Immunological Responses and Internal Microbes of *Eisenia fetida*

A Thesis submitted to the

University of North Bengal

For the Award of

Doctor of Philosophy (Ph.D.)

in

Biotechnology

by

Tilak Saha

Under the supervision of

Prof. Ranadhir Chakraborty

Department of Biotechnology
University of North Bengal
June, 2018

Dedicated to
"The Three 'M's of my life
My mother, Maya Saha,
My wife, Mousumi &
My daughter, Manyata"

- without whom neither me nor my study exist.

DECLARATION

I hereby declare that the research work embodied in this thesis has been carried out by me in the Department of Biotechnology, University of North Bengal, Darjeeling-734 013, West Bengal, India, under the supervision of Prof. (Dr.) Ranadhir Chakraborty, Department of Biotechnology, University of North Bengal, Darjeeling-734 013, West Bengal, India. I also affirm that this work is original and has not been submitted before in part or full for any degree/diploma or any other academic award to this or any other University or Institution.

Tilak Saha
29.05.2018

(Tilak Saha)

Department of Biotechnology

University of North Bengal -734 013

UNIVERSITY OF NORTH BENGAL

DEPARTMENT OF BIOTECHNOLOGY
DARJEELING-734 013, WEST BENGAL, INDIA

Prof. (Dr.) Ranadhir Chakraborty

CERTIFICATE

The research work presented in this thesis entitled "**Immunological Responses and Internal Microbes of *Eisenia fetida***" has been carried out under my direct supervision by **Mr. Tilak Saha**. This work is original and has not been submitted for any degree or diploma to this or any other University or Institution.

Date: 29.5.2018

(Ranadhir Chakraborty)
Supervisor
Ranadhir Chakraborty, Ph.D
Professor, Biotechnology
University of North Bengal
Siliguri-734013, West Bengal

ANTI-PLAGIARISM REPORT

URKUND

Urkund Analysis Result

Analysed Document:	tilaksaha_biotechnology.pdf (D39985758)
Submitted:	6/7/2018 1:42:00 PM
Submitted By:	nbuplg@gmail.com
Significance:	3 %

Sources included in the report:

http://www.biotech-asia.org/vol6no1/procoagulant-and-hemolytic-activities-of-coelomic-fluid-of-earthworms-eisenia-foetida-and-perionyx-excavatus/
http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0079257
https://www.ncbi.nlm.nih.gov/pubmed/9733802
https://is.cuni.cz/webapps/zzp/download/120036896
https://www.ijcmas.com/vol-3-3/J.%20Marlin%20Cynthia,%20et%20al.pdf
https://digital.library.unt.edu/ark:/67531/metadc2602/m2/1/high_res_d/thesis.pdf
http://www.mdpi.com/2072-6651/5/8/1392/htm
https://link.springer.com/chapter/10.1007/978-1-4419-8059-5_4
https://medworm.com/536564637/chryseomicrobium-excrementi-sp-nov-a-gram-stain-positive-rod-shaped-bacterium-isolated-from-an-ea/
https://link.springer.com/article/10.1007%252FBF00303095
https://link.springer.com/article/10.1007/BF02708433
https://link.springer.com/article/10.1007/s11368-011-0437-1
https://pdfs.semanticscholar.org/f618/979a9717b1730975643807 55f0836495367b.pdf
https://www.ncbi.nlm.nih.gov/pubmed/29749923
https://www.pubfacts.com/search/Draft+genome+sequence
https://worldwidescience.org/topicpages/e/earthworm+eisenia+andrei.html
https://www.science.gov/topicpages/e/earthworm+eisenia+fetida
https://www.ijcmas.com/6-8-2017/Jyoti%20Yadav,%20et%20al.pdf
https://www.ias.ac.in/article/fulltext/jbsc/028/06/0723-0731
http://geenmedical.com/search?wd=Chakraborty+R[Author]
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4960643/

Instances where selected sources appear:

57

SCHOLAR

Tilak Saha
07.06.2018

SUPERVISOR

7.6.2018

Ranadhir Chakraborty, Ph.D
Professor, Biotechnology
University of North Bengal
Siliguri-734013, West Bengal

ABSTRACT

Eisenia fetida, which crawls through and characteristically consumes its own home/habitat, is dependent on microbial associations for its growth and reproduction. It devours microbe-rich compost and animal manures, thus constantly exposing itself to pathogens. It hosts a good number of bacteria in its gut and coelom from its habitat. The gut contains more (many times) live bacteria than coelom. A certain fraction of the innumerable diverse bacteria present in the earthworm feed, during its passage through the gut, enjoys a selective advantage in the gut environment to contribute transformation of the soil biogeochemistry. As members of *Firmicutes* constituted the major fraction of the cultivable bacterial diversity, we focused our study on *Eisenia-Firmicute* association more specifically *E. fetida-Bacillus* spp. association. In the present study, the interrelationship between habitat (processed cow dung) and the earthworm (*E. fetida*) has been studied in the laboratory mimicking the *in-situ* conditions. The undigested residue of consumed food material excreted by the herbivorous cows, defined as cow dung, contains high titres of culturable *Bacillus* spp. ($> 10^{13}$ g^{-1} processed cow dung) throughout the period of processing until it gets suitable for feeding the earthworms. Since several *Bacillus* species present in the gut of *E. fetida* play a major role in the degradation of polymeric materials, they are predominant with titres $> 10^{11}$ g^{-1} gut content. In this study, twenty bacterial strains from the different regions of gut of *E. fetida* were isolated followed by molecular characterization by 16S rRNA genes. Most of the strains appear to be novel at the species level and at least one at the genus level. Further biochemical tests and chemotaxonomic studies were carried out for the bacterial strains ET03^T, EPG1^T, EAG2^T and EAG3^T to confirm the molecular findings.

Community composition, functional and metabolic dynamics were assessed by metagenomic analyses. Whole metagenome sequences were derived from experiments carried out on Illumina MiSeq platform. Taxonomic hit distribution at phylum level shows that the metagenome has 28.5% *Proteobacteria*, 15.2 % *Firmicute*s, 13.1% *Actinobacteria* and 13% *Bacteroidetes*. This finding has validated the culture-dependent data; where we found *Firmicutes* (~40%), *Proteobacteria* (~30%) and *Actinobacteria* (~30%); data on *Bacteroidetes*, being anaerobic, could not be found in the study (culture-dependent)as it was limited to mainly the aerobic and

microaerophilic cultivable bacteria. Whole Genome Sequences (WGS) of the unique *Eisenia*-associated-bacterial strains, ET03^T, EPG1^T, EAG2^T and EAG3^T, have been derived from Illumina NextSeq 500 NGS platform.

Coelome microbiome of *E. fetida* was explored by both culture-dependent and independent methods. Coupled with this aspect was the study related to the function of the host coelomocytes. For the culture-dependent exploration of bacterial diversity, we aseptically harvested the coelomic fluid (CF) in sterile capillary tubes. Serially diluted CF was spread on Luria Agar (LA) plates for heterotrophic bacterial growth and Hichrome Bacillus Agar (HBA) plates for enumeration of *Bacillus* specific population load in CF. Colonies from HBA plates were picked, purified and identified by means of 16S rRNA gene phylogeny. To estimate culture independent bacterial diversity, CF was used as the template for PCR amplification of 16S rRNA genes. The amplicons were cloned and randomly selected for restriction digestion using *Hae* III endonuclease for RFLP analysis. A UPGMA tree was generated from the restriction map to reveal the diversity of 16S rRNA gene sequences. The clusters obtained from UPGMA tree were cross-validated with 16S rRNA sequences derived from pure cultures resulting from culture-dependent experiments.

Coelomic space of the composter earthworm *E. fetida* is not sterile. Three *Bacillus* spp. viz. *Bacillus megaterium, B. cereus, B. pumilus* have been prominently observed at varied concentrations in the coelomic fluid of *Eisenia* during various stages of vermicomposting at regulated laboratory conditions. *B. coagulans* remained in the processed raw cow dung (PrCD) as the most preferred group (10^7-10^8 cfu/g PrCD) but was absent in the coelomic space of *E. fetida*. *B. megaterium* is invariably found in coelomic fluid of *Eisenia* (10^4 cfu/ml) along with variable presence of other co-species. *Bacillus thuringiensis,* which is an opportunistic invertebrate pathogen, occasionally observed in the processed raw cow dung (PrCD) and consequently in the coelomic fluid. High dose of *B. thuringiensis* inoculation in the PrCD leads to infection of coelomic fluid with the same. *E. fetida* has evolved various immune-defense mechanisms which are assigned to coelomocytes, present in it's coelomic fluid. Light microscopic studies revealed the following morphologically distinct groups of coelomocytes in *E. fetida* - amoebocytes (23±9%), granulocytes(18±7%) eleocytes(6±3%). Amoebocytes with phagocytotic activity occasionally bear inclusion bodies. Coelomic fluid contains 6×10^5/ml naturally occurring bacteria. The number

of potential phagocytic cells is >10 higher than the microbial population. Phagocytosis by coelomocytes can be modulated by humoral components and thus phagocytosis is promoted at certain states to check the pathogen spp. Forceful introduction of *B. thuringiensis* in the coelomic space of *E. fetida* increases the rate of phagocytosis.

E. fetida, also, has gained importance as a model in regeneration biology. Detailed histological shreds of evidence of tissue-level dynamics during regeneration are not available. The present study has been undertaken to describe tissue reorganization after amputation in *E. fetida*. Transverse amputation of adult *Eisenia fetida* at different regions of the body followed by survival and development studies revealed that anterior fragments can regenerate missing posterior regions when amputations are done at least beyond the clitella. Internal tissue reorganization and formation of the major tissues are complete during posterior regeneration within 11^{th} day post amputation. Histological studies reveal that neoblast cells originate from de-differentiation of the longitudinal muscle cells, basal epithelial cells and cells of visceral peritoneum. The blastemal mass comprising chloragogue tissue has also been observed. Histological study of the anterior amputees revealed gross tissue orientation deformities. The absence of de-differentiated muscle cells and no growth in visceral peritoneum are distinctly observed; blastema did not form. The population of *B. cereus* and *B. megaterium* is significantly reduced and the number of eleocytes was significantly increased during the middle stage of posterior regeneration. Finally differential gene expression in coelomocytes under bacterial challenge in comparison to control was analysed via genome-wide transcriptomics.

———

PREFACE

Since the dawn of immunology research and in spite of the dominant spotlight on mammalian and in particular human immunology, it may appear startling that invertebrates—and among others earthworms have continued to be an important model to solve several mysteries. In the sixties when transplantation experiments paved the path of understanding the recognition of self/non-self, it has also set the stone rolling for next decades to intensify research on earthworm immune mechanisms that evolved to prevent the invasions of pathogens. The skin constituted of epidermis and cuticle drained with muco-polysaccharides, casing the earthworm's body, is considered to be the first nonspecific antimicrobial barrier. The epidermis is outlined by a mono-layer epithelium of supporting cells, basal cells and secretor cells. The basal cells play a central role in wound healing and graft rejection, often wielding phagocytic activity. These phagocytes are coelomocytes which are freely flowing cells in the coelomic cavity. Recently, these coelomocytes have received particular attention in order to study immunity processes. The communication with the external environment is maintained by every segment of the coelomic cavity and so the skin fails to fully prevent the microbes from entering the coelomic cavity. As a result, the coelomic cavity is not aseptic and always contains microorganisms from the outer environment. Yet, there are well-organized mechanisms that keep the growth of microorganisms under control.

Moreover, earthworms play a key role in soil biology by providing ideal conditions for the growth of microorganisms. Biomass of soil, in many ecosystems, constituted of invertebrate components is largely dominated by earthworms. In the present study, the interrelationship between habitat (processed cow dung) and the earthworm (*Eisenia fetida*) has been studied in the laboratory mimicking the *in-situ* conditions. *In situ* environmental conditions versus those of the animal gut present distinct habitats and challenges for microbes due to differences in water activity regimens, pH, oxygen levels, etc. Cow-dung was used as a model system to study survivability of the randomly ingested microbes (bacteria, Achaea, fungi, and protozoa) during passage through the gut of the dung feeder earthworm *Eisenia fetida*. The cow-dung contains high titres of cultivable *Bacillus* spp. (> 10^{13} g^{-1} processed

cow dung) throughout the period of processing until it gets suitable for feeding the earthworms. Microbes of aerated cow-dung, after being devoured by the earthworm, passes through an itinerant anoxygenic micro-zone through suffering assaults or being selectively favoured by the gut *in-situ* factors. Since several *Bacillu*s species are predominant with titres > 10^{11} g^{-1} gut content, it was hypothesized that the journey of *Bacillus* population from cow dung to earthworm gut and back to the environment has an important regulatory effect on the overall dynamics of *Bacillus* population. The thesis has elucidated in details the swings in population dynamics of the subset of ingested dung-microbes under unique micro-conditions of the earthworm gut. Knowing fully well the limitation of culture-dependent studies, whole metagenome sequence of *E. fetida* gut content was analysed to understand the spectra of microbial diversity prevailing within earthworm system. Attempting to answer another fundamental question-'How do these worms manage to regulate the bacterial population under control?' was also one of the other major objectives set in this thesis. Side-by-side, wound healing including regeneration (after amputation) with concurrent monitoring of innate immune status and indigenous micro-flora of the coelomic fluid have been studied. Finally differential gene expression in coelomocytes under bacterial challenge was analysed via genome-wide transcriptomics.

———

ACKNOWLEDGEMENT

I am indebted to a great many people for their help and support during my Ph.D work.

First and foremost, I register my deepest gratitude and thanks to my supervisor ***Prof. Ranadhir Chakraborty,*** *Department of Biotechnology, University of North Bengal, Siliguri, West Bengal, India. Not only is he a valuable information source, but he is also an excellent personality. It has been my pleasure to work under someone who truly loves what he does. I consider myself fortunate enough to be associated with him whose vast scientific knowledge, personal care and spirit has resulted in fulfillment of this thesis.*

I register my gratefulness to Prof. Ashim Kumar Chakravarty, who was always there to help and encourage me, during my days of hardship. If he hadn't taken me to my mentor at the right time, I may never be where I am today.

I wish to express my deep sense of gratitude and respect to Dr. Dipanwita Saha, Dr. Shilpi Ghosh and Dr. Anoop Kumar for their help, constant encouragement and blessings.

I record my special thanks to all my senior research scholars and friends of my laboratory: Dr. Bhaskar Bhadra, Dr. Shriparna Mukherjee, Dr. Suparna Saha Bhowal, Dr. Arvind Kumar, Dr. Krishna Kant Yadav, Dr. Amit Kumar Mandal, Dr. Bipransh Kumar Tiwary for their support and help. My labmates Vivek Kumar Ranjan, Suvonil Chakraborty, Sudipta Das, Subhajit Sen and lab attendant Mr. Arup Ghosh deserve my heartiest thanks for their kind help and co-operation during the research work. I should also thank Post M.Sc. and M.Sc. (desertation) students-Nazli, Amit, Dibyendu, Nibendu, Avishek , Partha, Arun for their help and support. Special thanks go to the co-worker of my initial days in the lab, Anup Biswas.

I wish to thank my teachers, colleagues and other staff of the Department of Zoology, University of North Bengal for their co-operation during my research work. I would also like to take this opportunity to thank Sri. Biswanath Chakraborty, Central Instruments Laboratory, University of Kalyani, West Bengal, India for helping me in Scanning Electron Microscopy. I wish to thank the authority of the Universitity of North Bengal for providing me with the required facilities and support.

I can't express my profound regards in words only to my parents Sri Surendranath Saha and Smt. Maya Saha; my father-in-law Sri Nikhil Sarkar and mother-in-law Smt. Tanushree Sarkar for their support and unconditional love overlooking my shortcomings. I have constantly received significant support from my elder brother Pulak saha and my brother-in-law Sri Purnendu Das and Sri Tanmoy Sarkar. My wife, Mousumi Saha, has been my constant source of encouragement during this journey. She had shouldered my responsibilities, enabling me to perform my experiments in the laboratory till late night. It is my sweet daughter, Manyata who deserves my best complements and blessings for letting me work without boredom. Truly, I believe that completion of my Ph. D. work is the outcome of the blessings and love of all my family members.

Lastly, I would like to thank everybody who was important for the successful realization of this thesis; I express my apology that I could not mention their names personally.

Tilak Saha

Date: 29.05.2018
Place: NBU

(Tilak Saha)

CONTENTS

PAGE NO.

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AAF	Assembly and alignment-free
ANI	Average nucleotide identity
ARDRA	Amplified ribosomal DNA restriction analysis
°C	Degree Celsius
cDNA	Cyclic deoxyribo nucleic acid
CI	Competitive index
CMB	Coelomic fluid Mimicking Broth
CCF	Coelomic cytolytic factor
COFAM	Centre for Floriculture and Agro-business Management
COG	Clusters of Orthologous Groups
CPCSEA	Committee for the Purpose of Control And Supervision of Experiments on Animals
Da	Dalton
DMSO	Dimethyl sulfoxide
DPG	Diphosphatidylglycerol
g	Gram
GGDC	Genome-to-genome direct comparison
GO	Gene ontology
h	Hour(s)
HBA	Hichrome bacillus agar
IAEC	Institutional Animal Ethics Committee
Kg	Kilogram
kDa	Kilo Dalton
KO	KEGG (Kyoto Encyclopedia of Genes and Genomes) orthology
LD_{50}	Lethal dose for 50% of treated individuals
M	Molar
µg	Microgram
µl	Microliter
µm	Micrometer
µM	Micromolar
Mg	Milligram
MK	Menaquinone
m	Minute
m-DAP	Mmeso-diaminopimelic acid
ml	Mililiter
Mm	Millimeter
NGS	Next generation sequencing
NCBI-NR	National Center for Biotechnology Information non-redundant database
NOG	Non-supervised Orthologous Groups
mM	Millimolar
PBS	Phosphate buffered saline
PE	Phosphatidyl ethanolamine
PIM	Phosphatidyl inositolmannoside
pRCD	Processed cow dung
RCD	Raw cow dung
rRNA	Ribosomal ribo nucleic acid
Rpm	Rotation per minute
SDS-PAGE	Sodium dodecyl sulfate poly-acrylamide gel electrophoresis
SEM	Scanning electron microscope
TLC	Thin layer chromatography
UV	Ultraviolet
WGS	Whole genome sequencing

LITERATURE REVIEW

Literature review

I. Introduction to *Eisenia fetida*

Earthworms constitute one of the major invertebrate biomass living in the soil. Among the different species of these oligochaete worms, *Eisenia fetida* (Savigny) is the most renowned one for its ability to efficiently transform organic matters like peels of vegetables, fruits and leaves of plants etc. into vermicompost, the organic manure rich in nitrogen, phosphate and other essential components useful for proper plant growth. This cosmopolitan red worm is also naturally found in the upper part of the soil of mountainous regions and the adjoining terrains of North Bengal where cultivation of famous 'Darjeeling tea' is exercised (Halder, 1999).

