*(8), 588–598.

Lulamba, T. E., Green, E., and Serepa-Dlamini, M. H. (2021). Genome assembly and annotation of *Photorhabdus heterorhabditis* strain ETL reveals genetic features involved in pathogenicity with its associated entomopathogenic nematode and anti-host effectors with biocontrol potential applications. *Gene*, *795*, 145780.

Machado, R. A. R., Somvanshi, V. S., Muller, A., Kushwah, J., and Bhat, C. G. (2021). *Photorhabdus hindustanensis* sp. Nov., *Photorhabdus akhurstii* subsp. *akhurstii* subsp. Nov., and *Photorhabdus akhurstii* subsp. *bharatensis* subsp. Nov., isolated from *Heterorhabditis* entomopathogenic nematodes. *International Journal of Systematic and Evolutionary Microbiology*, *71*(9).

Mandel, M., and Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology*, *53*(1), 159–162.

Maneesakorn, P., An, R., Daneshvar, H., Taylor, K., Bai, X., Adams, B. J., Grewal, P. S., & Chandrapatya, A. (2011). Phylogenetic and cophylogenetic relationships of entomopathogenic nematodes (*Heterorhabditis*: Rhabditida) and their symbiotic bacteria (*Photorhabdus*: Enterobacteriaceae). *Molecular Phylogenetics and Evolution*, *59*(2), 271–280.

Margulis, L., and Fester, R. (Eds.). (1991). *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis*. Cambridge, MA: MIT Press. https://mitpress.mit.edu/books/symbiosis-source-evolutionary-innovation

Marsh, E. K., Berg, M. C. W. van den, and May, R. C. (2011). A Two-Gene Balance Regulates *Salmonella typhimurium* Tolerance in the Nematode *Caenorhabditis elegans*. *PLoS One*, *6*(3), e16839.

Maule, A. G., McVeigh, P., Dalzell, J. J., Atkinson, L., Mousley, A., and Marks, N. J. (2011). An eye on RNAi in nematode parasites. *Trends in Parasitology*, *27*(11), 505–513.

McEwan, D. L., Kirienko, N. V., and Ausubel, F. M. (2012). Host Translational Inhibition by *Pseudomonas aeruginosa* Exotoxin A Triggers an Immune Response in *Caenorhabditis elegans*. *Cell Host & Microbe*, *11*(4), 364–374.

McFall-Ngai, M. (2008). Are biologists in “future shock”? Symbiosis integrates biology across domains. *Nature Reviews Microbiology*, *6*(10), 789–792.

Mckenzie Bird, D. (2005). Model systems in agriculture: Lessons from worms. *Annals of Applied Biology*, *146*(2), 147–154.

McLean, F., Berger, D., Laetsch, D. R., Schwartz, H. T., and Blaxter, M. (2018). Improving the annotation of the *Heterorhabditis bacteriophora* genome. *GigaScience*, *7*(4), giy034.

Means, T. K., Mylonakis, E., Tampakakis, E., Colvin, R. A., Seung, E., Puckett, L., Tai, M. F., Stewart, C. R., Pukkila-Worley, R., Hickman, S. E., Moore, K. J., Calderwood, S. B., Hacohen, N., Luster, A. D., and El Khoury, J. (2009). Evolutionarily conserved recognition and innate immunity to fungal pathogens by the scavenger receptors SCARF1 and CD36. *The Journal of Experimental Medicine*, *206*(3), 637–653.

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

Melo, J. A., and Ruvkun, G. (2012). Inactivation of conserved genes induces microbial aversion, drug detoxification, and innate immunity in *C. elegans*. *Cell*, *149*(2), 452–466.

Meneely, P. M., Dahlberg, C. L., and Rose, J. K. (2019). Working with Worms: *Caenorhabditis elegans* as a Model Organism. *Current Protocols Essential Laboratory Techniques*, *19*(1), e35.

Mizuno, T., Hisamoto, N., Terada, T., Kondo, T., Adachi, M., Nishida, E., Kim, D. H., Ausubel, F. M., and Matsumoto, K. (2004). The *Caenorhabditis elegans* MAPK phosphatase VHP-1 mediates a novel JNK-like signaling pathway in stress response. *The EMBO Journal*, *23*(11), 2226–2234.

Montiel, R., Lucena, M. A., Medeiros, J., and Simões, N. (2006). The complete mitochondrial genome of the entomopathogenic nematode *Steinernema carpocapsae*: Insights into nematode mitochondrial DNA evolution and phylogeny. *Journal of Molecular Evolution*, *62*(2), 211–225.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., and Kanehisa, M. (2007). KAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, *35*, W182–W185.

Morris, R., Wilson, L., Sturrock, M., Warnock, N. D., Carrizo, D., Cox, D., Maule, A. G., and Dalzell, J. J. (2017). A neuropeptide modulates sensory perception in the entomopathogenic nematode *Steinernema carpocapsae*. *PLoS Pathogens*, *13*(3), e1006185.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, *5*(7), 621–628.

Moshayov, A., Koltai, H., and Glazer, I. (2013). Molecular characterisation of the recovery process in the entomopathogenic nematode *Heterorhabditis bacteriophora*. *International Journal for Parasitology*, *43*(10), 843–852.

Muñoz-López, M., and García-Pérez, J. L. (2010). DNA Transposons: Nature and Applications in Genomics. *Current Genomics*, *11*(2), 115–128.

Murfin, K. E., Dillman, A. R., Foster, J. M., Bulgheresi, S., Slatko, B. E., Sternberg, P. W., and Goodrich-Blair, H. (2012). Nematode-Bacterium Symbioses—Cooperation and Conflict Revealed in the "Omics" Age. *The Biological Bulletin*, *223*(1), 85–102.

Nawrocki, E. P., and Eddy, S. R. (2013). Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*, *29*(22), 2933–2935.

Nicholas, H. R., and Hodgkin, J. (2004). Responses to infection and possible recognition strategies in the innate immune system of *Caenorhabditis elegans*. *Molecular Immunology*, *41*(5), 479–493.

Nielsen, H. (2017). Predicting Secretory Proteins with SignalP. *Methods in Molecular Biology, 1611*, 59–73.

Nigon, V. M., and Félix, M.-A. (2018). History of research on *C. elegans* and other free-living nematodes as model organisms. In WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.181.1, http://www.wormbook.org.

Nuez, I., and Félix, M.-A. (2012). Evolution of Susceptibility to Ingested Double-Stranded RNAs in *Caenorhabditis* Nematodes. *PLoS One*, *7*(1), e29811.

Nyholm, S. V., and Graf, J. (2012). Knowing your friends: Invertebrate innate immunity fosters beneficial bacterial symbioses. *Nature Reviews Microbiology*, *10*(12), 815-827.

Ogier, J.-C., Pagès, S., Frayssinet, M., and Gaudriault, S. (2020). Entomopathogenic nematode-associated microbiota: From monoxenic paradigm to pathobiome. *Microbiome*, *8*(1), 25.

O'Leary, S. A., and Burnell, A. M. (1997). Isolation of mutants of *Heterorhabditis megidis* (Strain UK211) with increased desiccation tolerance. *Fundamental and Applied Nematology, 20*(2), 197-205.

Otero, L., Romanelli-Cedrez, L., Turanov, A. A., Gladyshev, V. N., Miranda-Vizuete, A., and Salinas, G. (2014). Adjustments, extinction, and remains of selenocysteine incorporation machinery in the nematode lineage. *RNA, 20*(7), 1023–1034.

Parks, S. C., Nguyen, S., Nasrolahi, S., Bhat, C., Juncaj, D., Lu, D., Ramaswamy, R., Dhillon, H., Fujiwara, H., Buchman, A., Akbari, O. S., Yamanaka, N., Boulanger, M. J., & Dillman, A. R. (2021). Parasitic nematode fatty acid- and retinol-binding proteins compromise host immunity by interfering with host lipid signaling pathways. *PLoS Pathogens*, *17*(10), e1010027.

Parra, G., Bradnam, K., and Korf, I. (2007). CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics, 23*(9), 1061–1067.