The individuals belonging to the genus *Eisenia*, Malm, 1877 (Subfamily- Eiseninae, Family- Lumbricidae, Order- Haplotaxida, Class- Clitellata, Phylum- Annelida) display a reddish colour, epilobous prostomium, closely paired setae, first dorsal pore around annule number 4/5, calciferous gland without pouches in 10, (Bouche, 1972; Gates, 1975; Csuzdi and Zicsi, 2003). There was some heterogeneity and confusion in the somatic characters in the species included in this genus. Many researchers believed that *E. fetida* and *E andrei* were phenotypic variants. Andre (1963) made them conspecific by naming the uniformly pigmented form var. *unicolor* and the striped form var. *typica*. Considering the unavailability of natural hybrids Bouche (1972) gave the uniformly pigmented worms the new subspecific name *Eisenia fetida andrei*. Apart from the difference in pigmentation the two forms cannot be distinguished. Both forms have a mean length of 60-120 mm, a diameter of 3-6 mm and a segment number varying between 80 and 120. The saddle-shaped clitellum covers 6-8 segments. Sims (1983) postulated the uniformly pigmented forms to be derived from the striped forms. Sheppard (1988) recommended two species of *Eisenia* distinguishable by pigmentation. He considered the banded worms to be *E. fetida* (brandling or tiger worms) while uniformly pigmented worms as *E. andrei* (red worm). Sheppard (1988) found that while both forms had similar cocoons laying ability and cocoon viability, but the numbers of progeny per cocoon produced by the two forms were significantly different. The two forms also differed biochemically (Roch *et al.*, 1980; Valembois *et al.*, 1982). Jeanike (1982) electrophoretically proved that at least three loci of the two species differed, indicating reproductive isolation. Jaenike (1982) first separated the two species genetically (Henry, 1999; McElroy and

Diehl, 2001, Perez-Losada *et al*., 2005). The genetic isolation is also correlated with physiological, ecological and behavioural differences, making it easy to distinguish between the two 'forms.

II. Cultivable bacterial diversity associated with *E. fetida*

Eisenia has been observed to continuously devour organic debris or to move ahead to search suitable food. Soil consumption for *E. fetida*, an earthworm is estimated to be 16 mg soil/ individual/day (Shaheen *et al.*, 2010). The nitrogen excretion rate of *E. fetida* has been estimated at 0.4 mg/g/day, which is very high relative to other earthworm species (Stafford and Edwards, 1985). Aerobic intestinal bacterial community structure of earthworm, *E. fetida*, was investigated by Kim *et al*. in 2004 based on 16S rRNA gene analysis. Ninety-one different colonies grown on Brain Heart Infusion medium were randomly isolated under aerobic condition. Based on partial sequence analysis of PCR-amplified 16S rRNA gene for strains, earthworm intestinal aerobic bacteria (EIAB) were divided into 12 groups, and each group was further divided into subgroups. Groups included 6% *Aeromonas*, 3% Agromyces, 31% *Bacillus*, 1% *Bosea*, 6% *Gordonia*, 6% *Klebsiella*, 7% *Microbacterium*, 2% *Nocardia*, 10% *Pseudomonas*, 19% *Rhodococcus*, 2% *Tsukamurella*, and 7% *Streptomyces*, with *Bacillus* being dominant group.

Kim *et al*. in 2010 identified *Brevibacillus agri*, *Bacillus cereus*, *Bacillus licheniformis*, and *Brevibacillus parabrevis* from earthworm viscera by 16S rRNA sequencing. These bacteria may be employed in the conversion of fish wastes into fertilizer. In 2013 Zhang *et al*. investigated change of intestinal bacteria community of *Eisenia* on *E coli* O157:H7 challenge by PCR-DGGE analysis. The result demonstrated that the intestinal bacteria of earthworm had the ability to adjust community structure to eliminate the pathogen *E. coli* within three days, and the amount of bacteria *Bacillus* increased significantly, which might be the positive antagonism to *E. coli*. Kwang-Hee Shin *et al.* in 2004 isolated one hundred different anaerobic bacterial colonies grown on Brain Heart Infusion medium. With the aid of partial sequence analysis of PCR-amplified 16S rDNA, these earthworm intestinal bacteria were identified to be *Clostridium bifermentans*, *C. butyricum*, *C. glycolicum*, *C. celerecrescens*, *C. lituseburense*, *Staphylococcus epidermidis*, *Propionibacterium acnes* (97% or more similarity) and a dominant group (49% of all isolates) of unique pyretic line with 90-95% similarity to *C. subterminale*.

Go´mez-Brando´n *et al*., 2011 found that some specific bacteria are activated and selected than others during passage through the gut. Monroy *et al.* (2011) observed a reduction in the density of total coliforms by 98%, after the passage of pig slurry through the gut of *E. fetida*. This was not related to decreases in bacterial biomass indicating a selection of bacterial group. Bacteria like *Escherichia coli* BJ18 in cattle dung is selectively reduced during passage through the gut of earthworms of the genus *Lumbricus*. Such reduction in the coliform bacteria did not influence other bacterial groups although a shift in the composition of the microbial community is observed. This selective effect on microbial composition of the ingested material through the earthworm gut may be because of competitive interactions between the transiently ingested and the endosymbiotic microbes residing in the gut (Kim *et al.* in 2010).

Phylogenetic distribution of 16S rRNA cDNA sequences obtained from [^{12}C]- and [^{13}C] glucose treatment as Phylogenetic affiliation (total relative abundance of phylum) revealed that the major phyla of earthworm gut contents included *Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, Firmicutes, Gemmatimonadetes, Nitrospirae, Planctomycetes,* (Alpha-, Beta-, Gamma- and Delta-) *Proteobacteria, Tenericutes and Verrucomicrobia*, taxa common to soils identified as revealed by 16S rRNA analysis. Alpha-, Beta- and Gamma-*Proteobacteria, Bacteroidetes, Actinobacteria and Firmicutes* were potentially dominant taxa in the earthworm gut, as determined on the basis of 16S rRNA gene analyses (Furlong *et al.,* 2002; Singleton *et al.,* 2003; Knapp *et al.,* 2009). Several authors have reported bacterial species belonging to the familis *Actinobacteriacae Aeromonadaceae, Comamonadaceae, Enterobacteriaceae, Flavobacteriaceae, Moraxellaceae, 'Paenibacillaceae', Pseudomonadaceae, Rhodocyclaceae, Sphingobacteriaceae* etc. to be present in the alimentary canal of earthworm (Ihssen *et al*., 2003; Horn *et al*., 2005; Byzov *et al*., 2009; Knapp *et al*., 2009). Bacteria of such diverse taxa which if metabolically active in the earthworm gut, might participate in hydrolysis of ingested biopolymers like cellulose, lignin etc. (Martin-Carnahan and Joseph, 2005; Bernadet and Bowman, 2006; Priest, 2009).

III. Culture independent studies of microbial population

Gut bacterial communities of *E. fetida* and *Perionyx excavates* grown on lignocellulosic biomass, was analysed by Singh *et al.,* 2015 16S rRNA gene based

clonal survey of gut metagenomic DNA. Total 67 clonal sequences belonging to *E. fetida* and 75 to *P. excavatus* were taxonomically annotated using MG-RAST. Most of the sequences were annotated to *Proteobacteria* (38-44%), and Firmicutes (9-11%) with a substantial presence of unclassified bacteria (14-18%). *P. excavatus* also had higher abundance of *Actinobacteria* along with *Firmicutes*.

Microbial composition can truly be studied by sequencing the microbiome with modern approaches such as-16S rRNA amplicon sequencing, shotgun DNA sequencing and shotgun RNA sequencing. But, anaerobic microbiome studied by high throughput sequencing strategies had certain biases (Campanaro *et al.,* 2018). This biasness in the abundance of some taxa (*Euryarchaeota* and *Spirochaetes*) was because of the inefficiency of universal primers to hybridize all the templates. Number of hypervariable regions under investigation influences the reliability of the results obtained. Shotgun DNA and 16S rRNA gene amplicon sequencing may underestimate the *Methanoculleus* genus, probably due to the low 16S rRNA gene copy number encoded in this taxon.

Rosselli 2016 commented that DNA-based methods may not provide information on the actual physiological relevance of each taxon within an environment and are affected by the variable number of rRNA operons in different genomes. To overcome these drawbacks an approach was proposed for direct sequencing of 16S ribosomal RNA without any primer or PCR-dependent step. So, the analyses of environmental microbial communities by direct 16S rRNA-seq from bacterial communities were attempted. The method was tested on a microbial community developing in an anammox bioreactor sampled at different time-points.

Development of metagenomics, although only a decade old, has made it possible to investigate microbes in their natural environments

IV. *Eisenia fetida* and microbe interaction

Gut wall-associated bacterial communities of *Lumbricus terrestris* and *L. friendi* (endogeic) were detected by Thakuria *et al*. 2010 using automated ribosomal intergenic spacer analysis of 16S and 23S rRNA genes. Prevalence of specific gut wall-associated bacteria, belonging mainly to the phyla *Proteobacteria, Firmicutes* and *Actinobacteria*, was ecological group dependent. Five different types of food

have been shown to cause shifts in gut wall-associated bacterial community, but the ecological group were same.

Aira *et al.,* 2015 studied effects of earthworms (after passage through its gut) on the taxonomic and phylogenetic bacterial composition of animal manures. They described cast microbiome of earthworm. When *E. andrei* was fed with the cow, horse and pig manures the emitted cast strongly differed in their taxonomic composition, but these differences were markedly reduced once transformed into earthworm cast microbiomes after passage through the earthworm gut. They reported 30 OTUs (2.6% of OTUs from cast samples) to be the core earthworm cast microbiome comprised, of which 10 are possibly native to the earthworm gut. Cast microbiome of earthworm is constituted mainly by the members of Phyla *Actinobacteria* and *Proteobacteria* where as for most other animals core gut microbiomes are composed mainly of Firmicutes and Bacteroidetes. So it seems that earthworms control the microbial composition of cast by selecting from the bacterial pool present in the ingested material.

V. Brief introduction to coelomic fluid and coelomocytes

Cotuk and Dales (1984) examined that the earthworm *E. fetida* has both cellular and humoral defences against bacteria. Later, R. P. Dales and Y. Kalac, 1991 explained that the antibacterial property of coelomic fluid of *E. fetida* is because of three host microbes. *Aeromonos hydrophila, Serratia marcestens, Yersinia ruckeri* are those three microbes which occur naturally in the coelomic fluid of *E. fetida*. Among them, *A. hydrophila* and *S. marcestens*are were known pathogen, and *Y. ruckeri* was later on found to be pathogenic to another earthworm, *Dendrobaena ueneta* (Cotuk and Kala, 1990). Antibacterial substances present in the coelomic fluid of *Eisenia* have been postulated to inhibit growth of some bacteria. Infection of the earthworm leads to enhancement of these substances in the coelomic fluid (Lassegues *et al.,* 1989). Valembois *et al.* (1982, 1986) demonstrated that lipoprotein in the coelomic fluid has antibacterial properties against bacteria which share at least one common antigen with sheep red blood cells, which imparts coelomic fluid a property to lyses RBC. Along with this, coelomic fluid of earthworms *E. fetida exerts* a large variety of biological effects including bacteriostatic, proteolytic, and cytolytic activities (Bilej *et al.,* 1991) Kauschke and Mohrig (1987) described the toxic effect of *E. fetida* coelomic fluid on different cell types, such as chicken fibroblasts, guinea-pig polymorphonuclear leukocytes, and insect haemocytes. They also found that this toxic effect was not

exerted against coelomocytes of other lumbricids as well as the cells of some molluscs, nematodes, and protozoans. The toxic effect seems to be correlated with the haemolytic activity, since 3 out of 7 haemolytic fractions exerted cytotoxic activity. Furthermore, compounds with antitumor activities have been isolated from body homogenates of lumbricids *E. fetida, Lumbricus rubelus,* and *Lumbricus terrestris.* Bilej *et al.,* (1990) examined that the coelomocytes have antigen binding property with akinetic approach., Hanus'ova *et al.,* (1998) suggested that coelomic fluid contains a ubiquitous PLA2-like component which might be involved in inflammatory reactions in earthworms (*E. fetida*). They reported the isolation of a factor from the coelomic fluid of the earthworm *E. fetida* that exerts mitogenic activity on murine splenocytes. This factor, CMF (coelomic mitogenic factor) was found to bind concanavalin A (ConA) and to block ConA-induced spleen cell proliferation. In contrast CMF synergizes with lipopolysaccharide (LPS) to trigger the proliferation of the same cell type. N-terminal amino acid sequencing revealed that CMF displays significant homology with phospholipase A2 (PLA2). CMF-enriched coelomic fluid fraction exerts phospholipase enzymatic activity.

Different authors have classified coelomocytes into types on the basis of morphological, cytochemical criteria and functional attributes (antibacterial property). In general, there are three main types — eleocytes, the free chloragogen cells with functions in nutrition, the other two types are either hyaline or granular amoebocytes, representing effector immunocytes. The amoebocytes seem to be involved in a broad range of immunological functions like phagocytosis. Amoebocytes can engulf materials like inert carbon particles, microbial cells and foreign cells. (Bilej *et al.*,1990). Homa *et al*., 2012 described five types of coelomocytes- leucocytes type I (basophilic) and II (acidophilic), neutrophils, granulocytes and eleocytes. All types of coelomocytes except eleocytes can phagocytose and encapsulate foreign materials.

Kurek *et al.,* (2007) put some light on the origin of amoebocytes from the coelomic mesenchymal lining and eleocytes (chloragocytes) from chloragogen cells that mainly cover the alimentary tract on coelomic side. Amoebocytes recognise and eliminate foreign material, mainly by phagocytosis and encapsulation. They are also involved in clotting, wound healing, cytotoxicity, inflammation, graft rejection, granuloma formation and coelomic fluid coagulation. Eleocytes, also called chloragocytes have numerous spherical granules, the chloragosomes. They have the capacity to store

endogenous materials such as glycogen and lipids as well as pigments including riboflavin. Eleocytes take part in ion and pH balance of the coelomic fluid, in some aspects of nutrition and excretion, but are also associated with immune defence. They are involved in encapsulation and brown body formation. Secretions of eleocytes have bacteriostatic properties. Moreover they take part in detoxification of earthworm tissues and heavy metals accumulation. Different types of stress (i.e. heat, pollution) lead to heat shock proteins expression in coelomocytes.

VI. Modulation of coelomic fluid constituents under bacterial challenge

Coelomic cavity of *E. fetida* communicates with outer environment by dorsal pores. In consequence, the coelomic cavity is not aseptic and always contains bacteria, protozoans and fungi from the outer environment. But, there are efficient mechanisms by which the growth of microorganisms are kept under control (Dales *et al.*,1992). It was reported that coelomic fluid contains naturally occurring bacteria in the magnitude of $6X10^5$ /ml ($0.9X10^5$ /adult individual) while the number of potentially phagocytic cells is more than ten times higher (Ko¨hlerova´, 2004). If the bacterial burden surpasses the optimum limit it may cause disease in earthworm. Smirnoff and Heimpel (1961) reported that when large doses of *Bacillus thuringiensis* invaded the body cavity of the earthworm, it caused an extensive septicemia and eventual death. Hiempel (1966) reported that blister disease of *E. fetida* was found to contain crystalliferous bacteria in all the lesions. Two strains of bacteria were isolated, identified, and typed according to their H-antigen and their vegetative cell esterases. Both strains appeared to be *Bacillus thuringiensis*. Superior number of phagocytes along with various humoral factors prevents the microorganisms from outgrowth (Bilej *et al*., 2000). These phagocytes are coelomocytes which are freely flowing cells in the coelomic cavity. Recently, these coelomocytes have received particular attention (Goven *et al.*, 1987; Rodriguez-Grau *et al.*, 1989; Fitzpatrick *et al.*, 1990). To study immunological processes, techniques for isolation of coelomocytes by dissection, electrical stimulation or withdrawal by fine capillary tubes have been developed. In 1973, Valembois *et al.* have described mechanical and electrical excitation techniques to induce, extrusion of coelomocytes through pores in *E. fetida* integument. A year later, Cooper (1974) proposed the recovery of *Lumbricus terrestris* coelomocytes by puncture through the integument with a sharpened pasteur pipette. A non-invasive extrusion technique for the harvesting of *Lumbricus terrestris*

coelomocytes was proposed by Rodriguez-Grau *et al*. (1989), a technique fully described by Eyambe *et al.* (1991). Its novelty was in part the use of ethanol as an irritating agent and the addition of the mucolytic agent guaiacol glyceryl ether (GGE).

VII. Regeneration in *Eisenia fetida*

Annelids are a large and diverse group of typically segmented worms found in marine, freshwater, and terrestrial environments, with over 17,000 species described (Brusca and Brusca, 2003; Zhang, 2013). The main part of the body is usually composed of a series of repeated segments, ranging from just a few to several hundred (depending on species and age), with an asegmental cap of tissue present at both the anterior and posterior ends. Annelids typically have a large fluid-filled coelom surrounded by a muscular body wall, a complete gut with an anterior mouth (within the anterior asegmental cap) and a posterior anus (within the posterior cap of tissue), and a nervous system composed of paired cerebral ganglia, segmental ventral ganglia, and peripheral segmental nerve rings. Some species possess body wall outgrowths such as lateral, segmentally iterated parapodia (paddle-like outgrowths) used for locomotion; anterior tentacles, palps, or proboscises that aid in sensation, respiration and/or feeding; opercula that seal tube entrances; and lateral or posterior gills that aid in respiration. Based on their body architecture, the most common injuries are expected to involve transverse cuts of the main body (removing head and/or tail) and amputation of the various body wall outgrowths (Bely *et al*., 2014).

Most annelids add segments from a sub-terminal posterior growth zone present ahead to the non-segmental posterior cap (pygidium) to grow (Hyman, 1940; Brusca and Brusca, 2003). Some species reach a fixed number of segments after maturity, but, most others do not. Few annelids appear to grow continuously throughout their life. Under starvation conditions, degrowth in overall size occurs in at least some invertebrates like naidids and degrowth resulting in a reduction of segment number may also occur (Bely, 2014).

Annelid phylogeny was of ever uncertainty until recent molecular studies that have greatly clarified the relationships among annelid groups (Weigert *et al.*, 2014), setting the stage for interpretation of regenerative variation in a phylogenetically. Annelids which are classified into ~100 families fall into the two major clades, the *Errantia* and the *Sedentaria*. Only a few annelid lineages branch outside of these, as basal lineages

(Weigert *et al*., 2014). The *Errantia* includes many mobile species and ancestrally this clade might have pronounced anterior palps and lateral parapodia. The most common groups included in *Errantia* are nereids like, *Platynereis*, *Neanthes*, *Nereis etc*. The *Sedentaria* (meaning sedentary) includes many burrowing species and the ancestor of this clade is said to have reduced palps and parapodia because of burrowing lifestyle. Groups like *Capitellids, Spionids, Terebellids, Serpulids, Sabellids* etc. are included in *Sedentaria*. *Clitellata*, includes groups such as earthworms, aquatic "oligochaetes", and leeches. Annelids vary widely in regenerative ability. Some species can regenerate every part of the body, with extremes which can even regenerate from a single isolated segment. Some are even incapable of regenerating a single lost segment (Bely, 2006). Regeneration of anterior (towards the head) and posterior (towards the tail) segments is well documented in a wide range of annelids.

The ability to regenerate posterior segments is very broadly distributed across the phylum suggesting this ability likely to be ancestral for the phylum. Posterior regeneration is described in as many as 23 annelid families, of both the groups *Errantia* and *Sedentaria*. Only a few groups of *Sedentaria* is reported to be incapable of regenerating posterior segments which may be a secondarily lost character. Anterior regeneration is variable in annelids. Failure to regenerate anterior segments has been shown in many families of Annelida. This may be because of the fact that anterior segment regeneration ability has been lost many times within the phylum (Bely, 2006) or alternatively may have been regained in few groups.

Regeneration of structures like prostomium, feeding palps or operculum has also been shown in many groups of annelids. In species particularly having anterior regeneration capacity can also regenerate the anterior non-segmental tip i.e. prostomium and peristomium along with any head appendages. Such regeneration has been documented in few clitellates. Feeding palp regeneration has been documented in spionids (Lindsay *et al.*, 2008) and operculum regeneration in serpulids (Szabó and Ferrier, 2014). Amputation and regeneration of parapodia have been documented in one species of nereid (Boilly and Boilly-Marer, 1995).

The techniques for assessing the wound healing along with transplantation experiments have earlier been described by Cooper and Roth (1984). They observed healing under a dissecting microscope at 24 hr after wounding or grafting in lumbricid annelids. Well-healed grafts were recognized by a smooth, flat appearance,

homogeneous pigmentation and were outlined by newly formed connective tissue. In the case of well-healed wounds, a tight cicatrix developed on the wound.

Regeneration at molecular level was studied in few groups of annelids, with identification of the genes involved in the regeneration process (Bely and Sikes, 2010). Some annelids exhibit regenerative abilities very similar to planarians (Bode *et al*., 1973; Newmark *et al*., 2000), which can completely regenerate a new organism from small body fragment. However, the mechanisms of regeneration in the two groups are quite different. In planarians totipotent stem cells are widely distributed throughout their bodies that can lead the process of regeneration from every splitted fragment of the body (Redien and Alvaro, 2004). In annelids that do not have such distribution of totipotent cells throughout the body, regeneration is hypothesized to occur primarily by cellular dedifferentiation and redifferentiation (Thouveny and Tassava, 1998). To understand mechanisms of regeneration in annelids, the gene expression pattern during regeneration was studied on model animal *Enchytraeus japonensis* (Myohara *et al*., 2006). Besides the known genes ECM, glutamine synthetase, NICE-5, glucosidase, a new gene, Ejrup1-5 was observed to upregulate during regeneration. Structural analyses of the products of these genes, their functions in transportation and binding, transcriptional regulation, protein interaction and cell adhesion could be speculated. A strong expression of glutamine synthetase gene occuring in the blastemal regions of regenerating *E. japonensis* could also be observed by in situ hybridization (Niv *et al*., 2008). This suggests that at least in *E. japonensis* glutamine synthetase may play roles in regeneration. Cho *et al*., 2009 monitored the expression pattern of three *labial* genes (Pex-la-01, Pex-lab02, Pex-lab03) during regeneration in *Perionyx excavatus*. These genes were found to express only in the head-regenerating tissues. Overall, earthworm provides a unique and valuable model to investigate the mechanism of regeneration because this process is rapid.

VIII. Modulation of coelomic fluid constituents during regeneration

Wound healing in the earthworms is proposed to have use as a biomarker for assessing chemical toxicity (Cooper and Roch, 1986). The wound healing process involves the inflammatory response and various cell types, including coelomocytes. This process was monitored in earthworms like *Lumbricus terrestris* (Cikutovic *et al*., 1999) after exposure to a variety of chloride compounds. The authors found that

increasing concentration and duration of exposure significantly reduced the wound healing process (Ville *et al.*, 1995). Chloride compounds could have interfered with the membrane of coelomocytes suppressing their function in healing process. Some organics (polychlorinated biphenyl, pentachlorophenol, chlordane) suppress phagocytosis of coelomocytes in earthworm. Cell division, which is important in the wound healing process, might be affected with heavy metal (such Cd^{2+} or Cu^{2+}) pollution in soil (Cikutovic *et al.*, 1993). Such metallic ions can interfere with an enzymatic pathway in coelomocytes. Production of superoxide (O^{2-}), required for killing phagocytosed microorganisms may be inhibited. Injured animals are more susceptible to pathogens, so the chemicals that suppress the healing process may have actual effect on immunoactive cells responsible for phagocytizing and killing microorganisms.

The illumination of the regeneration process in annelid is valuable to us to explore strategies to enhance the regenerative capabilities in vertebrates as both these groups, being triploblastic eucoelomates, have similarity in basic body plan. The cellular events like muscle dedifferentiation which occur during wound healing, are similar to those that take place during tail regeneration of lizards. Additional studies in this area would enable us to model the complex mechanism of regeneration of vertebrates.

IX. Ethical permission for experiments on animal

This study was carried out following the recommendation from the Committee for the Purpose of Control And Supervision of Experiments on Animals (CPCSEA), New Delhi. The study did not use any endangered or protected species. Application for permission for animal experiments had been filed to the Institutional Animal Ethics Committee (IAEC), Department of Zoology, University of North Bengal, Siliguri 734013, West Bengal, India (Registration No. 840/ac/04/CPCSEA dt. 01/10/2004).

Eisenia fetida **e**arthworms were collected from Vermiculture Facility, Centre of Floriculture and Agribusiness Management (COFAM), Department of Biotechnology, University of North Bengal, Siliguri-734013, West Bengal, India. Detailed study plan had been given while describing of all surgical procedures (including methods of asepsis) and post-operative care.

X. Objectives of the study.

The present work had been undertaken with the following objectives-

1. To enumerate cultivable bacterial population in gut and coelomic space of *Eisenia fetida*.
2. To accomplish whole-metagenome-analyses of the gut content of *E. fetida* (by Next Generation Sequencing) leading to the evaluation of microbial diversity along with metabolic diversity with the aid of suitable bioinformatics workflow.
3. To enumerate population of different species of *Bacillus* in the coelomic fluid and to study the dynamics of indigenous *Bacillus* species in coelomic fluid in the face of challenge (by forced injection) by *Bacillus thuringiensis*.
4. To study modulation of coelomic constituents following *B. thuringiensis* challenge.
5. To study wound healing including regeneration (after amputation) with concurrent monitoring of innate immune status and indigenous microflora of the coelomic fluid.

CHAPTER - 1

CULTIVABLE BACTERIA IN THE GUT, COELOMIC SPACE AND HABITAT OF *EISENIA FETIDA* WITH SPECIAL EMPHASIS ON CELL NUMBER OF VARIOUS *BACILLUS* SPP. IN A COW DUNG-MICROBE-EARTHWORM SYSTEM

1.1 Introduction

Microbes have seemingly endless capacity to transform the world around them. All life on Earth depends directly or indirectly on microbes for many essential functions. All living organisms including plants and animals have closely associated microbial communities which make life possible for them by means of making nutrients, metals, and vitamins accessible or neutralizing toxins from the host body and its environment or even making the host resistant to other pathogenic microbes. Eukaryotic invertebrate hosts, in many occasions, function or demonstrate phenotypes which are impossible unless supported by the physiological activities of the bacteria that are allowed purposefully to reside within the system of organs. On the other hand, the eukaryotic host thrusts an effect on the dynamics of free –living microbial community (Fraune and Bosch, 2010). The interactions of the host with the environmental microbes and the ephemeral microbial communities help not only in shaping the microbial landscape but also to understand how the obligate and symbiotic microbes influence phenotypic expression of the host as well. One of the classical examples of symbiosis in invertebrates where the bacterium in return of its permanent shelter and food-security renders protection against protozoan infection and synthesizes vitamin for the host is that of *Glossina morsitans* (tsetse fly) and *Wigglesworthia glossindia* (a Gram-negative bacterium) (Heller, 2011). In *Drosophila*, the presence or absence of specific *Wolbachia* species determines resistance or susceptibility to viral infection (Faria *et al.,* 2016). The actinobacterial symbionts, *Coriobacterium glomerans* and *Gordonibacter* sp., residing in the mid-gut of the cotton stainer *Dysdercus fasciatus* contribute to strengthening the immune status of the host by supplementing B vitamins (Salem *et al.,* 2014).

Biomass of soil, in many ecosystems, constituted of invertebrate components is largely dominated by earthworms. A vertically transmitted bacterium *Verminephrobacter*, made its presence in the excretory organ (nephridia) of nearly all earthworms from the evolutionary past. In the act of give-n-take policy of symbiosis, *Verminephrobacter* spp. utilizes waste products of digestion and in return provides support to host reproduction (Lund *et al.,* 2014). A second vertically transmitted symbiotic bacterium, *Flexibacter*-like, in some of the earthworms, has also been reported (Møller *et al.,* 2015). The

characteristic body cavities, gut and coelom, of the typical tube-within-tube architecture of earthworm have enabled the researchers to investigate the immune system and its communication with the dynamic microbial population present within the cavities (Engelmann *et al.,* 2011). Earthworm hosts quite a sizable number of bacteria in its gut and coelom from its habitat (soil). The gut contains more (many times) live bacteria than coelom (Edwards and Lofty, 1977). A certain fraction of the innumerable diverse bacteria present in the earthworm feed, during its passage through the gut, enjoys the selective advantage in the gut environment to contribute transformation of the soil chemistry (Govindarajan and Prabhaharan, 2014). The bacterial density in the coelomic fluid is regulated by various factors including the presence of phagocytic cells and antimicrobials (Schindler, 2004). Since expression of defence molecules, coelomic cytolytic factor (CCF) in particular, in the gut epithelial tissue is high; immune response to an incessant flow of high-density microbial presence is less affected. The scenario of the cellular immune response in the coelomic cavity is different because of the coelomocytes, liberated from the coelome-mesenchyma, are the major players of the immune system (Dvořák *et al.,* 2016). Within the coelomic cavity the freely suspended amoebocytes, also called as immunocytes can be regarded as functionally equivalent to that of macrophages present in the vertebrates (Dhainaut and Scaps, 2001). Chloragosomes stores riboflavin (vitamin B2). Incursion by soil-borne pathogens is thwarted by eleocytes resulting from chloragocytes. When there is a possibility of huge upsurge due to the mobilization of pathogenic bacteria, pathogens in mass are captured or in other words encapsulated by multicellular entities produced by amoebocytes and eleocytes (collectively called coelomocytes). The plausible immunomodulatory function of riboflavin (stored in chloragosomes) was demonstrated by establishing it as the chemo attractant for coelomocyte-taxis (Sulik *et al.,* 2012). Eleocytes accumulates free riboflavin in chloragosome. It was hypothesized earlier that the riboflavin content is perturbed under stress or else dependent on the balance between pathogen load and immune system of the earthworm (Plytycz and Morgan, 2011**)**. Several speculations were made to reveal how the level of riboflavin is maintained in chloragocyte and eleocytes in the light of the existing knowledge that riboflavin biosynthesis does not occur in animals. It was observed that even starvation for more than a month took a heavy toll on earthworm's

body weight and reproduction but coelomocytes and riboflavin storage or accretion remained largely unaffected (Sulik *et al.*, 2012). Hence, the suspect of endosymbionts or restricted harbouring of riboflavin-synthesizing bacteria in the coelomic cavity of the earthworm becomes very strong. Even if the host's compulsion for allowing a proliferation of those preferred bacteria in the coelomic cavity to obtain the sustained supply of riboflavin is assumed or hypothesized as the prime reason for symbiosis, solid microbiological evidence are necessary to accept the hypothesis. Several pertinent issues are to be addressed in order to establish this phenomenon of purposeful hosting of soil-borne bacteria in the coelomic fluid. What attracts these environmental bacteria to direct their senses to reach the organic destiny within a biologically active niche (coelomic cavity in the present study)? What is the host's compulsion to allow the proliferation of those preferred bacteria and promote relaxation to evade immune-barricade? This study has adequately addressed these questions.

In the present study, the interrelationship between habitat (processed cow dung) and the earthworm (*Eisenia fetida*) has been studied in the laboratory mimicking the *in-situ* conditions. The undigested residue of consumed food material excreted by the herbivorous cows, defined as cow dung, contains high titres of culturable *Bacillus* spp. ($> 10^{13}$ g^{-1} processed cow dung) during processing until it gets suitably processed for feeding the earthworms. Earthworms consume their own home/habitat and their growth is dependent on microbial associations. Since several *Bacillus* species present in the gut of *E. fetida* play a major role in the degradation of polymeric materials, they are predominant with titres $> 10^{11}$ g^{-1} gut content. We hypothesized that the journey of *Bacillus* population from cow dung to earthworm gut and back to the environment has an important regulatory effect on the overall dynamics of *Bacillus* population even in case of invasion of a pathogenic species, *Bacillus thuringiensis*, and re-entry or colonization into the host system. At least three different *Bacillus* species were found in differential numbers in the circulating coelomic fluid of the host. This phenomenon has enabled us to uncover the basis of symbiosis in the coelom of *E.fetida*.

1.2 Materials and methods

1.2.1 Rearing of *Eisenia fetida* in laboratory

Fresh *Eisenia fetida* specimens were collected from Centre for Floriculture and Agro-business Management (COFAM), University of North Bengal, Darjeeling. The worms were maintained in the laboratory with dried cow dung at 22°C and 70-80% humidity in plastic tubs (Edwards and Bohlen, 1996).

1.2.1.1 Processing of cow dung for feeding E. fetida

Raw cow dung (RCD) samples from healthy cows (*Bos taurus indicus*) were collected and made into chips; sun-dried (8 h /day) for two days in the open air to transform them into semi-dried, odourless chips. The chips were broken into small pieces, moistened by spraying sterile distilled water (in order to increase the moisture content to 70%) and then left in the incubator at 25°C for 48 h. A quantity of 100 g of this semi-processed RCD was taken per sterile Petri-plate (diameter = 14 cm) to prepare the processed RCD (PrCD) earthworm bed for further experiments.

1.2.1.2 Monitoring of cultivable bacterial load and major Bacillus spp. *diversity during making of processed RCD*

Set of three semi-processed RCD containing Petri-plates were taken for this experiment. The plates were incubated at 25°C. Pinch sampling of the content of Petri-plate was followed in order to achieve a representative coverage of the cultivable bacterial load and diversity of *Bacillus* spp. Four separate aseptic recoveries, approximately 0.1 g per recovery per quadrant, were combined (quantity of mixture was determined) to make one 'sample'. Each PrCD sample, taken in a 2.0 ml microfuge tube, was dissolved (50% w/v) in phosphate buffer saline, pH 7.2 and vigorously mixed for 1.5 min using a vortex mixer (Tarson). The insoluble part of the mixture was allowed to gravitate in order to collect the PrCD aqueous solution (stock solution; undiluted). A series of dilution tubes per sample was set up to obtain dilutions of 10^{-1} through 10^{-7}. Multiple dilutions and plating on Luria agar (LA), diluted (0.1x LA) Luria agar (DLA) and Hichrome Bacillus Agar (HBA; Cat. No. M1615; Himedia, India) were used to ensure about countable plates, after incubation

in aerobic and anaerobic chambers at 25 °C, for studying the population dynamics of cultivable bacterial load and major *Bacillus* spp. diversity. Day 0 sample is defined as the sample drawn on the day immediately after placing the semi-processed PrCD in the Petri-plates. Sampling after a passage of 24 h from the time of day 0 was considered as day 1 sample. Likewise, samples at regular 24 h intervals up to day 6 were analyzed.

1.2.2 Estimation of cultivable bacterial load and major *Bacillus* spp. diversity in habitat of *E. fetida*

Two sets, three PrCD containing Petri-plates per set, were taken for this experiment. Set I was kept without earthworms. In the other set, Set II, each Petri-plate was inoculated with twenty-five adult earthworms (having clitella; body weight 0.3 ± 0.05 g). The plates were incubated at 25°C and observed at regular intervals to check escape of any earthworm from their PrCD beds. Pinch sampling of PrCD (earthworm bed) was followed in order to achieve a representative coverage of the cultivable bacterial load and diversity of *Bacillus* spp. Four separate aseptic

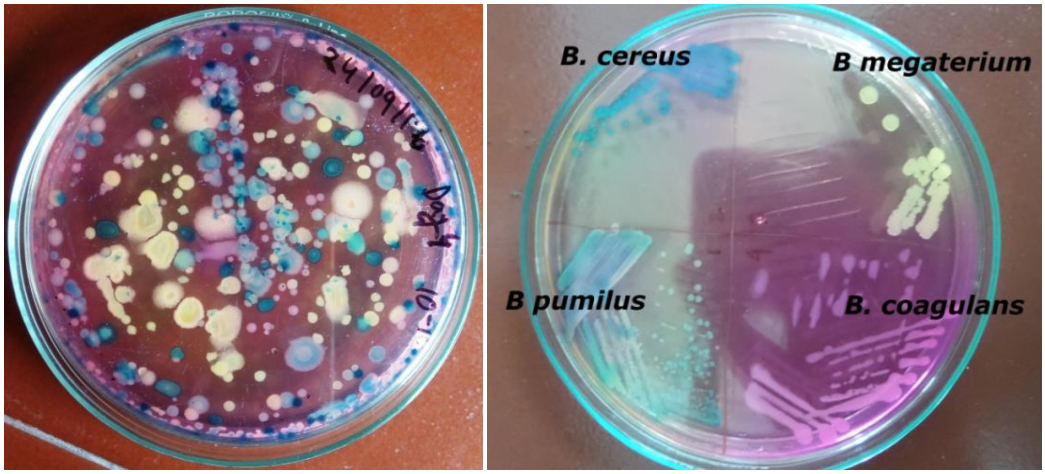


Fig. 1.1: Culture plates showing growth of major Bacillus spp. on HBA media

recoveries, approximately 0.1 g per recovery per quadrant, were combined (quantity of mixture was determined) to make one ‘sample’. Each PrCD sample, taken in a 2.0 ml microfuge tube, was dissolved (50% w/v) in phosphate buffer saline, pH 7.2 and vigorously mixed for 1.5 min using a vortex mixer. The insoluble part of the mixture was allowed to gravitate in order to collect the PrCD aqueous solution (stock solution;

undiluted). A series of dilution tubes per sample was set up to obtain dilutions of 10^{-1} through 10^{-7}. Multiple dilutions and plating on Luria agar (LA), diluted (0.1x LA) Luria agar (DLA) and Hichrome Bacillus Agar (HBA; Cat. No. M1615; Himedia, India) were used to ensure about countable plates, after incubation in aerobic and anaerobic chambers at 25 °C, for studying the population dynamics of cultivable bacterial load and major *Bacillus* spp diversity. Day 0 sample is defined as the sample drawn on the day immediately after the inoculation of *E. fetida* in PrCD bed. Sampling after a passage of 24 h from the time of day 0 was considered as day 1 sample. Likewise, samples at regular 24 h intervals up to day 6 were analyzed. Parallel samples (on the same sampling day) were drawn simultaneously from the Set I for evaluation of cultivable heterotrophic bacterial population along with the predominant *Bacillus* spp.

1.2.3 Estimation of cultivable bacterial load and major *Bacillus* spp. diversity in the gut of *E. fetida*

At different days, day 0 – 6, two earthworms (*E. fetida*) were taken from each of the three experimental plates of Set II and washed several times with sterile distilled water. Thoroughly washed earthworms (two from each PRCD plate) were transferred to sterile dry tissue paper kept inside the sterile Petri plates and observed for 20-30 min in the laminar hood. Fresh cast pellets of *E. fetida* expelled on the tissue paper were aseptically collected in pre-weighed sterile microfuge tubes. After measuring the quantity of the gut sample, dissolution of the same in PBS (50% w/v) was made. Dilution plating, spreading and incubation were done as described in the preceding section. The above-mentioned procedure was done on Day 0, 1, 2, 4 and 6.

1.2.4 Estimation of cultivable bacterial load and major *Bacillus* spp. diversity in coelomic fluid of *E. fetida*

Two earthworms (*E. fetida*), at different days (day 0 – 6), were taken from each of the three experimental plates of Set II and washed several times with sterile distilled water. Thoroughly washed individual earthworm was held firmly in a sterile tissue paper and coelomic fluid (≤ 20.0 µl) was collected in a fine- sharp sterile capillary tube (diameter ≤ 1.0 mm) by puncturing the coelomic cavity. The coelomic fluid content of the capillary

tube was serially diluted in sterile PBS. Multiple dilutions and plating on LA, DLA and HBA plates were done to enumerate the total number of heterotrophic bacteria (CFUs) and differentiation between various species of *Bacillus* along with viable count respectively.