Pees, B., Yang, W., Kloock, A., Petersen, C., Peters, L., Fan, L., Friedrichsen, M., Butze, S., Zárate-Potes, A., Schulenburg, H., and Dierking, K. (2021). Effector and regulator: Diverse functions of *C. elegans* C-type lectin-like domain proteins. *PLoS Pathogens*, *17*(4), e1009454.

Perreau, J., and Moran, N. A. (2022). Genetic innovations in animal-microbe symbioses. *Nature Reviews. Genetics*, *23*(1), 23–39.

Pierleoni, A., Martelli, P. L., and Casadio, R. (2008). PredGPI: a GPI-anchor predictor. *BMC Bioinformatics, 9*(1), 1-11.

Poinar, G. O. (1975). Description and Biology of a New Insect Parasitic Rhabditoid, *Heterorhabditis bacteriophora* N. Gen., N. Sp. (Rhabditida; Heterorhabditidae N. Fam.). *Nematologica*, *21*(4), 463–470.

Poinar, G. O., Karunakar, G. K., and David, H. (1992). *Heterorhabditis indicus* n.sp. *(Rhabditida:* Nematoda) from India: Separation of *Heterorhabditis* spp. by infective juveniles. *Fundamental and Applied Nematology, 15*(5), 467-472.

Propheter, D. C., Chara, A. L., Harris, T. A., Ruhn, K. A., and Hooper, L. V. (2017). Resistin-like molecule β is a bactericidal protein that promotes spatial segregation of the microbiota and the colonic epithelium. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(42), 11027–11033.

Rahimi, F. R., McGuire, T. R., & Gaugler, R. (1993). Morphological mutant in the entomopathogenic nematode, *Heterorhabditis bacteriophora*. *The Journal of Heredity*, *84*(6), 475–478.

Rancurel, C., Legrand, L., and Danchin, E. G. J. (2017). Alienness: Rapid Detection of Candidate Horizontal Gene Transfers across the Tree of Life. *Genes*, *8*(10), 248.

Rasmann, S., Köllner, T. G., Degenhardt, J., Hiltpold, I., Toepfer, S., Kuhlmann, U., Gershenzon, J., and Turlings, T. C. J. (2005). Recruitment of entomopathogenic nematodes by insect-damaged maize roots. *Nature*, *434*(7034), 732–737.

Ratnappan, R., Vadnal, J., Keaney, M., Eleftherianos, I., O'Halloran, D., and Hawdon, J. M. (2016). RNAi-mediated gene knockdown by microinjection in the model entomopathogenic nematode *Heterorhabditis bacteriophora*. *Parasites & Vectors*, *9*(1), 160.

Rawlings, N. D., Barrett, A. J., Thomas, P. D., Huang, X., Bateman, A., and Finn, R. D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research*, *46*(D1), D624–D632.

Regeai, S. O., Fitzpatrick, D. A., Burnell, A. M., and Kakouli-Duarte, T. (2021). Characterisation of the complete mitochondrial genome of the insect-parasitic nematode *Heterorhabditis bacteriophora*: An idiosyncratic gene order and the presence of multiple long non-coding regions. *Nematology*, *24*(1), 37–54.

Rehman, S., Gupta, V. K., and Goyal, A. K. (2016). Identification and functional analysis of secreted effectors from phytoparasitic nematodes. *BMC Microbiology*, *16*(1), 48.

Riehl, K., Riccio, C., Miska, E. A., and Hemberg, M. (2021). TransposonUltimate: Software for transposon classification, annotation and detection. *bioRxiv*. https://doi.org/10.1101/2021.04.30.442214

Robertson, H. M., and Thomas, J. H. (2006). The putative chemoreceptor families of *C. elegans*. In WormBook, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.66.1, http://www.wormbook.org.

Rödelsperger, C., Ebbing, A., Sharma, D. R., Okumura, M., Sommer, R. J., and Korswagen, H. C. (2020). Spatial Transcriptomics of Nematodes Identifies Sperm Cells as a Source of Genomic Novelty and Rapid Evolution. *Molecular Biology and Evolution*, *38*(1), 229–243.

Rougon-Cardoso, A., Flores-Ponce, M., Ramos-Aboites, H. E., Martínez-Guerrero, C. E., Hao, Y.-J., Cunha, L., Rodríguez-Martínez, J. A., Ovando-Vázquez, C., Bermúdez-Barrientos, J. R., Abreu-Goodger, C., Chavarría-Hernández, N., Simões, N., and Montiel, R. (2016). The genome, transcriptome, and proteome of the nematode *Steinernema carpocapsae*: Evolutionary signatures of a pathogenic lifestyle. *Scientific Reports*, *6*(1), 37536.

Ruby, E. G. (2008). Symbiotic conversations are revealed under genetic interrogation. *Nature Reviews Microbiology*, *6*(10), 752–762.

Ryu, J. H., Kim, C. H., and Yoon, J. H. (2010). Innate immune responses of the airway epithelium. *Molecules and Cells*, *30*(3), 173–183.

Sajnaga, E., and Kazimierczak, W. (2020). Evolution and taxonomy of nematode-associated entomopathogenic bacteria of the genera *Xenorhabdus* and *Photorhabdus*: An overview. *Symbiosis*, *80*(1), 1–13.

Samarasinghe, B., Knox, D. P., and Britton, C. (2011). Factors affecting susceptibility to RNA interference in *Haemonchus contortus* and in vivo silencing of an H11 aminopeptidase gene. *International Journal for Parasitology*, *41*(1), 51–59.

Sandhu, S. K., Jagdale, G. B., Hogenhout, S. A., and Grewal, P. S. (2006). Comparative analysis of the expressed genome of the infective juvenile entomopathogenic

nematode, *Heterorhabditis bacteriophora*. *Molecular and Biochemical Parasitology*, *145*(2), 239–244.

Schroeder, N. E., and MacGuidwin, A. E. (2010). Behavioural quiescence reduces the penetration and toxicity of exogenous compounds in second-stage juveniles of *Heterodera glycines. Nematology, 12*(2), 277-287.

Seppey, M., Manni, M., and Zdobnov, E. M. (2019). BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods in Molecular Biology, 1962*, 227–245.

Serra, L., Macchietto, M., Macias-Muñoz, A., McGill, C. J., Rodriguez, I. M., Rodriguez, B., Murad, R., and Mortazavi, A. (2019). Hybrid Assembly of the Genome of the Entomopathogenic Nematode *Steinernema carpocapsae* Identifies the X-Chromosome. *G3, 9*(8), 2687–2697.

Seybold, A. C., Wharton, D. A., Thorne, M. A. S., & Marshall, C. J. (2016). Establishing RNAi in a Non-Model Organism: The Antarctic Nematode *Panagrolaimus* sp. DAW1. *PLoS One*, *11*(11), e0166228.

Shapira, M., Hamlin, B. J., Rong, J., Chen, K., Ronen, M., and Tan, M.-W. (2006). A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(38), 14086–14091.

Shapiro-Ilan, D. I., Han, R., and Dolinksi, C. (2012). Entomopathogenic Nematode Production and Application Technology. *Journal of Nematology*, *44*(2), 206–217.

Smythe, A. B., Holovachov, O., and Kocot, K. M. (2019). Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny. *BMC Evolutionary Biology*, *19*(1), 121.

Soderlund, C., Bomhoff, M., and Nelson, W. M. (2011). SyMAP v3.4: A turnkey synteny system with application to plant genomes. *Nucleic Acids Research*, *39*(10), e68–e68.

Sommer, R. J., and Bumbarger, D. J. (2012). Nematode model systems in evolution and development. *Wiley Interdisciplinary Reviews. Developmental Biology*, *1*(3), 389–400.

Sommer, R. J., and McGaughran, A. (2013). The nematode Pristionchus pacificus as a model system for integrative studies in evolutionary biology. *Molecular Ecology*, *22*(9), 2380–2393.

Somvanshi, V. S., Dubay, B., Kushwah, J., Ramamoorthy, S., Vishnu, U. S., Sankarasubramanian, J., Rajendhran, J., and Rao, U. (2019). Draft Genome Sequences for Five *Photorhabdus* Bacterial Symbionts of Entomopathogenic *Heterorhabditis* Nematodes Isolated from India. *Microbiology Resource Announcements*, *8*(4).