1.2.5. Formulation of a medium -Coelomic fluid Mimic Broth (CMB)

The coelomic fluid from fifty earthworms was collected (procedure described earlier) in a sterile microfuge tube. Coelomocytes were discarded as pellet and the supernatant was collected in a fresh sterile microfuge tube after centrifugation for 10 min at 500 x g, 4 ^{0}C. Thus the CF devoid of coelomocytes was analyzed for carbohydrate, protein, fat, urea, mineral and free amino acid contents following standard methods. With reference to the observed contents of the cell-free coelomic fluid, a medium was formulated with the following composition: (in w/v.): 0.05% Dextrose, 0.77% Beef extract, 0.075% Tween 20, 0.002% Urea, 0.14% Yeast extract, 0.7% NaCl and 1.5% Agar (used only for making plates).

1.2.6. Culture-independent assessment of eubacterial diversity in the coelomic fluid of *E. fetida*

Since aseptically collected coelomic fluid, used directly as the source of target DNA from resident coelomate bacteria, repeatedly failed in PCR amplification of 16S rRNA gene sequences, enrichment procedure was followed not only to increase template numbers but also to decrease PCR inhibitory substances, present if any in the coelomic fluid. 5 ml of CMB medium was inoculated with 5 µl of aseptically collected coelomic fluid and incubated overnight at 25°C. The bacterial cell pellet was collected after centrifugation of the culture at 5000 x g for 5 min. The pellet was resuspended in 500 µl sterile distilled water, lysed by boiling at 100 ^{0}C in a water bath for 20 min. The lysate was taken as templates to PCR amplify 16S rRNA gene sequences following standard method (ref). The PCR product was gel-analyzed for specificity and yield before cloning. Blunt-end PCR products generated by proofreading DNA polymerase were directly ligated with pJET 1.2/ blunt cloning vector following protocols supplied by the manufacturer of Clone JET PCR Cloning Kit #K1231 (Thermo scientific). *E. coli* DH5α was directly

transformed with the ligation product following $CaCl_2$ transformation method. The recombinant colonies were picked up by sterile toothpicks to construct master plates for building the clone library. Each of the library members was assigned a number. A set of 14 random numbers ($1/10^{th}$ of total clones) were generated using research randomizer (http://www.randomizer.org/). Prior to clone analyses of the 14 colonies, as per numbers (random) from the previously marked colonies (1.2,3….140) on master plates, short strikes of individual colonies were propagated on ampicillin plates. Small amounts of each were used for colony PCR using pJET1.2 forward and reverse sequencing primer following protocols supplied by the manufacturer (Thermoscientific). Colony PCR products were digested with *Hae*III in accordance with the manufacturer's (Merck Genei) instruction, electrophoresed on a 2% (w/v, solution of agarose in TAE buffer) agarose gel, and band sizes were determined by using 100 bp DNA ladder as size standards. Amplified Ribosomal DNA restriction analysis (ARDRA) data from 14 clones were analyzed. The phylogenetic tree was generated by PyElph (Pavel and Vasile, 2012) (a software tool for gel images analysis) using neighbour joining method.

1.2.7. Culture-dependent isolation of *Bacillus* spp. from coelomic fluid of *E. fetida* and phylogenetic characterization of the isolates

At different time intervals, two earthworms (*E. fetida*) were taken from each of the three experimental plates, washed thoroughly several times with sterile distilled water. The coelomic fluid was collected in a sharpened capillary tube by carefully puncturing the body wall from each worm held firmly in a sterile tissue paper. The content of the capillary tube was liberated in PBS and multiple dilutions and plating on HBA plates were done to differentiate as per colour of colonies of *Bacillus* spp. The most frequently occurring colonies representing three different colours and morphology (yellow mucoid colonies for *B. megaterium*; blue flat colonies for *B. cereus*; light green to green colonies for *B. pumilis*) were isolated and dilution streaked on HBA to obtain pure cultures of the isolates. Pure cultures were maintained in Luria agar slants. Cell pellet (from 0.5 ml log-phase culture) suspended in 200 µl sterile distilled water, boiled for 1.5 minutes at 850 watts in microwave was centrifuged at 8,000 rpm for 2 min to obtain the supernatant; 2 µl of the resulting supernatant was used as template DNA for PCR amplification of 16S

rRNA gene using a set of primers as described by Lane (Lane, 1991). The PCR product was gel-analyzed for specificity and yield before cloning. The blunt-ended PCR product was ligated with pJET 1.2 vector and transformed into *E. coli* DH5α to obtain recombinant colonies, followed by sequencing of the cloned inserts in ABi3730XI by Sanger sequencing methodology.

1.2.7.1 Phylogenetic affiliations of three different strains of Bacillus, Ah4, BP, and BCR

Multiple alignments and phylogenetic analyses of the 16S rRNA gene sequences of Ah4, BP, BCR and all recognized species of genus *Bacillus* (ascertained from nBlast result against the query Ah4/BP/BCR sequence) were conducted in the software package MEGA (version 7.0) (Kumar *et al*., 2016). Multiple alignments of sequences were done with CLUSTAL W (Thompson *et al*., 1994) and the resulted multiply aligned sequences were corrected, edited and approximately 1423 base pair long nucleotide stretch of all the *Bacillus* sp. was selected for further analysis. Distances were calculated according to the Jukes-Cantor parameter. Phylogenetic analyses were performed using two tree-making algorithms: the neighbour-joining (N-J) (Saitou and Nei, 1987) and UPGMA (Unweighted Pair Group Method with Arithmetic mean) methods to ensure consistency of the clusters formed (data not shown) (Sneath and Sokal, 1973). All gaps and missing data in the aligned sequence were eliminated from the dataset (complete deletion option). Tree topology was evaluated by the bootstrap resampling method of Felsenstein (1985) based on 1000 replications.

1.2.8. Individual growth curves of *B. megaterium* (Ah4), *B. cereus* (BCR), *B. pumilus* (BP), *and B. coagulans* (BCO) [four predominant *Bacillus* species in RCD or PRCD] in CMB

Growth curves of four different strains of *Bacillu*s, Ah4, BP, BCR, and BCO in CMB medium were compared. A 100-ml flask containing 10 ml CMB medium was inoculated with Ah4 or BP or BCR or BCO. The culture was incubated overnight at 25°C. 10 ml of sterile CMB (in a 100 ml flask) was inoculated with 100 µl of the overnight grown culture and incubated at 25 °C for 4 h with shaking (80 rpm) to reach the log phase. An aliquot of 3.0 ml of log phase culture was then used to inoculate two 500 ml flasks each containing 300 ml sterile CMB. The contents of each flask were thoroughly mixed and

distributed aseptically into 60 sterile flasks such that each flask contains 10 ml of inoculated culture. One flask from the set was immediately placed on ice for dilution plating (0 h plating) and the rest were kept in the incubator at 25 °C. At different time intervals, flasks were withdrawn from the incubator, placed on ice for dilution plating on Luria agar (LA) plates. For each such culture, grown for a defined time at 25 °C, a series of dilution tubes was set up to obtain dilutions of 10^{-1} through 10^{-7} of the *Bacillus* strains. Multiple dilutions and plating were used to ensure countable plates for plotting the growth curves.

1.2.9. Dual-species competition among four *Bacillus* spp.; three most prevalent coelomic *Bacillus* species (*B. megaterium*, *B. cereus*, *B. pumilus*) and one chance intruder *B.thuringiensis*

Log phase cells from monocultures of four *Bacillus* species [*B. megaterium* (Ah4); *B. cereus* (BCR); *B. pumilus* (BP); *B. thuringiensis* (BT)] were centrifuged to remove the supernatants, and the cell pellets were resuspended in fresh CMB broth. The cell suspension was adjusted to OD_{600} 0.5, and mixed in 1:1 ratio of Ah4 : BCR; Ah4 : BP; Ah4 : BT; BCR : BP; BCR : BT; and BP : BT. The 1:1 mixture of dual-species contained approximately 10^{10}-10^{11} cfu ml^{-1} of each species. 1 ml of each mixture was inoculated into 99 ml CMB in a 250 ml Erlenmeyer flask. The contents of each flask were thoroughly mixed and distributed aseptically into 10 sterile 125 Erlenmeyer flasks such that each flask contains 10 ml of inoculated culture. One flask from the set was immediately placed on ice for dilution plating (0 h plating) and the rest were kept in the shaker incubator at 25 °C. At different time intervals, flasks were withdrawn from the incubator, placed on ice for dilution plating on HBA plates. For each such culture, grown for a defined time at 25 °C, a series of dilution tubes was set up to obtain dilutions of 10^{-1} through 10^{-7} of the *Bacillus* strains . Multiple dilutions and plating on HBA were used to ensure countable plates for plotting the growth curves. Competitive index (CI) values were determined for every combination of dual-species. CI is calculated as species A/species B ratio within the output sample(x) divided by the corresponding ratio in the inoculum (y). CI = x/y. If the CI value is 1, it would reflect that species A is able to grow as efficiently as the species B, and a CI < 1 would indicate that growth of species A is attenuated.

1.2.10. Growth of *B. megaterium* (Ah4), *B. cereus* (BCR), *B. pumilus* (BP), and *B. coagulans* (BCO) in mixed culture

Log phase cells from monocultures of four *Bacillus* species [*B. megaterium* (Ah4); *B. cereus* (BCR); *B. pumilus* (BP); *B. coagulans* (BCO)]were centrifuged to remove the supernatants, and the cell pellets were resuspended in fresh CMB broth. The cell suspension was adjusted to OD_{600} 0.5, and mixed in 1 : 1: 1: 1 ratio of Ah4 : BCR : BP: BCO. The 1:1:1:1 mixture of quadruple-species contained approximately 10^8-10^9 cfu ml^{-1} of each species. 2 ml of the mixture was inoculated into 198 ml CMB in a 500 ml Erlenmeyer flask. The contents of the flask were thoroughly mixed and distributed aseptically into 20 sterile 125 Erlenmeyer flasks such that each flask contains 10 ml of inoculated culture. One flask from the set was immediately placed on ice for dilution plating (0 h plating) and the rest were kept in the shaker incubator at 25 °C. At different time intervals, flasks were withdrawn from the incubator, placed on ice for dilution plating on HBA plates. For each such mixed-culture, grown for a defined time at 25 °C, a series of dilution tubes was set up to obtain dilutions of 10^{-1} through 10^{-7} of the *Bacillus* strains . Multiple dilutions and plating on HBA were used to ensure countable plates for plotting the growth curves.

1.2.11. Chemotaxis assay

For chemotaxis experiments, four PRCD dwelling Gram positive cells of different species of the genus *Bacillus* (*B. megaterium*, *B. cereus*, *B. pumilus*, *B. coagulans*), one soil-borne pathogen (*B. thuringiensis*) and one Gram negative soil inhabiting species (*Pseudomonas aeruginosa*) were used. Cells of *Bacillus* spp. were grown in CMB medium while *P. aeruginosa* cells were grown in LB. Overnight grown bacteria in CMB or LB medium (0.2 ml) were inoculated into 10 ml sterile CMB or LB in a 250 ml Erlenmeyer flask and kept on a rotary shaker at 30 ^{0}C until $O.D_{600}$ reached 0.2 to 0.4 to attain the log phase. Cells from log phase were transferred (1% v/v) to 5 ml sterile medium (CMB/LB) in a 125 ml Erlenmeyer flask and then incubated as described above until $O.D_{600}$ reached 0.5 to 0.6. Bacterial cultures were centrifuged at 8000 x g for 5 min at 4 ^{0}C. Pellets were washed twice with PBS buffer (= chemotaxis buffer) and the re-

suspended pellet was added up to 2 ml of PBS to bring bacteria to approximately 10^9 per ml. The microfuge tubes, each containing 200 µl bacterial suspensions, were horizontally placed in sterile Petri-plates under the laminar hood. Coelomic fluid (CF), ≤ 5 µl, was collected in a sharpened capillary tube (length, ≤ 4 cm; internal diameter of the piercing end to draw CF, ≤ 0.5 mm) by carefully puncturing the body wall from each worm. After drawing CF, the other ends of the capillary tubes were thoroughly sealed with molten wax. The control capillary contained sterile PBS instead of CF. The open end of each capillary (rinsed with PBS) was then inserted into the microfuge tube containing the bacterial suspension. The apparatus used in this study (Fig. 1.) was basically a hybrid of two earlier apparatus for chemotaxis assay (Adler, 1973; Chakraborty and Roy, 1990).

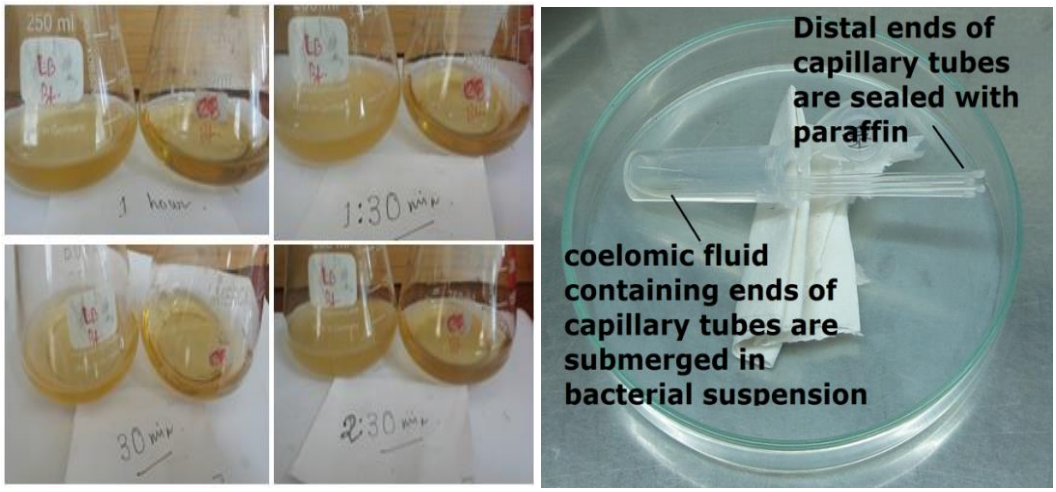


Fig. 1.2: Photographs of experiments on bacterial competition and chemotaxis. A- Culture flasks showing growth of major *Bacillus* spp. in coelom mimicking broth. B- Chemotaxis assay.

After incubation at different time intervals, the capillaries were removed and their exterior surfaces were washed with a thin stream of sterile water, the sealed ends were broken over sterile microfuge tubes containing 1.0 ml PBS such that contents of the open ends were fully released. Suitable dilutions were made in sterile PBS and spread plated in HBA plates (for *Bacillus* spp.) or LA (for *P. aeruginosa*). After incubation overnight at 30 ^{0}C, colonies were counted.

1.2.12. Riboflavin synthesis of *B. megaterium* (Ah4), *B. cereus* (BCR), *B. pumilus* (BP), and *B. coagulans* (BCO) in CMB

Overnight grown cultures of *B. megaterium* (Ah4), *B. cereus* (BCR), *B. pumilus* (BP), *and B. coagulans* (BCO) in CMB were inoculated (2% v/v) into 10 ml sterile CMB in a 250 ml Erlenmeyer flask, and kept on a rotary shaker at 30 ^{0}C until $O.D_{600}$ reached to 0.4 to attain the log phase. Cells from log phase were transferred (1% v/v) to 5 ml CMB kept in culture tubes. Four sets of culture tubes (one set per bacterial culture consisting of 10 culture tubes) wrapped in aluminum foil, to prevent photo-degradation of riboflavin produced by the bacteria, were grown under mild shaking condition (80 rpm) at 25°C. 2 ml cultures from one of each tube of four different sets were taken out at different time intervals in aluminum foil-wrapped microfuge tubes and centrifuged at 8000 x g for 5 min. From each microfuge tube, 1 ml of the supernatant was collected in a fresh aluminium foil-wrapped microfuge tube and mixed with 1 ml of sodium borate buffer. OD was measured at 440 nm in the nano-spectrophotometer (Make) and concentration of riboflavin was determined by plotting the OD value in a standard curve of riboflavin (prepared, prior to the experiment, by following standard method) (Bartzatt and Follis, 2014).

1.2.13. Statistical analysis

All the experiments were carried out in triplicate (n=3). The results are expressed as mean ± SD. The observed p-value corresponding to the F-statistic of one-way ANOVA was ≤ 0.01, which strongly suggested that one or more pairs of the variables were significantly different. One way ANOVA with post-hoc Tukey HSD Test (Kramer, 1956) was done where necessary to compare results of the initial (day 0 or day 1) and final day. The level of significance at 99% or 95% was indicated in the corresponding line or bar graphs as **, $p < 0.01$; and *, $p < 0.05$. Similar statistical analysis was performed and interpreted accordingly on experimental results described in other chapters also.

1.2.14 Isolation and taxonomic identification of bacteria from the gut of *E. fetida*

For isolation of the strains, the earthworms were washed repeatedly with sterile distilled water and then anesthetized by placing into 10% ethanol. The posterior (gut region in the last 1/3rd of the worm's body) gut was aseptically removed from the anaesthetized worm. After surface cleaning by sterile distilled water the gut along with its content was mixed

with 500µL of phosphate buffer saline (pH 7.2) using a micro pestle in a 1.5ml microfuge tube. Coelomic fluid did not require mixing with a pestle. The mixed content was serially diluted (in PBS) and spread on Luria agar and incubated at 25ºC for 48h in aerobic and anaerobic conditions. Distinct and unique colonies that developed on agar plates were picked, single colony purified, put in glycerol stock (80%) and refrigerated (-70 °C) for future study.

1.2. 14.1 Identification and phylogenetic analysis based on 16S rRNA gene

The 16S rRNA genes of strains were amplified from the genomic DNA prepared by standard protocol (Farlong *et al.* 1996). The purified amplicon was ligated to pJET 1.2 blunt vector (ThermoFisher, K1232) and cloned in *E. coli* DH5α. The inserted part was sequenced from both forward and reverse ends to obtain an almost complete 16S rRNA gene sequence. The sequence comparisons with entries in the GenBank and EMBL databases were performed with the BLASTn program (Pearson, 1990; Altschul *et al*., 1990; Altschul *et al*., 1997). To determine the phylogenetic affiliation, the 16S rRNA gene sequences of the strains were aligned with the sequences of members of the closely related bacteria with the CLUSTAL W program (Thompson *et al.*, 1994). Evolutionary relationships of members of the strains were inferred using three different tree-making algorithms: the neighbour-joining (Saitou and Nei, 1987), maximum likelihood (Felsenstein, 1981)and maximum-parsimony (Fitch, 1971) in Mega Ver 7.0 (Kumar *et al.*, 2016). Phylogenetic analyses and the fidelity of the tree topologies were evaluated by bootstrap analysis with 1000 replicates (Felsenstein, 1981; Tamura *et al.*, 2007).