Somvanshi, V. S., Gahoi, S., Banakar, P., Thakur, P. K., Kumar, M., Sajnani, M., Pandey, P., and Rao, U. (2016). A transcriptomic insight into the infective juvenile stage of the insect parasitic nematode, *Heterorhabditis indica*. *BMC Genomics*, *17*(1), 166.

Somvanshi, V. S., Kaufmann-Daszczuk, B., Kim, K.-S., Mallon, S., and Ciche, T. A. (2010). *Photorhabdus* phase variants express a novel fimbrial locus, mad, essential for symbiosis. *Molecular Microbiology*, *77*(4), 1021–1038.

Somvanshi, V. S., Sloup, R. E., Crawford, J. M., Martin, A. R., Heidt, A. J., Kim, K., Clardy, J., & Ciche, T. A. (2012). A Single Promoter Inversion Switches Photorhabdus Between Pathogenic and Mutualistic States. *Science*, *337*(6090), 88–93.

Srinivasan, J., Dillman, A. R., Macchietto, M. G., Heikkinen, L., Lakso, M., Fracchia, K. M., Antoshechkin, I., Mortazavi, A., Wong, G., and Sternberg, P. W. (2013). The Draft Genome and Transcriptome of *Panagrellus redivivus* Are Shaped by the Harsh Demands of a Free-Living Lifestyle. *Genetics*, *193*(4), 1279–1295.

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

Stock, S. P. (2005). Insect-parasitic nematodes: From lab curiosities to model organisms. *Journal of Invertebrate Pathology*, *89*(1), 57–66.

Stock, S. P. (2015). Diversity, Biology and Evolutionary Relationships. In R. Campos-Herrera (Ed.), *Nematode Pathogenesis of Insects and Other Pests: Ecology and Applied Technologies for Sustainable Plant and Crop Protection* (pp. 3–27).

Springer International Publishing. https://doi.org/10.1007/978-3-319-18266-7_1

Stock, S. P., and Goodrich-Blair, H. (2012). Nematode parasites, pathogens and associates of insects and invertebrates of economic importance. *Manual of Techniques in Invertebrate Pathology*, 373–426.

Stock, S. P., and Reid, A. P. (2004). Biosystematics of entomopathogenic nematodes (Steinernematidae, Heterorhabditidae): Current status and future directions. *Proceedings of the Fourth International Congress of Nematology, 8-13 June 2002, Tenerife,* 435–446.

Strauch, O., and Ehlers, R. (1998). Food signal production of *Photorhabdus luminescens* inducing the recovery of entomopathogenic nematodes *Heterorhabditis* spp. In liquid culture. *Applied Microbiology and Biotechnology*, *50*(3), 369-374.

Sumaya, N. H., Olarewaju, O., Ambaw, G., Yali, A., Godina, G., Vandenbossche, B., Barg, M., Doerfler, V., Ehlers, R.-U., and Molina, C. (2020). Gene expression analysis of oxidative stress tolerance in the entomopathogenic nematode *Heterorhabditis bacteriophora*. *Nematology*, *23*(1), 15–31.

Sun, S., Rödelsperger, C., and Sommer, R. J. (2021). Single worm transcriptomics identifies a developmental core network of oscillating genes with deep conservation across nematodes. *Genome Research*, *31*(9), 1590–1601.

Takeda, S., Kadowaki, S., Haga, T., Takaesu, H., and Mitaku, S. (2002). Identification of G protein-coupled receptor genes from the human genome sequence. *FEBS Letters*, *520*(1), 97–101.

Tan, M.-W., and Shapira, M. (2011). Genetic and molecular analysis of nematode-microbe interactions. *Cellular Microbiology*, *13*(4), 497–507.

The *C. elegans* sequencing consortium. (1998). Genome Sequence of the Nematode *C. elegans*: A Platform for Investigating Biology. *Science*, *282*(5396), 2012–2018.

Theodoropoulou, M. C., Tsaousis, G., Litou, Z., Bagos, P., and Hamodrakas, S. (2013). *GPCRpipe:* A pipeline for the detection of G-protein coupled receptors in proteomes. In *Proceedings of the 21st annual International Conference on*

Intelligent Systems for Molecular Biology (ISMB) and 12th European Conference on Computational Biology (ECCB), Berlin, Germany (pp. 21-23).

Thomas, G. M. (1979). *Xenorhabdus* gen. Nov., a Genus of Entornopathogenic, Nematophilic Bacteria of the Family Enterobacteriaceae. *International Journal of Systematic and Evolutionary Microbiology, 29(4), 352-360.*

Tobias, N. J., Heinrich, A. K., Eresmann, H., Wright, P. R., Neubacher, N., Backofen, R., and Bode, H. B. (2017). *Photorhabdus*-nematode symbiosis is dependent on hfq-mediated regulation of secondary metabolites. *Environmental Microbiology*, *19*(1), 119–129.

Tusnády, G. E., and Simon, I. (1998). Principles governing amino acid composition of integral membrane proteins: Application to topology prediction. *Journal of Molecular Biology*, *283*(2), 489–506.

Urban, M., Cuzick, A., Seager, J., Wood, V., Rutherford, K., Venkatesh, S. Y., De Silva, N., Martinez, M. C., Pedro, H., Yates, A. D., Hassani-Pak, K., and Hammond Kosack, K. E. (2020). PHI-base: The pathogen–host interactions database. *Nucleic Acids Research*, *48*(D1), D613–D620.

Vadnal, J., Ratnappan, R., Keaney, M., Kenney, E., Eleftherianos, I., O'Halloran, D., and Hawdon, J. M. (2017). Identification of candidate infection genes from the model entomopathogenic nematode *Heterorhabditis bacteriophora*. *BMC Genomics*, *18*(1), 8.

Vaishnava, S., Yamamoto, M., Severson, K. M., Ruhn, K. A., Yu, X., Koren, O., Ley, R., Wakeland, E. K., and Hooper, L. V. (2011). The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. *Science, 334*(6053), 255–258.

van der Hoeven, R., McCallum, K. C., and Garsin, D. A. (2012). Speculations on the activation of ROS generation in *C. elegans* innate immune signaling. *Worm*, *1*(3), 160–163.

Vanlalhlimpuia, L., and Hrang Chal Lalramnghaki, V. (2018). Morphological and molecular characterization of entomopathogenic nematode, *Heterorhabditis baujardi* (Rhabditida, Heterorhabditidae) from Mizoram, northeastern India. *Journal of Parasitic Diseases, 42*(3), 341–349.

Viney, M. E., and Thompson, F. J. (2008). Two hypotheses to explain why RNA interference does not work in animal parasitic nematodes. *International Journal for Parasitology*, *38*(1), 43–47.

Visvikis, O., Ihuegbu, N., Labed, S. A., Luhachack, L. G., Alves, A.-M. F., Wollenberg, A. C., Stuart, L. M., Stormo, G. D., & Irazoqui, J. E. (2014). Innate host defense requires TFEB-mediated transcription of cytoprotective and antimicrobial genes. *Immunity*, *40*(6), 896–909.

Warnock, N. D., Atcheson, E., McCoy, C., Whiteside, L., and Dalzell, J. J. (2021). Highly divergent neuropeptide – non-coding RNA regulatory networks underpin variant host-finding behaviours in *Steinernema* species infective juveniles. *International Journal for Parasitology*, *51*(9), 693–698.

Waterfield, N. R., Ciche, T., and Clarke, D. (2009). *Photorhabdus* and a Host of Hosts. *Annual Review of Microbiology*, *63*(1), 557–574.

Watson, R. J., Joyce, S. A., Spencer, G. V., and Clarke, D. J. (2005). The *exbD* gene of *Photorhabdus temperata* is required for full virulence in insects and symbiosis with the nematode *Heterorhabditis*. *Molecular Microbiology*, *56*(3), 763–773.