1.2. 14.2 Phenotypic and chemotaxonomic study of selected novel bacterial strains

Growth of the strains were tested at 4 ºC, 10 ºC, 20 ºC, 28 ºC, 30 ºC, 35 ºC, 37 ºC and 42 ºC (±1 ºC). For salt tolerance tests 2%, 4%, 6%, 8%, 10%, 15% or 20 % (w/v) NaCl was added to peptone-yeast extract (PY) medium (composition: 10 g peptone, 5 g yeast extract) devoid of NaCl or KCl. To assess growth at different pH levels, the pH of the sterile LB medium was adjusted from pH 3.0 to 12.0 by using either 0.1 M HCl or 0.1 M NaOH. Results were obtained after 48 h incubation at 28 ºC. The Gram test was performed by the KOH lyses method (Murray *et al.*, 1994) and further confirmed by the

Gram-staining method of Claus (1992). Catalase activity was examined by the production of bubble after addition of few drops of 3% (v/v) H_2O_2 on the full grown colony of strains in LA slant. Voges-Proskauer test was performed by observation of colour development after addition of alpha-naphthol and potassium hydroxide to the Voges-Proskauer broth of the stain's culture. Ability to hydrolyze starch was determined by assessing the development of clear zones (after treatment with Grams iodine solution) around the streaked culture on starch agar plates (nutrient agar 2.3%; soluble starch 0.5%; pH 7.2). Urease test was performed by observing the development of red or pink colour at the periphery of the bacterial colonies on Christensen's urea agar plates (MacFaddin, 2000). Other biochemical characteristics or activities such as β-Galactosidase, nitrate reduction, arylsulfatase (3d), acid phosphatase, pyrazinamidase, α-esterase and β-esterase, citrate utilization, ONPG test, nitrate reduction, and abilities to ferment fructose, inulin, lactose, maltose, mannitol, raffinose, ribose, salicin, sorbitol, sucrose, trehalose etc. were examined by following standard methods (Gordon *et al.*, 1974; Lanyi, 1987; Smiberg and Krieg, 1994). Phenotypic characterization of ET03^T was performed using the Biolog GEN III MicroPlate following manufacturer's instructions. Briefly, bacterial suspension, prepared in special "gelling" inoculating fluid, was transferred to GEN III MicroPlate, (100 µl per well). Incubation was carried out in an aerobic atmosphere for 48 h. Increased respiration due to the growth of bacteria using the single carbon source provided in each well caused reduction of the tetrazolium redox dye, forming a purple colour. The reactions were read using the fully automated OmniLog system. Carbon source utilization assays were double-checked by using HiCarbohydrate kit parts A, B and C (Cat. No. #KB009; HiMedia) according to the manufacturer's protocol. Antibiotic susceptibility (specific for oligotrophic bacteria) was determined according to the method described by Kumar *et al*. (2010).

1.2.14.3 Determination of Respiratory Quinone, Polar lipid, FAME and GC mol% of the strains

For the study of quinones and polar lipids, two-stage lipid extraction method as described by Tindall *et al*. (1989) was followed with volumetric modifications. Extraction with methanol: hexane (2:1 v/v) followed by thin layer chromatography (TLC) was performed.

Briefly, the menaquinone part was purified by running the hexane fraction on TLC silica gel 60 F254 (Merck) using petroleum benzene: di-ethyl-ether (0.85:0.15) as the solvent. Further development of menaquinone components was done using acetone: water - 0.99:0.01 as the solvent and observed under UV_{254}. The methanolic part was taken for extraction of polar lipids using chloroform: methanol: 0.3%NaCl (1:2:0.8) as the extraction medium (Bligh and Dyer1959, Minnikin *et al*., 1984; Tindall, 1985). Polar lipids were separated by two-dimensional TLC. In the first dimension, chloroform: methanol: water (65: 25:4, v/v), and in the second dimension, chloroform: methanol: acetic acid: water (80:12:15:4, v/v) were used as the solvent to partition the polar lipids on silica gel plate. Lipid functional groups were identified using spray reagents specific for phospholipids (Mb-Blue), free amino groups (ninhydrin) and sugars (α-naphtol).

Fatty acids were extracted from 36 h old (exponentially growing) cells grown in tryptone-soy-agar (M290; HiMedia) at 28 °C and then esterified to form fatty acid methyl esters (FAMEs). The FAMEs were then analysed by GC (Hewlett Packard 5890 II plus) and the Sherlock Microbial Identification System using version 4.10 of the TSBA40 library (Microbial ID) as described previously (Kaur *et al.,* 2012).

1.2. 14.4 Scanning Electron Microscopy (SEM) of the strains.

Samples were fixed in glutaraldehyde (3%) and 2% osmium tetraoxide according to the standard protocol The fixed cells were mounted on small glass coverslips, dehydrated using increasing ethanolic gradients. After sputter gold coatingthe details of the cell shape were ascertained with help of a scanning electron microscope (EVO LS10, Zeiss, Germany; JS MIT 100, JEOL Ltd., Japan).

1.2. 14.5 Submission of the strains in recognized 'Type' culture collection centres

The novel strains were submitted to the internationally recognised 'Type' culture collection centres,with prior permission of the National Biodiversity Authority (NBA), Chennai, India.

1.3. Results and Discussion

1.3.1. Cultivable bacteria in cow dung during processing in the laboratory

On the zero day of the processing of RCD, the order of abundance of four different species of *Bacillus* was *B.coagulans* > *B. cereus* > *B. pumilus* > *B. megaterium* (Fig. 1.3-A). With time, up to 4th day of processing, there have been increases observed in case of *B. coagulans* (> 10^{13} cfu g^{-1} on 4 d) and *B. cereus* (> 10^{12} cfu g^{-1} on 4 d). The populations of the other two species, *B. pumilus and B. megaterium*, after getting increased up to the 2nd day of processing, have shown a progressive decrease in cell number on 4th and 6th day (Fig. 1.3-A).

1.3.2. Cultivable bacteria in the habitat of *E. fetida*

In absence of *E. fetida*, the fall and rise in populations of *B.coagulans*, *B. cereus,* and *B. pumilus* in PrCD resulted in a typical signature, fall on day 2↓ rise on day 4↑ fall on day 6↓, reflecting a net decline in numbers compared to their zero-day enumeration. The dynamics of *B. megaterium* numbers resulted into a signature (rise on day 2↑ fall on day 4↓ rise on day 6↑) opposite to other three *Bacillus* spp., showing a marginal increase in number on day 6 compared to day 1 (Fig. 1.3-C).

When earthworms (*E. fetida*) were released on PrCD beds, on 0 days, *B. megaterium* was least in numbers among the four *Bacillus* spp. With time, in presence of *E. fetida*, there has been an almost logarithmic rise in *B. megaterium* reaching highest cell density; while the population of *B. coagulans* declined logarithmically to become least in numbers among the four *Bacillus* spp. (Fig. 1.3-B). The cell density of *B. cereus,* after an initial fall in number on day 2, equalized *B. megaterium* on day 6. Till day 4, the population of *B. pumilus* decreased in orders (similar to that of *B. coagulans),* but has shown an increment on day 6 (Fig. 1.3-B). The results have shown that presence of *E. fetida* and their continuous feeding on PrCD influenced significant change in the population dynamics of four *Bacillus* spp. *E. fetida* had the remarkable positive effect on growth of *B. megaterium* providing the survival advantage. Conversely, the population of most dominant *Bacillus* species in RCD and PrCD, *B. coagulans*, faced extreme selection leading to the fast diminution in the viable count in the presence of the earthworms (*E. fetida*).

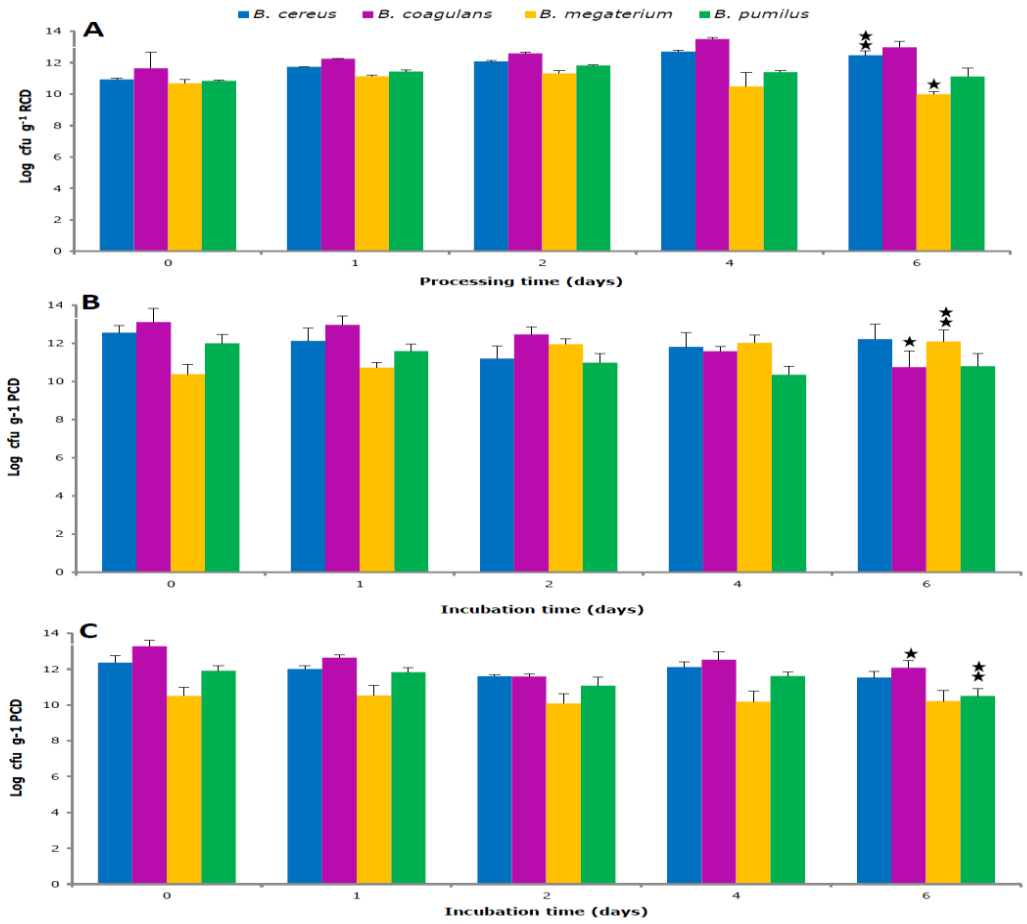


Fig.-1.3: Prevalence of the four most abundant bacilli present in cow dung that: (A) is under processing for 6 days without *E. fetida;* (B) had been processed for 6 days and was then incubated @25°C with *E. fetida* (100 animals KG^{-1}) upto 6 days; (C) had been processed for 6 days and was then incubated @25°C without *E. fetida.* [n=3, **, $p < 0.01$; and *, $p < 0.05$]. (RCD- Raw cow dung, PrCD- Processed cow dung).

1.3.3. Cultivable bacteria in gut of *E. fetida*

When earthworms (*E. fetida*) from the stock culture (maintained in vermiculture pit in the field) were thoroughly washed with sterile distilled water and released in fresh PrCD bed mixed with laboratory-grown cells of *B. thuringiensis*, gut sample collected immediately after the seeding (day 0 sample) contained no detectable number of *B. thuringiensis* while considerable number of it ($\approx 10^{10}$ cfu / g cast) was present on day 1. Population dynamics of *Bacillus* spp. in the gut sample of *E. fetida* on day 1 revealed an

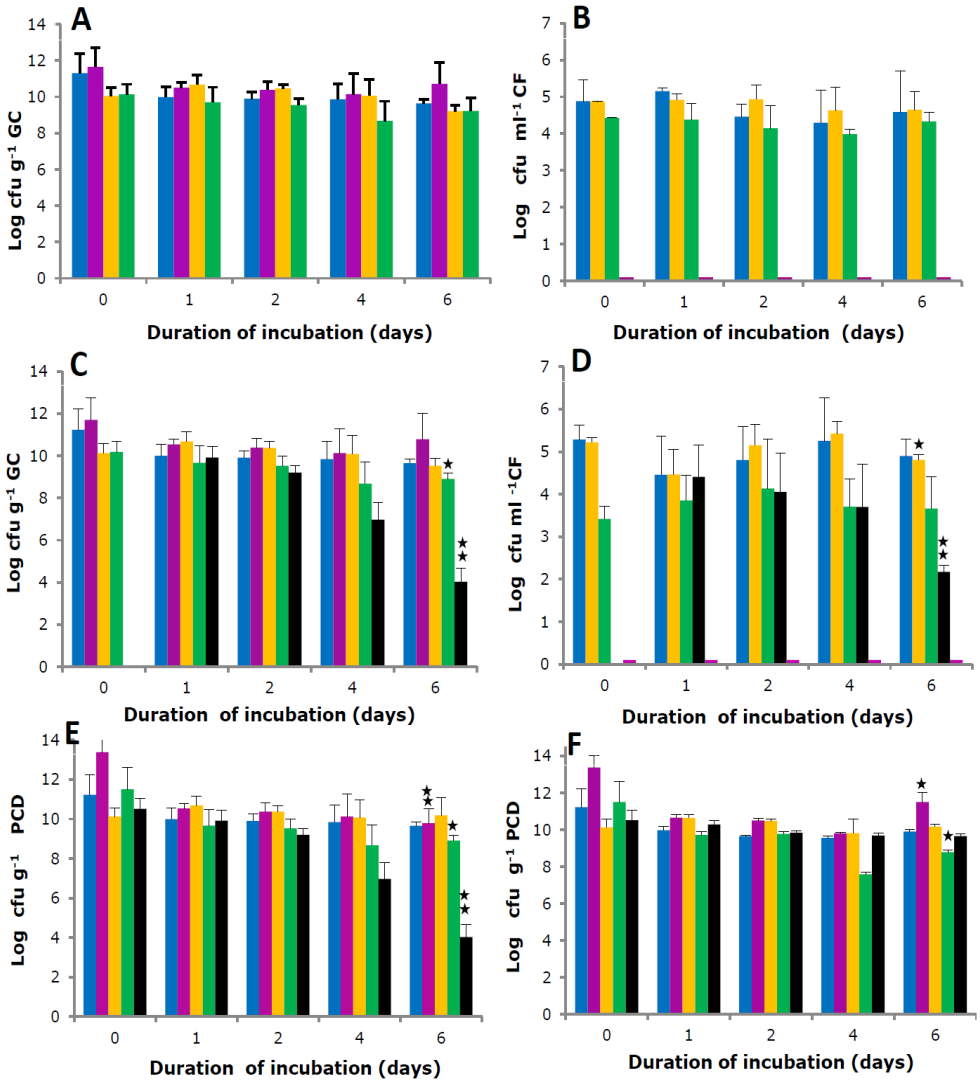


Fig.1.4: Prevalence of the dominant bacilli present in the: (A) gut of *E. fetida* inhabiting processed cow dung (PrCD); (B) coelomic fluid of *E. fetida* inhabiting PrCD; (C) gut of *E. fetida* inhabiting PrCD that had been supplemented with *Bacillus thuringiensis* (a chance pathogen of *E. fetida*); (D) coelomic fluid of *E. fetida* inhabiting PrCD that had been supplemented with *B. thuringiensis;* (E) PrCD that had been supplemented with *B. thuringiensis* and was inhabited by *E. fetida* and (F) PrCD that had been supplemented with *B. thuringiensis* but was not inhabited by *E. fetida.* Prior to supplementation with *B. thuringiensis* and/or introduction of *E. fetida*, all cow dung had been processed for 6 days. For inoculating dung with *B. thuringiensis*, a suspension of exponential phase cells (optical density [600 nm] 0.5) was used. For dung inhabited by *E. fetida*, the worm was immediately introduced (100 animals KG^{-1}) after processing. All dung samples were incubated (@25°C, in the dark) keeping in glass petriplates (diameter=15cm) for up to 6 days. The bacilli were *B. thuringiensis* [■], *B. megaterium* [■], *B. cereus* [■], *B. pumilus* [■] and *B. coagulans* [■].Values are means of three independent replicates (n=3), and bars indicate standard deviation [**, $p < 0.01$; and *, $p < 0.05$].

important fact that though *B. thuringiensis* was abundant, populations of *B. megaterium* dominated over other species of *Bacillus*. The dominance of *B. megaterium* continued on day 2 and 4 with a concurrent progressive reduction in *B. thuringiensis* populations (Fig. 2.4-C). Gut samples on day 6 have shown least representation of *B. thuringiensis* populations (reduced to 10^4 cfu g^{-1} cast) while the abundance of the other four *Bacillus* species was like this: *B. coagulans* > *B. megaterium* (≈ *B. cereus*)> *B.pumilus* (Fig. 1.4-E). Gut samples collected immediately from *E. fetida* just after the seeding (day 0 sample) in PrCD bed without the addition of laboratory grown cells of *B. thuringiensis*, contained no detectable number of *B. thuringiensis* as expected but abundances of *B. megaterium* and *B. pumilus* (≈ 10^{10} cfu / g^{-1} cast) were less by more than an order to that of *B. cereus* and *B. coagulans* (> 10^{11} < 10^{12} cfu g^{-1} cast) (Fig. 1.4 E). Decrease in the populations of *B. cereus, B. pumilus* and *B. coagulans* with the characteristic increase in *B. megaterium* was noted in the gut samples collected on day 1. The dominance of *B. megaterium* persisted on day 2 gut samples of *E. fetida*. Reduction in the population of *B. pumilus* was prominent in gut samples of day 4. Populations of *B. coagulans* exceeded other three Bacillus species in gut samples of day 6 (Fig. 1.4-E).

1.3.4. Cultivable bacteria in coelomic fluid of *E. fetida*

Coelomic fluids (CFs) aseptically withdrawn from multiple numbers of *E. fetida*, in different volumes, for different purposes, and examined many a time under the microscope or plated on rich medium (Luria-Bertani agar) have confirmed the presence of live bacteria. Colonies picked up randomly from such plates were clonally purified for studying their cell's Gram character, morphology, and few physiological and biochemical characteristics. After ascertaining that the majority of them could be putatively assigned to the genus- *Bacillus*, detailed enumeration of heterotrophic bacteria on LA and *Bacillus* spp. on selective HBA were done. Heterotrophic bacterial density in CF was in the range of $2.47 \times 10^5 \pm 1.9 \times 10^4$ cfu ml^{-1}. The density of *Bacillus* spp. in CF was in the range of $1.66 \times 10^5 \pm 1.4 \times 10^4$ cfu ml^{-1} which is approximately 66% of total bacterial load in the CF (Fig 1.5). The differential load of major *Bacillus* spp. was as follows: $6.31 \times 10^4 \pm 1.4 \times 10^3$ cfu ml^{-1} for *B. megaterium*, $4.77 \times 10^4 \pm 2.2 \times 10^3$ cfu ml^{-1} for *B. cereus*, and $1.81 \times 10^4 \pm 1.5 \times 10^3$ cfu ml^{-1} for *B. pumilus.*

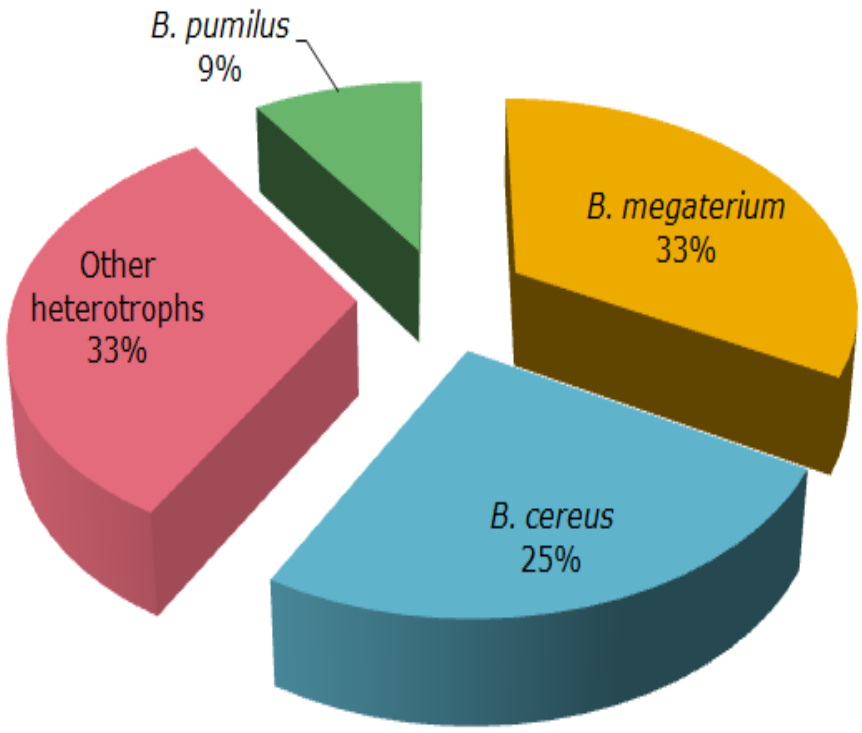


Fig. 1..5: Pie chart showing microbial load in coelomic fluid of *E. fetida* incubated at 25°C with processed cow dung (100 animals kg^{-1}).

The population densities of *B. megaterium* and *B. cereus* in the coelomic fluid of *E. fetida* were almost equal and found higher than *B. pumilus* on day 0. The cells of *B. coagulans*, although dominant in gut microbiota were completely absent in the CF of *E. fetida*. In numbers, *B. cereus* had a slight edge over *B. Megaterium* in CF collected on day 1 (Fig. 1.4-D). The most dominant population observed on day 2 and 4 was *B. megaterium*. The most variable population among the three residents was *B. cereus* (as adjudged by the standard deviation) but the mean value of *B. megaterium* had the slight edge over *B. cereus* on day 6 (Fig. 1.4-D). The population of *B. pumilus* remained always the least in the CF of *E. fetida*. The results indicate that the coelomic fluid appears selective and restrictive (rarely allowed cell densities of any of the three species to go above 10^5 cfu ml^{-1} CF) in hosting the three distinct species of *Bacillus*.

1.3.4.1. Population dynamics of three *Bacillus* species residing in the coelomic fluid of PrCD grown *E. fetida*

The population densities of *B. megaterium* and *B. cereus* in the coelomic fluid of *E. fetida* were almost equal and found higher than *B. pumilus* on day 0. The cells of *B. coagulans*, although dominant in gut microbiota were completely absent in the CF of *E. fetida*. In numbers, *B. cereus* had a slight edge over *B. megaterium* in CF collected on day 1 (Fig. 1.4D). The most dominant population observed on day 2 and 4 was *B. megaterium*. The most variable population among the three residents was *B. cereus* (as adjudged by the standard deviation) but the mean value of *B. megaterium* had the slight edge over *B. cereus* on day 6 (Fig. 1.4D). The population of *B. pumilus* remained always the least in the CF of *E. fetida*. The results indicate that the coelomic fluid appears selective and restrictive (rarely allowed cell densities of any of the three species to go above 10^5 cfu ml^{-1} CF) in hosting the three distinct species of *Bacillus*.

1.3.4.2. Dynamical changes in the population structure of the coelomic *Bacillus* species in the fate of forced introduction of *B. thuringiensis* in the coelome of PrCD grown *E. fetida*

When a high dose of *B. thuringiensis* (a known pathogen of invertebrates from *Bacillus* spp.) suspended in PBS, was injected into the coelomic cavity of PrCD grown *E. fetida*, the worms survived showing abundance of the intruding pathogen in good numbers (> 10^4 < 10^5 c. f.u ml^{-1}) in their CF, 8 h after injection on Day 0. There was a concomitant increase in the populations of *B. megaterium* and *B. cereus* (> 10^5 cfu ml^{-1}) and sinking down of *B. pumilus* population to marginally little above 10^3 cfu ml^{-1} (Fig. 1.6). On day 2, the populations of *B. thuringiensis*, *B. megaterium* and *B. cereus* almost equalized (>10^4 < 10^5 cfu ml^{-1}) and the population of *B. pumilus* got slightly elevated (10^4 cfu ml^{-1}) in the CF of *E. fetida*. Populations of *B. thuringiensis* dropped to 10^4 cfu ml^{-1} and came down to the level of *B. pumilus*, whereas cell density of *B. megaterium* reached the peak followed by *B. cereus* in CF on day 4. Further downsizing of the populations of *B. thuringiensis* in the CF took place on day 6, while the populations of *B. megaterium* and *B. cereus* were restricted to the level just below 10^5 cfu ml^{-1} (Fig. 1.6). The results have shown that coelomic fluid is not only capable of supporting the growth of *Bacillus* spp but also most likely host few bacteria selectively for its own defence against pathogens and/or to derive some other benefit.

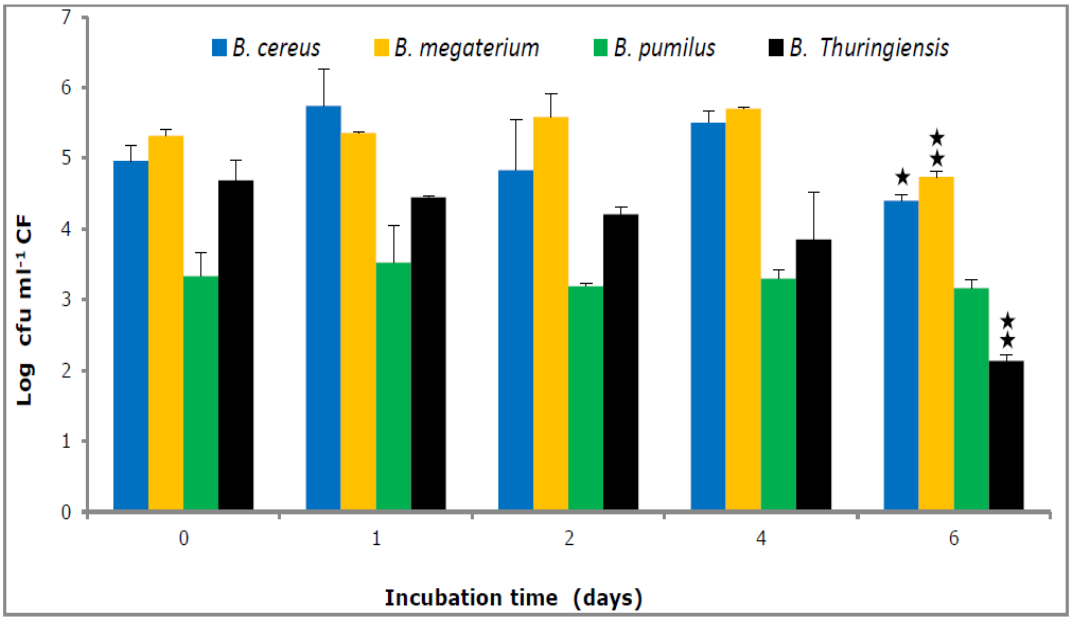


Fig. 1.6: Quantities of three dominant bacilli, *B. cereus*, *B. megaterium*, and *B. pumilus* in CF when coelom of *E. fetida* was injected with laboratory grown *B. thuringiensis* cells [n=3, **, $p < 0.01$; and *, $p < 0.05$].

1.3.5. Culture-independent assessment of eubacterial diversity in the coelomic fluid of *E. fetida*

A culture-independent molecular method based on the analyses of 16S rDNA obtained from total DNA has been used to better understand the diversity of bacteria inhabiting coelomic fluid of PrCD grown *E. fetida*. A random selection of 10 % clones, from the clone library, constituted of 140 clones from 16S rDNA, was made. All recombinant clones yielded a band of ≈ 1.5 kb in size after colony PCR amplification using pJET1.2 forward and reverse sequencing primers. Restriction digestion of the amplicons with *HaeIII* resulted in the number of pattern type. The comparison of ARDRA profiles was performed on the basis of the presence (1) or absence (0) of fragments was generated by PyElph 1.4 software[15]. The similarity between the DNA sequences was computed and used to generate a phylogenetic tree based on clustering methods applied to the distance matrix. The genetic distances are displayed on the branches (Fig. 1.7-A). Six clusters /groups or operational taxonomic units (OTUs) have resulted from this analysis; group I

(Ah4 group) comprised of lanes 12, 13, 14, 11* and 10*; group II (BCR group) consisting of lanes 4,3, and 6*; group III (BP group) comprised of lanes 1 and 7; group IV consisting of lanes 8 and 9; group V represented by lane 2; and group VI represented by lane 5. From the six OTUs, organism belonging to three groups, group I, II, and III were identified, from culture-dependent studies and phylogenetic characterization of the isolates, as *B. megaterium*, *B. cereus*, and *B. pumilus* respectively.

1.3.6. Culture-dependent isolation of *Bacillus* spp. from coelomic fluid of *E. fetida* and phylogenetic characterization of the isolates

Dilution plating of CF on HiChrome Bacillus agar (HBA) plates readily identified three types of colonies. The yellow mucoid colonies were characteristic of *B. megaterium* because of their property to ferment mannitol contained in the medium indicated by the change of the color of phenol red. The colonies of *B. cereus* are flat with distinct blue centers because of its possession of β-glucosidase that cleaves the chromogenic mixture present in the medium. The colonies of *B. pumilus* are raised and pink in colour. More than two colonies, from each of the three colony types, were repeatedly dilution streaked on HBA to obtain pure cultures. The strains, Ah4, BCR, and BP, representative of putative *B. megaterium*, *B. cereus*, and *B. pumilus* respectively, were phylogenetically characterized. During 16S rRNA gene sequence analysis, the pairwise comparison indicated that strain Ah4 shared similarity with *B. megaterium* NBRC 15308 (99.5%); strain BCR shared similarity with *B. cereus* ATCC 14579 (99.8%); strain BP shared similarity with *B. pumilus* NBRC 12092 (99.7%). Other species have shown a lower level of similarity to strain Ah4, BCR, and BP. Phylogenetic trees were constructed by using 16S rRNA gene sequences with neighbor-joining (Fig. 1.7-B), maximum likelihood and maximum parsimony methods (not shown) in MEGA (ver 7.0). Regardless of different evolutional comparisons, similar topology was obtained in all phylogenetic trees, which indicates that the strains Ah4, BCR, and BP belong to the *B. megaterium*, *B. cereus*, and *B. pumilus* species cluster respectively.

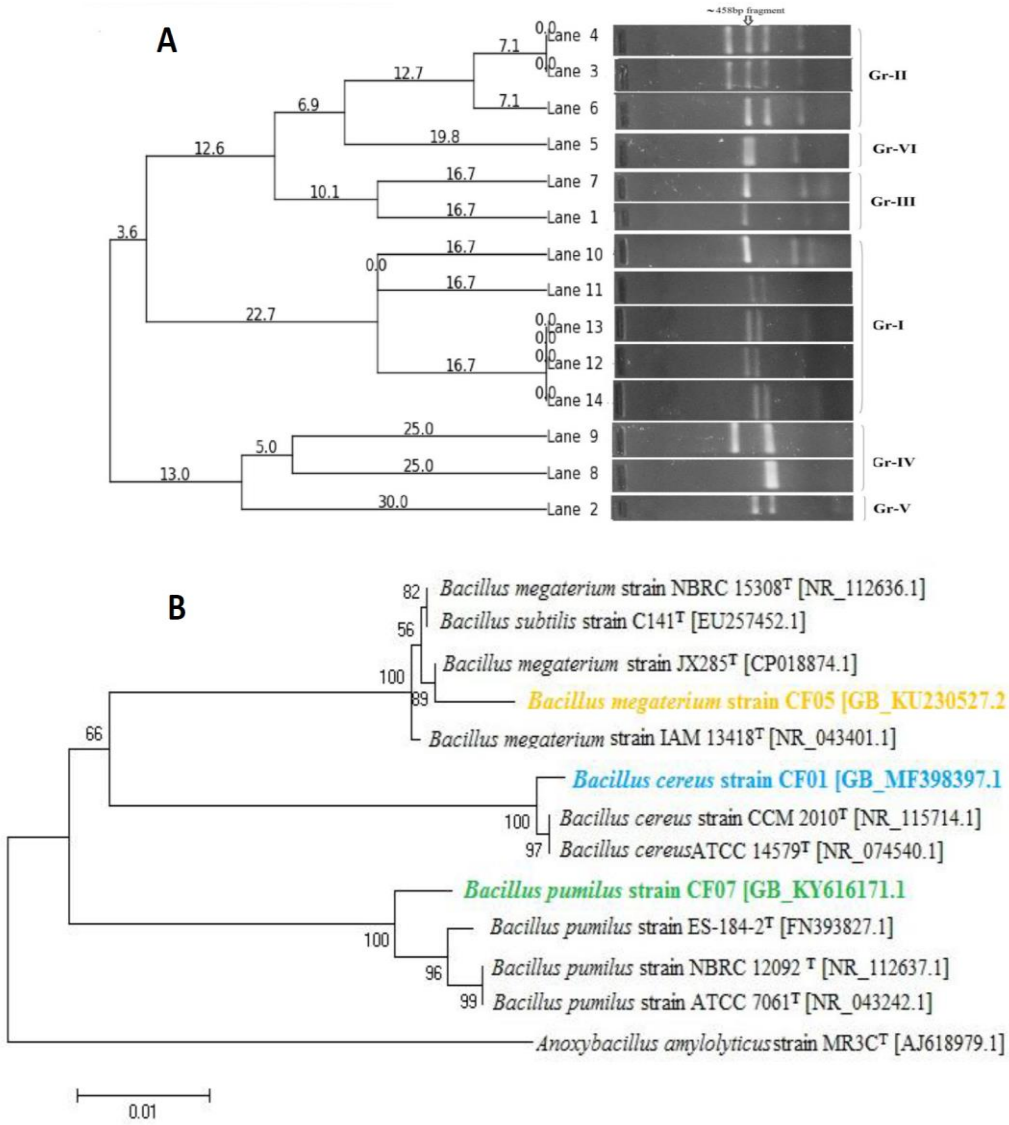


Fig. 1.7: Bacilli present in coelomic fluid of *E. fetida* assessed by: (A) Culture independent and (B) Culture dependent analysis. For culture independent assessment a phyllogram based on UPGMA matrices was computed (using PyElph software) from gel image of HaeIII digest of 16S rRNA gene amplicons derived from randomly selected 14 clones of the clone library created from metagenome extracted from coelomic fluid of *E. fetida.* Clusturing of homologous groups was based on genetic distance (expressed by numerical values on the dendrogram) in respect to sharing of restriction sites; Six groups have been identified- Gr-I [Lane numbers- 12,13,14, 11 and 10], Gr-II [Lane numbers- 4,3 and 6], Gr-III [Lane numbers- 1 and 7], Gr-IV [Lane numbers- 8 and 9], Gr-V [Lane number- 2] & Gr-VI [Lane number- 5]. Group-I,II and III have similar band patterns corresponding to HaeIII digest of 16S rRNA genes from *B. megaterium*, *B. cereus* and *B. pumilus* respectively (data not shown). For culture dependent assessment phylogeny of the three bacilli (indicated by coloured fonts) isolated from coelomic fluid of *E. fetida* has been computed based on Neighbour Joining matrices. Bootstrap values (expressed as percentages of 1000 replications) >50% are shown at branch points. Bar, 1 substitution per 100 nucleotide position.

1.3.7. Individual growth curves of *B. megaterium* Ah4, *B. cereus* BCR, and *B. pumilus* BP, and *B. coagulans* BCO (four predominant Bacillus species in RCD or PrCD) in CMB

It was evident that there is a dynamic presence of four *Bacillus* species, *B. megaterium*, *B. cereus*, *B. pumilus*, and *B. coagulans* in RCD, PrCD, and gut samples of *E. fetida*. On the contrary, coelomic fluid of the worms selectively hosted three of the four *Bacillus* spp., rejecting *B. coagulans*. Hence, carrying out *in-vitro* growth kinetics of each of the four *Bacillus* species in Coelomic fluid Mimicking Broth (CMB) became necessary to understand the cause of rejection. CMB was formulated (detailed composition given in the materials and methods) on the basis of determination of mean content (g ml^{-1}) of sugar, protein, free amino acids, triglycerides, urea, and uric acid in the coelomic fluid of *E. fetida* fed with PrCD, along with data from published literature on small molecule composition of CF (Bundya *et al.*, 2001).

The composition of the medium was largely determined by the objective of the experiment (to simulate mixed culture condition in the coelomic fluid of *E. fetida*), and the properties of the strains. Also, it was presumed that the metabolic activity of the cells of the individual species population shall modify the composition of CMB. It was observed that CMB supported the growth of all the four *Bacillus* spp. (Fig.1.8-Ai-iv). Taking cell densities, at two time points, 0.5 h (initial, t_1) and 3.0 h (final, t_2) as logA and logB respectively, and number of generations (n) worked out by the equation: n = (logB – logA)/log2, the generation time was calculated as: $(t_2 - t_1)/n$. The generation time of *B. cereus*, *B. megaterium*, *B. pumilus*, and *B. coagulans*, corresponding to their growth curves (Fig. 6 a, b and c), were 33.4, 24.6, 127.0, and 86.6 min respectively. Although the results have shown that in the CMB medium, *B. megaterium* and *B. pumilus* were the fastest and slowest growing *Bacillus* species respectively, yet the growth performance of *B. coagulans* was found better than *B. pumilus*. Hence, it appears that some other factor(s) like proneness to be phagocytosed by the immunocytes or any metabolic deficiency of *B. coagulans* in producing the specific product (upon which host may be benefitted) have had led to evolutionary selection/rejection in the coelom of *E. fetida*.

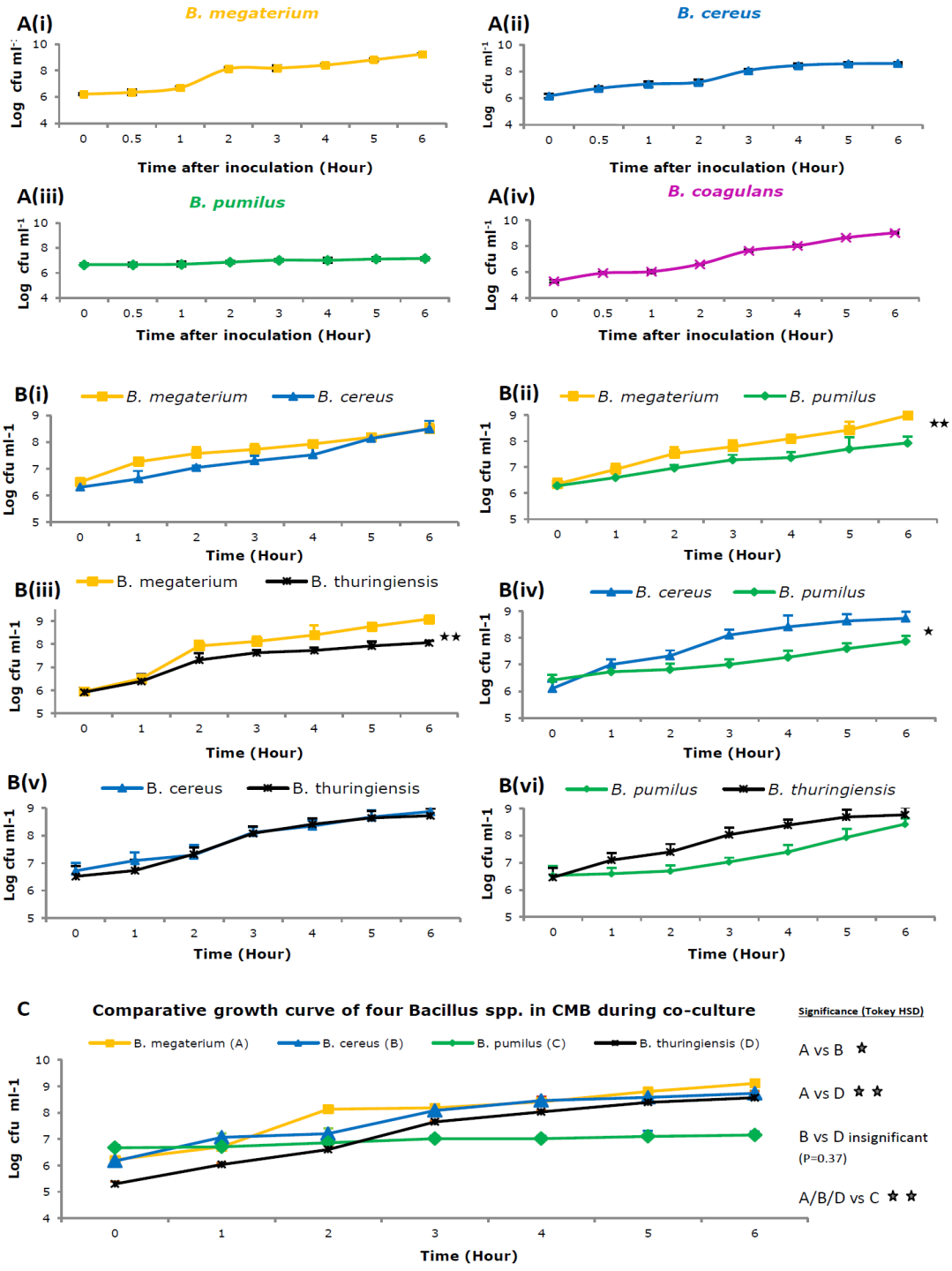


Fig. 1.8: Growth of pure cultures (of CF isolates) in Coelome mimicking broth(CMB) (6Ai to iv):Growth of cocultures in CMB (6Bi to vi): Growth of mixed culture in CMB (6C). [n=3, **, $p < 0.01$; and *, $p < 0.05$].