Watson, R. J., Millichap, P., Joyce, S. A., Reynolds, S., and Clarke, D. J. (2010). The role of iron uptake in pathogenicity and symbiosis in *Photorhabdus luminescens* TT01. *BMC Microbiology*, *10*(1), 177.

Wheeler, D., Darby, B. J., Todd, T. C., and Herman, M. A. (2012). Several grassland soil nematodes species are insensitive to RNA-mediated interference. *Journal of Nematology*, *44*, 92–101.

Wilkinson, P., Paszkiewicz, K., Moorhouse, A., Szubert, J. M., Beatson, S., Gerrard, J., Waterfield, N. R., and Ffrench-Constant, R. H. (2010). New plasmids and putative virulence factors from the draft genome of an Australian clinical isolate of *Photorhabdus asymbiotica*. *FEMS Microbiology Letters*, *309*(2), 136–143.

Wilkinson, P., Waterfield, N. R., Crossman, L., Corton, C., Sanchez-Contreras, M., Vlisidou, I., Barron, A., Bignell, A., Clark, L., Ormond, D., Mayho, M., Bason, N., Smith, F., Simmonds, M., Churcher, C., Harris, D., Thompson, N. R., Quail, M., Parkhill, J., and ffrench-Constant, R. H. (2009). Comparative genomics of

the emerging human pathogen *Photorhabdus asymbiotica* with the insect pathogen *Photorhabdus luminescens*. *BMC Genomics*, *10*, 302.

Williams, J. S., Thomas, M., and Clarke, D. J. (2005). The gene *stlA* encodes a phenylalanine ammonia-lyase that is involved in the production of a stilbene antibiotic in *Photorhabdus luminescens* TT01. *Microbiology*, *151*(8), 2543–2550.

Wistrand, M., Käll, L., and Sonnhammer, E. L. L. (2006). A general model of G protein-coupled receptor sequences and its application to detect remote homologs. *Protein Science, 15*(3), 509–521.

Yoshimura, J., Ichikawa, K., Shoura, M. J., Artiles, K. L., Gabdank, I., Wahba, L., Smith, C. L., Edgley, M. L., Rougvie, A. E., Fire, A. Z., Morishita, S., and Schwarz, E. M. (2019). Recompleting the *Caenorhabditis elegans* genome. *Genome Research*, *29*(6), 1009–1022.

Yuan, D., Li, S., Shang, Z., Wan, M., Lin, Y., Zhang, Y., Feng, Y., Xu, L., and Xiao, L. (2021). Genus-level evolutionary relationships of FAR proteins reflect the diversity of lifestyles of free-living and parasitic nematodes. *BMC Biology*, *19*(1), 178.

Yunger, E., Safra, M., Levi-Ferber, M., Haviv-Chesner, A.,and Henis-Korenblit, S. (2017). Innate immunity mediated longevity and longevity induced by germ cell removal converge on the C-type lectin domain protein IRG-7. *PLoS Genetics*, *13*(2), e1006577.

Zhang, K., Baiocchi, T., Lu, D., Chang, D. Z., and Dillman, A. R. (2019). Differentiating between scavengers and entomopathogenic nematodes: Which is *Oscheius chongmingensis*? *Journal of Invertebrate Pathology*, *167*, 107245.

Zhu, J., Chen, G., Zhu, S., Li, S., Wen, Z., Bin Li, Zheng, Y., and Shi, L. (2016). Identification of Tissue-Specific Protein-Coding and Noncoding Transcripts across 14 Human Tissues Using RNA-seq. *Scientific Reports*, *6*(1), 28400.

Zilber-Rosenberg, I., and Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: The hologenome theory of evolution. *FEMS Microbiology Reviews*, *32*(5), 723–735.

Zioni, S., Glazer, I., and Segal, D. (1992). Life cycle and reproductive potential of the nematode *Heterorhabditis bacteriophora* strain HP88. *Journal of Nematology, 24*(3), 352.

Zou, C.-G., Ma, Y.-C., Dai, L.-L., and Zhang, K.-Q. (2014). Autophagy protects *C. elegans* against necrosis during *Pseudomonas aeruginosa* infection. *Proceedings of the National Academy of Sciences*, *111*(34), 12480–12485.