1.3.8. Dual-species competition among four *Bacillus* spp.; three most prevalent coelomic Bacillus species (*B. megaterium, B. cereus, B. pumilus*) and one chance intruder *B.thuringiensis*

Dual-species batch cultures, for all possible combinations (*B. cereus* and *B. megaterium*; *B. pumilus* and *B. megaterium*; *B. thuringiensis* and *B. megaterium*; *B. pumilus* and *B. cereus*; *B. thuringiensis* and *B. cereus*; & *B. pumilus* and *B. thuringiensis*) were studied to infer interaction among each other of the pair (Fig. 1.8). Competitive index (CI) value determined for the dual-species batch culture for *B. cereus* and *B. megaterium* or *B. thuringiensis* and *B. cereus*; was 1.0, which means that *B. megaterium* or *B. thuringiensis* is able to grow as efficiently as *B. cereus* in CMB (Fig. 1.8-A & E). For the dual-species batch-culture for *B. pumilus* and *B. megaterium* or *B. pumilus* and *B. cereus* or *B. pumilus* and *B. thuringiensis,* the C I value obtained was 0.87, 0.86, and 0.95 respectively (CI < 1.0 in all three cases), reflecting the fact that growth of *B. pumilus* is attenuated in all the three cases (Fig. 1.8-B, D, & F). CI value, obtained from dual-species batch culture for *B. thuringiensis* and *B. cereus,* was 0.93, which means that the growth of *B. thuringiensis* is attenuated (Fig. 1.8-C). The results obtained from dual-species competition have also revealed that the growth of *B. pumilus* was much better in dual-species batch culture, particularly combo with *B. thuringiensis,* than its performance as the solo or in mixed culture.

1.3.9. Growth of *B. megaterium* (Ah4), *B. cereus* (BCR), *B. pumilus* (BP), and *B. coagulans* (BCO) in mixed culture

The growth of the mixed culture of four different species of the genus *Bacillus* (Fig. 1.8-C) has supported that the complexity of CMB medium plays a role in determining inter-species interaction. At different time intervals, individual species in the mixed culture displayed differential growth rates. In the first one hour interval (0 -1h), highest growth rate was recorded for *B. cereus* (K = 2.96 h^{-1}) followed by *B. coagulans* (K = 2.40 h^{-1}), while in the following next hour (1 – 2h) *B. megaterium* displayed the highest growth rate (K = 4.75 h^{-1})compared to any other growth rate recorded for any of the remaining three species in successive time periods (Fig. 1.8-C). During 2-3h time period, *B. coagulans* (K = 3.44 h^{-1}) exceeded other three species in growth rate while in the following time period (3-4 h), *B. cereus* equalled the growth rate of *B. coagulans* (K = 1.25 h^{-1}). During 4-5 h

time period, *B. coagulans* maintained its dominance in growth rate (K = 2.06 h^{-1}). In the last lap, 5-6 h, the high growth rate was shown by *B. megaterium (*K = 1.42 h^{-1}) and *B. cereus (*K = 1.26 h^{-1}). Throughout the period of observation, the growth of *B. pumilus* remained least and inhibited, almost reached a plateau after 3 h of incubation. The results have clearly indicated the growth advantage of *B. coagulans* in mixed culture compared to its slow growth rate in solo culture. On the contrary, the growth of *B. pumilus* was at a high disadvantage, even found slower than its solo performance. Mixed culture of four different *Bacillus* species in CMB medium has exhibited synergism because it attained significantly higher cell density and produced more bacterial biomass (Fig. 1.8-C). Such synergistic growth may result when multiple species produce complementary enzymes and participate in the metabolite cross feeding enabling bacteria to consume substrates cooperatively. The problems of feedback inhibition and metabolite repression present in mono-species culture may be lessened in mixed culture.

1.3.10. Chemotactic responses of *B. megaterium*, *B. cereus*, *B. pumilus*, *B. coagulans*, *B.thuringiensis,* and *Pseudomonas aeruginosa* to coelomic-fluid of *E. fetida*

In order to investigate the phenotype involved in the establishment of cow-dung bacteria-*Eisenia* interactions and pathogenicity caused due to primary or secondary infection of soil-borne pathogens, chemotaxis has intuitively emerged in our study design. With the background information on small molecule composition of coelomic fluid of *E. fetida* being dominated by organic acids like succinate (being most prominent), malonate, acetate, α-ketoglutarate, and formate; amino acids like tyrosine and alanine; and other organic compounds like myo-inositol, glycerol, methanol, and nicotinamide mono nucleotide (Bundya *et al.,* 2001), taken together with earlier reports on taxis towards naturally occurring amino acids, sugars, or extracts of leaves of Birch, Poplar, Apple and Hawthorn by few species of *Bacillus* (Lebenko *et al.,* 2005)and identification of a chemoreceptor for TCA intermediates in *P. aeruginosa* (Lacal *et al.,* 2010) capillary assays were carried out to determine the repertoire of responses of the test organisms to coelomic fluid of *E. fetida*. The number of bacteria attracted into a capillary tube containing CF as attractant was measured. Kinetics of accumulation of the three different bacteria (*B. megaterium* strain Ah4, *B. cereus* strain BCR and *B. pumilus* strain BP isolated from CF of *E. fetida*) in capillaries was presented in Fig. 1.9. Highest

accumulation of cells in capillaries in 5 min was recorded for *B. megaterium* [significantly higher than *B. cereus* (**, $p < 0.01$) and *B. pumilus* (**, $p < 0.01$)]. Both *B. megaterium and B. cereus* have shown highest and 2nd highest accumulation up to fifteen minutes, although the difference was less significant (*, $p < 0.05$) at fifteenth minute; numbers decreased with time in *B. megaterium* while no characteristic change in

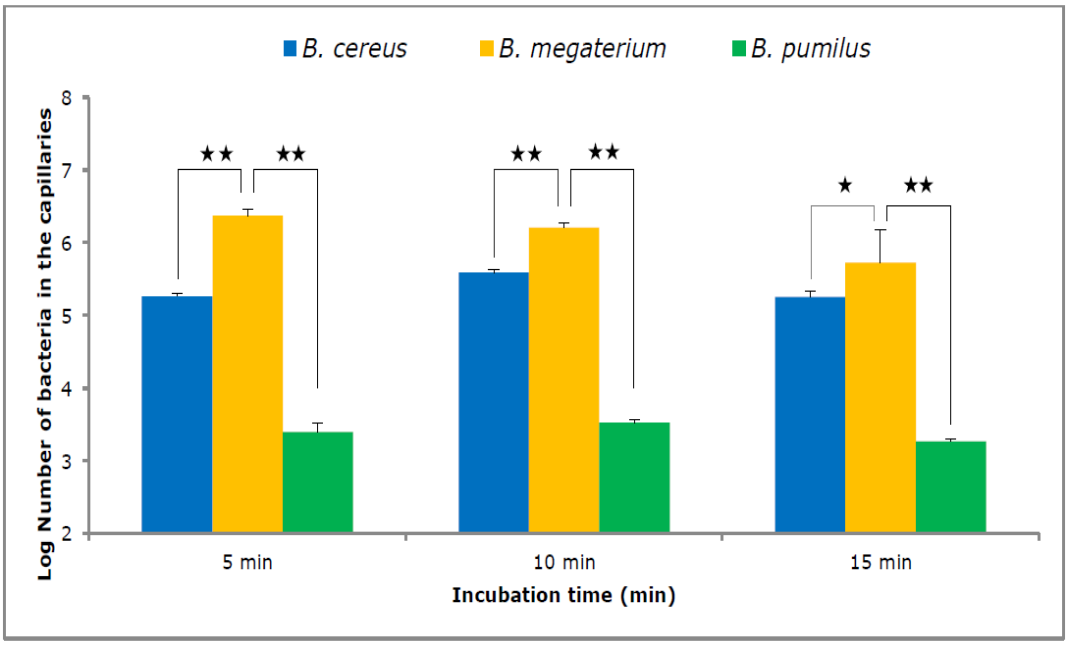


Fig. 1.9: Rate of accumulation of bacteria in capillaries containing coelomic fluid of *E. fetida* [n=3, **, $p < 0.01$; and *, $p < 0.05$].

numbers with time was noted in case of *B. cereus* The number of cells accumulated over time was least in case of *B. pumilus* (Fig. 1.9). To minimize error, testing with the blank was repeated at least 6 times. As a group, *Bacillus* exhibited differential levels of chemotaxis (chemotaxis of *B. megaterium* > chemotaxis of *B. cereus* > chemotaxis of *B. pumilus*) towards coelomic fluid of *E. fetida* where *B. megaterium* seems to be the most efficient in early Chemotaxis (Fig.1.9).

1.3.11. Riboflavin synthesis of *B. megaterium* Ah4, *B. cereus* BCR, *B. pumilus* BP, and *B. coagulans* BCO in CMB medium

Earthworms are vulnerable to invasion by bacterial pathogens because of their highly permeable integument. Yet, a high degree of bactericidal potency is also evident due to

the activities of the immune-competent cells which stores riboflavin to their advantage (Mazur *et al.*, 2011). Despite the inability of the host to synthesize riboflavin, its storage predominates in free coelomocytes of eleocyte-rich earthworm species like *E. fetida* (Santocki *et al.*, 2016). Thus, it appeared that riboflavin might be the key currency in host-bacteria symbiosis in the coelomic fluid. Agreements of symbiotic interactions between eukaryotic host and bacteria in several occasions have proven the involvement of bacterial riboflavin production and secretion (LeBlanc *et al.*, 2013). Hence, the ability of the *Bacillus* strains to secrete riboflavin in the CMB medium was assayed. *B. megaterium* exceeded the other three species (Fig. 1.10) In terms of secretion of riboflavin in the medium, with the characteristic cycle of release (highest secretion peak, > 3.0 µM, at 0.5 h) and uptake (signified by the decrease in the concentration). The other two strains, *B. cereus* and *B. pumilus*, which colonizes CF, have also shown their distinctive riboflavin secretion and uptake signatures.

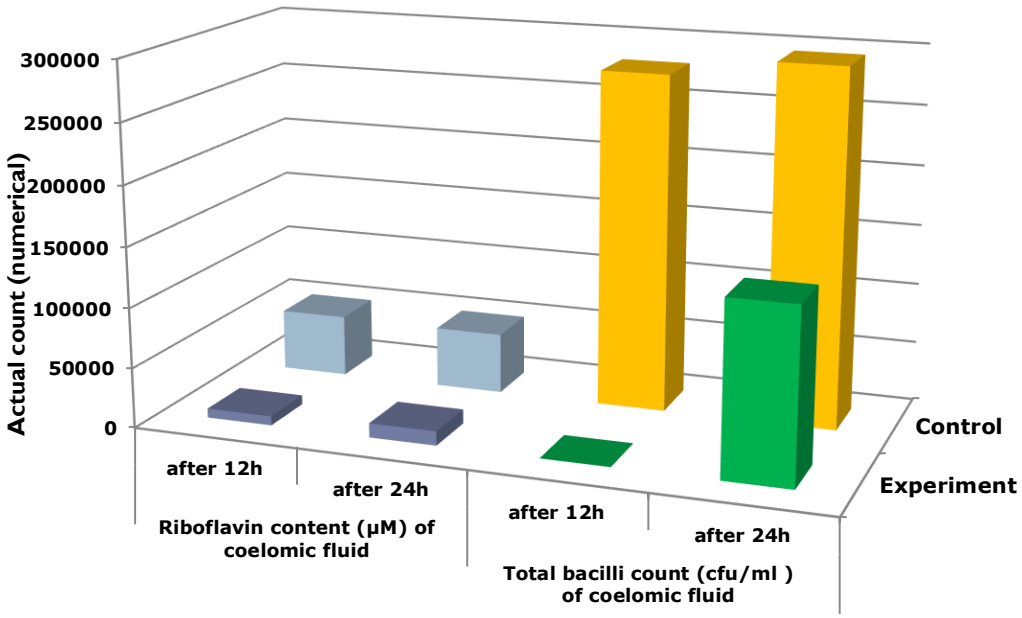


Fig. 1.10: Total bacilli count and concentration of riboflavin in the coelomic fluid of *E. fetida* after 12h and 24h of levofloxacin injection in comparison to the control animals without antibiotic treatment.

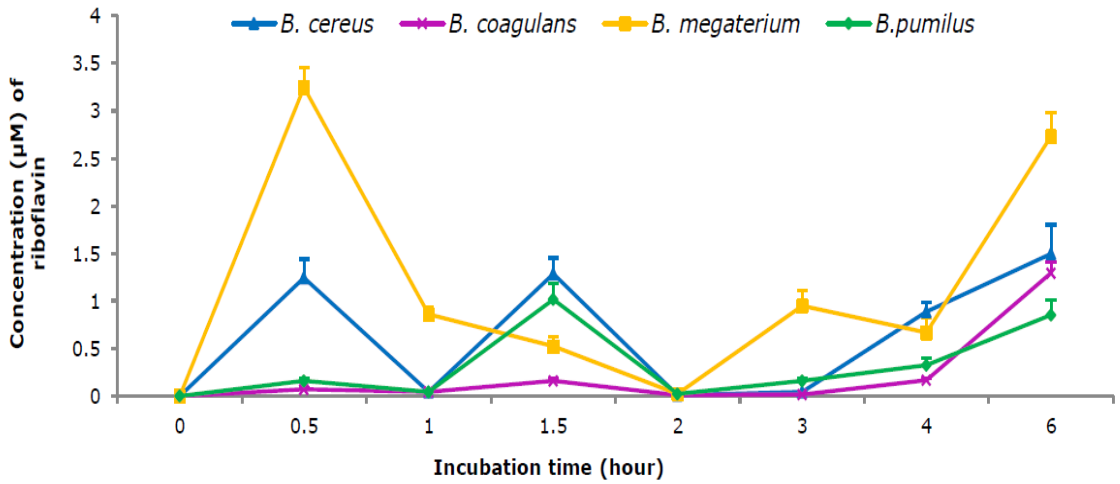


Fig. 1.11: Riboflavin secretion by pure cultures of three CF isolates (*B. cereus* CF01, *B. megaterium* CF05, and *B. pumilus* CF07) and one PrCD isolate (absent in CF), *B. coagulans* PrCD01, in CMB medium (n=3).

The secretion peaks (> 1.0 µM) at 0.5 and 1.5 h were noted for *B. cereus*, while only a conspicuous peak (< 1.0 µM) at 1.5 h was observed for *B. pumilus*. The species (which was not detected in CF but conspicuously present in PrCD and gut) *B. coagulans* secreted least or insignificant amount of riboflavin till 4 h of incubation in CMB (Fig. 1.11).

1.3.12. Isolation and taxonomic identification of bacteria from gut of *E. fetida*

Altogether 20 unique colonies (8 from cast, 3 from posterior gut, 5 from anterior gut and 4 from the whole gut in anaerobic condition) excluding 5 from coelomic fluid were isolated. 40% of the bacteria were Firmicutes; Proteobacteria and Actinobacteria were found in equal amount (30% each) (Fig. 1.12). A detailed phylogenetic tree (Fig. 1.13) shows the phylogeny of all the 20 isolates.

List of cultivable bacterial strains isolated from gut and coelomic fluid of *E. fetida* and their 16S rRNA gene accession numbers (NCBI) are provided in Table 1.1

Table 1.1: List of cultivable bacterial strains associated with *E. Fetida* and their 16S rRNA gene accession numbers.

Sl No.	Strain	NCBI Accession Numbers	16SrRNA gene sequence	Taxonomy
			From cast	
1	*Bacillus efetidiens* strain ET01 (=MCC3263, =LMG 29877, =JCM 32462, =KCTC 33945)	KU1751 76.2	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGGATCGATGGGAGCTTGCTC CCTGAGATCAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTA ATACCGGATAATTTAGTTCCTCGCATGAGGAACTGTTGAAAGGTGGCTTCGGCTACCACTTACAGATGGACCCGCGGCGCAT TAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGA GACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGT GAGTGAAGAAGGTTTTCGGATCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCCGTTCGAATAGGGCGGCACCTCGACGG TACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTG GGCGTAAAGCGCGCGCAGGTGGTTTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGG GAACTTGAGTGCAGAAGAGGAAAGTGGAATTCCAAGTGTAGCGGTGAAATGCGTAGATATTTGGAGGAACACCAGTGGCG AAGGCGACTTTCTGGTCTGTAACTGACACTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCAC GCCGTAAACGATGAGTGCTAAGTGTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGA GTACGGTCGCAAGACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAAC GCGAAGAACCTTACCAGGTCTTGACATCCTCTGACAACCCTAGAGATAGGGCTTTCCCCTTCGGGGGACAGAGTGACAGGT GGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAG CATTCAGTTGGGCACTCTAAGATGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTA TGACCTGGGCTACACACGTGCTACAATGGACGGTACAAAGGGCAGCGAGACCGCGAGGTTTAGCCAATCCCATAAAACCGT TCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAAT ACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGAGGTAACCTTTTGGAGC CAGCCGCCTAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAGGTAACCGTA	Phylum: Firmicutes Class:Bacilli Order: Bacillales Family: Bacillaceae (Sp. Nov.)
2	*Bacillus sp. strain ET02*	KU2305 22.2	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGGATCGATGGGAGCTTGCTCCCT GAGATCAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACC GGATAATTTAGTTCCTCGCATGAGGAACTGTTGAAAGGTGGCTTCGGCTACCACTTACAGATGGACCCGCGGCGCATTAGCTAG TTGGTGAGGTAACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCC CAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGAAGAAG GTTTTCGGATCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCCGTTCGAATAGGGCGGCACCTTGACGGTACCTGACCAGAA AGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCG CAGGTGGTTTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTTGGGTCTAGAGTACGGCAC AGGAAAGTGGAATTCCAAGTGTAGCGGTGAAATGCGTAGATATTTGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGTCTGT AACTGACACTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAA GTGTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTCGCAAGACTGAAACTC AAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGAC ATCCTCTGACAACCCTAGAGATAGGGCTTTCCCCTTCGGGGGACAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCG TGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCATTCAGTTGGGCACTCTAAGATGACTGCC GGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACG GTACAAAGGGCAGCGAGACCGCGAGGTTTAGCCAATCCCATAAAACCGTTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGC ATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACC ACGAGAGTTTGTAACACCCGAAGTCGGTGAGGTAACCTTTTGGAGCCAGCCGCCTAAGGTGGGACAGTTGATTGGGGTGAAGT CGTAACAAGGTAACCGTA	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae
3	*Chryseomicrobium excrementi* strain ET03^{T} (= KCTC 33943^{T} = LMG 30119^{T} = JCM 32415^{T})	KU2305 23.2	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGGATGACGGAGGAGCTTGCT CCTCTAGATTCAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCCTGTAGATTGGGATAACTCCGGGAAACCGGGGC TAATACCGAATAATCCTTCGGACCTCATGGTCCGATGTTGAAAGACGGCTCCGGCTGTCACTACAGGATGGGCCCGCGGCGC ATTAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACT GAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCACAATGGACGAAAGTCTGATGGAGCAACGCCGC GTGAGTGAAGAAGGTTCTCGGATCGTAAAACTCTGTTGTGAGGGAAGAACAAGTACCGGAGTAACTGCCGGTACCTTGACG GTACCTCATTAGAAAACCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATCATT GGGCGTAAAGCGCGCGCAGGCGGTCCCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGG GGGACTTGAGTGCAGAAGAGGAAAGCGGAATTCCAAGTGTAGCGGTGAAATGCGTAGAGATTTGGAGGAACACCAGTGGC GAAGGCGGCTTTCTGGTCTGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA CGCCGTAAACGATGAGTGCTAAGTGTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGG AGTACGGTCGCAAGACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCA ACGCGAAGAACCTTACCAGGTCTTGACATCCCGCCGACCGCCCAGGAGACTGGGCCTTCCCTTCGGGGACGGCGGTGACAG GTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCC AGCATTCAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCT TATGACCTGGGCTACACACGTGCTACAATGGATGGTACAAAGGGCTGCGAACCCGCGAGGGGGAGCCAATCCCATAAAACC ATTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGTGGATCAGCATGCCACGGTGA ATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGGGGTAACCTTTATGG AGCCAGCCGCCGAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAGGTAACCGTA	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Planococcaceae (Sp. Nov.)
4	*Bacillus eisenifilia* strain ET04 (=MCC 3370, =JCM 32463, =KCTC 33944, =LMG 30120)	KX8392 65.1	>GAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGGATTGATGGGAGCCTGCTCCC TGATATCAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAAT ACCGGATAACTCATTTCCTCGCATGAGGAAATGTTGAAAGGTGGCTTTTAGCTACCACTTACAGATGGACCCGCGGCGCATT AGCTAGTTGGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAG ACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACAGAGCAACGCCGCGTG AGTGATGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTGCCGTTCGAATAGGGCGGCACCTTGACGGT ACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTGG GGGTAAAGCGCGCGCAGGTGGTTCCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGG AACTTGAGTGCAGAAGAGGAAAGTGGAATTCCAAGTGTAGCGGTGAAATGCGTAGATATTTGGAGGAACACCAGTGGCGA AGGCGACTTTCTGGTCTGTAACTGACACTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACG CCGTAAACGATGAGTGCTAAGTGTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAG TACGGTCGCAAGACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACG CGAAGAACCTTACCAGGTCTTGACATCCTCTGACAACCCTAGAGATAGGGCTTTCCCCTTCGGGGGACAGAGTGACAGGTG GTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGC ATTCAGTTGGGCACTCTAAGATGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTAT GACCTGGGCTACACACGTGCTACAATGGACGGCACAAAGGGCAGCGAGACCGCGAGGTTTAGCCAATCCCATAAAACCGTT CTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATA CGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGAGGTAACCTTTTGGAGCC AGCCGCCTAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAGGTAACCGTAATCTTTCTAGAAGATTTCCCTCACCC	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae (Sp. Nov.)
5	*Pseeudomonas aeruginosa strain ET05 (=MCC 3462)*	MF3983 95.1	>AGAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGGATGAAGGGAGCTTGCTCCTG GATTCAGCGGCGGACGGGTGAGTAATGCCTAGGAATCTGCCTGGTAGTGGGGGATAACGTCCGGAAACGGGCGCTAATACCGC ATACGTCCTGAGGGAGAAAGTGGGGGATCTTCGGACCTCACGCTATCAGATGAGCCTAGGTCGGATTAGCTAGTTGGTGGGGT AAAGGCCTACCAAGGCGACGATCCGTAACTGGTCTGAGAGGATGATCAGTCACACTGGAACTGAGACACGGTCCAGACTCCTA CGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGAAGGTCTTCGGATT GTAAAGCACTTTAAGTTGGGAGGAAGGGCAGTAAGTTAATACCTTGCTGTTTTGACGTTACCAACAGAATAAGCACCGGCTAA CTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGTGGTTCAG CAAGTTGGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATCCAAAACTACTGAGCTAGAGTACGGTAGAGGGTGGTGGAA TTTCCTGTGTAGCGGTGAAATGCGTAGATATAGGAAGGAACACCAGTGGCGAAGGCGACCACCTGGACTGATACTGACACTGA GGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTCGACTAGCCGTTGGGGTCCT TGAGATCTTAGTGGCGCAGCTAACGCGATAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAATGAATTGACG GGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCTGGCCTTGACATGCTGAGAACTT TCCAGAGATGGATTGGTGCCTTCGGGAACTCAGACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTA AGTCCCGTAACGAGCGCAACCCTTGTCCTTAGTTACCAGCACCTCGGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCAG AGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGGCCAGGGCTACACACGTGCTACAATGGTCGGTACAAAGGGTTG CCAAGCCGCGAGGTGGAGCTAATCCCATAAAACCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCGTGAAGTCGGAAT CGCTAGTAATCGTGAATCAGAATGTCACGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTGGGTT GCTCCAGAAGTAGCTAGTCTAACCGCAAGGGGGACGGTTACCACGGAGTGATTCACGACTGGGGTGAAGTCGTAACAAGGTAA CCGTA	Phylum: Proteobacteria Class: Gammaproteobacteria Order: Pseudomonadales Family: Pseudomonadaceae

6	*Bacillus sp. strain ET08*	KU2305 24	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGGACAGAAGGGAGCTTGCTCCC GGATGTTAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGAGCTAATAC CGGATAGTTCCTTGAACCGCATGGTTCAAGGATGAAAGACGGTTTCGGCTGTCACTTACAGATGGACCCGCGGCGCATTAGCTA GTTGGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGC CCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAA GGTTTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTGCAAGAGTAACTGCTTGCACCTTGACGGTACCTAACCAGAAA GCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGGGCTCGC AGGCGGTTTCTTAAGTCTGATGTGAAAGCCCCCGGCTCAACCGGGGAGGGTCATTGGAAACTGGGAAACTTGAGTGCAGAAGA GGAGAGTGGAATTCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGACTCTCTGGTCTGTA ACTGACGCTGAGGAGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGT GTTAGGGGGGGTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTCGCAAGACTGAAACTC AAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGAC ATCCTCTGACAACCCTAGAGATAGGGCTTTCCCTTCGGGGACAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTG AGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCATTCAGTTGGGCACTCTAAGGTGACTGCCGG TGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGATAGA ACAAAGGGCTGCGAGACCGCAAGGTTTAGCCAATCCCACAAATCTGTTCTCAGTTCGGATCGCAGTCTGCAACTCGACTGCGTG AAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACG AGAGTTTGCAACACCCGAGGTCGGTGAGGTAACCTTTATGGAGCCAGCCGCCGAAGGTGGGGCAGATGATTGGGGTGAAGTCG TAACAAGGTAACCGTA	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae
7	*Bacillus sp. strain ET09*	KU2305 25	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGGACAGAAGGGAGCTTGCTCCC GGATGTTAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGAGCTAATAC CGGATAGTTCCTTGAACCGCATGGTTCAAGGATGAAAGACGGTTTCGGCTGTCACTTACAGATGGACCCGCGGCGCATTAGCTA GTTGGTGAGGTAACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGC CCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAA GGTTTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTGCAAGAGTAACTGCTTGCACCTTGACGGTACCTAACCAGAAA GCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGGGCTCGC AGGCGGTTTCTTAAGTCTGATGTGAAAGCCCCCGGCTCAACCGGGGAGGGTCATTGGAAACTGGGAAACTTGAGTGCAGAAGA GGAGAGTGGAATTCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGACTCTCTGGTCTGTA ACTGACGCTGAGGAGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGT GTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTCGCAAGACTGAAACTCAA AGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACAT CCTCTGACAACCCTAGAGATAGGGCTTTCCCTTCGGGGACAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAG ATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCATTCAGTTGGGCACTCTAAGGTGACTGCCGGTG ACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACAGAAC AAAGGGCTGCGAGATCGCAAGGTTAGCCAATCCCACAAATCTGTTCTCAGTTCGGATCGCAGTCTGCAACTCGACTGCGTGAAG CTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGA GTTTGCAACACCCGAAGTCGGTGAGGTAACCTTTATGGAGCCAGCCGCCGAAGGTGGGGCAGATGATTGGGGTGAAGTCGTAA CAAGGTAACCGTA	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae
8	*Celullosaemicrobium bengalensis* strain ET10 (=MCC 3389, =LMG 30121, =KCTC 49085)	KU2305 26.2	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGATGATGCCCAGCTTGCTGG GTGGATTAGTGGCGAACGGGTGAGTAACACGTGAGTAACCTGCCCTTGACTTCGGGATAACTCCGGGAAACCGGGGCTAAT ACCGGATATGAGCCGCCCTCGCATGGGGGTGGTTGGAAAGTTTTTCGGTCAGGGATGGGCTCGCGGCCTATCAGCTTGTTGG TGGGGTGATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCACACTGGGACTGAGACACGGCCC AGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAGTGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGATGAA GGCCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGCAAGTGACGGTACCTGCAGAAGAAGCGCCAGCTAACTACGT GCCAGCAGCCGCGGTAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGTAGGCGGTTTGTCGCG TCTGGTGTGAAAACTCGAGGCTCAACCTCGAGCTTGCATCGGGTACGGGCAGACTAGAGTGCGGTAGGGGAGACTGGAATT CCTGGTGTAGCGGTGGAATGCGCAGATATCAGGAGGAACACCGATGGCGAAGGCAGGTCTCTGGGCCGCAACTGACGCTGA GGAGCGAAAGCATGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGTTGGGCACTAGGTGTGGGGCTC ATTCCACGAGTTCCGTGCCGCAGCAAACGCATTAAGTGCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAA TTGACGGGGGCCCGCACAAGCGGCGGAGCATGCGGATTAATTCGATGCAACGCGAAGAACCTTACCAAGGCTTGACATGCA CGGGAAGCCACCAGAGATGGTGGTCTCTTTGGACACTCGTGCACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGAT GTTGGGTTAAGTCCCGCAACGAGCGCAACCCTCGTCCATGTTGCCAGCGGGTTATGCCGGGGACTCATGGGAGACTGCCG GGGTCAACTCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGTCTTGGGCTTCACGCATGCTACAATGGCCG GTACAAAGGGCTGCGATACCGTAAGGTGGAGCGAATCCCAAAAAGCCGGTCTCAGTTCGGATTGGGGTCTGCAACTCGACC CCATGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAACGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTC AAGTCACGAAAGTCGGTAACACCCGAAGCCCATGGCCCAACCGTTCGCGGGGGGAGTGGTCGAAGGTGGGACTGGCGATT GGGACTAAGTCGTAACAAGGTAACCGTA	Phylum: Actinobacteria Class: Actinobacteria Subclass: Actinobacteridae Order: Actinomycetales Suborder: Micrococcineae Family: Promicromonosporaceae (Sp. Nov.)
			from anterior gut	
9	*Pseudoxanthomonas* sp. strain EAG1	KX8392 66.1	>AGAGTTTGATCCTGGCTCAGAGTGAACGCTGGTGGTAGGCCTAACACATGCAAGTCGAACGGCAGCACAGGAGAGCTTGCTC TCTGGGTGGCGAGTGGCGGACGGGTGAGGAATACATCGGAATCTACTCTGTCGTGGGGGATAACGTAGGGAAACTTACGCTAA TACCGCATACGACCTACGGGTGAAAGTGGGGGACCGCAAGGCCTCACGCGATTGAATGAGCCGATGTCGGATTAGCTAGTTGG CGGGGTAAAGGCCCACCAAGGCGACGATCCGTAGCTGGTCTGAGAGGATGATCAGCCACACTGGAACTGAGACACGGTCCAG ACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGCAAGCCTGATTCCAGCCATACCGCGTGAGTGAAGAAGGCC GCTCGGGTTGTAAAGCTCTTTTGTTGGGAAAGAAATCCTGCCGGTTAATACCTGGCAGGGATGACGGTACCCAAAGAATAAGC ACCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTACTCGGAATTACTGGGCGTAAAGCGTGCGTAG GTGGTTGTTTAAGTCTGTCGTGAAAGCCCTGGGCTCAACCTGGGAATTGCGATGGATACTGGGCAACTAGAGTGTGGTAGAGG GTGGCGGAATTCCCGGTGTAGCAGTGAAAATGCGTAGAGGATCGGGAGGAACATCCGTGGCGAAGGCGGCCACCTGGGCCAA CACTGACACTGAAGGCACGAAAGCGTGGGGGAGCAGAACAGGGATTAGATACCCTGGTAGTCCACGCCCTAAACGATGCGAA CTGGAATTGTTGGGTTCAACTTGGAACCCAGTATCGAAGCTAACGCGTTAAGTTCGCCGCCTGGGGAGTACGGTCGCAAGACTG AAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGTATGTGGTTTAATTCGATGCAACGCGAAGAACCTTACCTGGT CTTGACATCCACGGAACTTTCCAGAGATGGATTGGTGCCTTCGGGAACCGTGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGT GTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCTTAGTTGCCAGCACGTAATGGTGGGAACTCTAAGGAG ACCGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGACCAGGGCTACACACGTACTACAA TGGGAAGGACAGAGGGCTGCAATCCCGCGAGGGGAGCCAATCCCAGAAACCTTCTCTCAGTCCGGATCGGAGTCTGCAACTC GACTCCGTGAAGTCGGAATCGCTAGTAATCGCAGATCAGCATTGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCG TCACACCATGGGAGTTTGTTGCACCAGAAGCAGGTAGCTTAACCTTCGGGAGGGCGCTTGCCACGGTGTGGCCGATGACTGGG GTGAAGTCGTAACAAGGTAACCGTA	Phylum: Proteobacteria Class:Gamma-Proteobacteria Order: Xanthomonadales Family: Xanthomonadaceae
10	*Streptomyces lapidiformes* strain EAG2^{T} (=LMG 30118^{T} = KCTC 49082^{T} = JCM 32459^{T})	KY5939 24	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGATGAAGCCCTTCGGGGTGG ATTAGTGGCGAACGGGTGAGTAACACGTGGGCAATCTGCCCTGCATTCTGGGACAAGCCCTGGAAACGGGGTCTAATACCG GATATGACTGTCCACCGCATGGTGGATGGTGTAAAGCTCCGGCGGTGCAGGATGAGCCCGCGGCCTATCAGCTAGTTGGTG AGGTAGTGGCTCACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCACACTGGGACTGAGACACGGCCCAG ACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGATGACGG CCTTCGGGTTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAAGTGACGGTACCTGCAGAAGAAGCGCCGGCTAACTACGTGC CAGCAGCCGCGGTAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGTAGGCGGCTTGTCACGTC GGTTGTGAAAGCCCGGGGCTTAACCCCGGGTCTGCAGTCGATACGGGCAGGCTAGAGTTCGGTAGGGGAGATCGGAATTCC TGGTGTAGCGGTGAAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGGATCTCTGGGCCGATACTGACGCTGAG GAGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGCACTAGGTGTGGGCAAC ATTCCACGTTGTCCGTGCCGCAGCTAACGCATTAAGTGCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAA TTGACGGGGGCCCGCACAAGCGGCGGAGCATGTGGCTTAATTCGACGCAACGCGAAGAACCTTACCAAGGCTTGACATACA CCGGAAACGTCTGGAGACAGGCGCCCCCTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATG TTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCCGTGTTGCCAGCAGGCCCTTGTGGTGCTGGGGACTCACGGGAGACC GCCGGGGTCAACTCGGAGGAAGGTGGGGACGACGTCAAGTCATCATGCCCCTTATGTCTTGGGCTGCACACGTGCTACAAT GGCCGGTACAATGAGCTGCGATACCGCGAGGTGGAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGATTGGGGTCTGCAACT CGACCCCATGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCATTGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCC CCGTCACGTCACGAAAGTCGGTAACACCCGAAGCCGGTGGCCCAACCCCTTGTGGGAGGGAGCTGTCGAAGGTGGGACTGG CGATTGGGACGAAGTCGTAACAAGGTAACCGTA	Phylum: Actinobacteria Class: Actinobacteria Order: Actinomycetales Family: Streptomycetaceae (Sp. Nov.)
11	*Pradoshia cofamensis* strain EAG3^{T} (=LMG 30312^{T} =JCM 32460^{T})	MF3983 91.1	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGAATCAGATGGGAGCTTGCT CCCGGATGATTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCCGTAAGATCGGGATAACTCCGGGAAACCGGTGC TAATACCGAATAGTTCTGACACATCATGGTGTCAGAGGGAAAGTTGGCTTCGGCTAACACTTACGGATGGGCCCGCGGCG CATTAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGAC TGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCG CGTGAGCGAAGAAGGCCTTCGGGTCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTGCCACTCGAATAGGCTGGCACCTTGA CGGTACCTGACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTA TTGGGCGTAAAGCGCGCGCAGGCGGTCCTTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACT GGAGGACTTGAGTGCAGAAGAAGAGTGGAATTCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCAGTG GCGAAGGCGACTCTTTGGTCTGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTC CACGCCGTAAACGATGAGTGCTAAGTGTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGG GGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAG	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae (Gen. Nov.,

			CAACGCGAAGAACCTTACCAGGTCTTGACATCCTCTGACAACCCTAGAGATAGGGCCTTCCCCTTCGGGGGACAGAGTGAC AGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCTTAGTTG CCAGCATTCAGTTGGGCACTCTAGGGAGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCC CTTATGACCTGGGCTACACACGTGCTACAATGGGTGGTACAAAGGGCAGCGAGACCGCGAGGTTAAGCAAATCCCATAAAA CCACTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGT GAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGGGGTAACCCTTAC GGGAGCCAGCCGCCGAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAGGTAACCGTA	Sp. Nov.)
12	*Mumia enteroni* strain EAG4 (=LMG 30314, =KCTC 49083)	MF3 983 93.1	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAGCGGTAAGGCCCTTCGGGGTAC ACGAGCGGCGAACGGGTGAGTAACACGTGAGTAATCTGCCCTTCACTCTGGGATAACCGCTGGAAACGGTGGCTAATACCG GATATGACCTTTGCCCGCATGGGTGTTGGTGGAAAGTTTTTCGGTGGAGGATGAGCTCGCGGCCTATCAGCTTGTTGGTGGG GTGATGGCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGACCGGCCACACTGGGACTGAGACACGGCCCAGAC TCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCAACGCCGCGTGAGGGATGACGGCC TTCGGGTTGTAAACCTCTTTCAGTAGGGACGAAGCACAAGTGACGGTACCTACAGAAGAAGGACCGGCCAACTACGTGCCA GCAGCCGCGGTAATACGTAGGGTCCGAGCGTTGTCCGGAATTATTGGGCGTAAAGGGCTCGTAGGCGGTTTGTCGCGTCGG GAGTGAAAACTTGGGGCTTAACCCCATTCGTGCTTCCGATACGGGCGGACTAGAGGCAGGCAGGGGAGAACGGAATTCCTG GTGTAGCGGTGGAATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGGTTCTCTGGGCCTGTTCTGACGCTGAGGAG CGAAAGCATGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGTTGGGCGCTAGGTGTGGGATCCATTC CACGGGTTCCGTGCCGCAGCTAACGCATTAAGCGCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGA CGGGGGCCCGCACAAGCGGCGGAGCATGCTGATTAATTCGATGCAACGCGAAGAACCTTACCTGGGTTTGACATACACCAG AAGCCTGCAGAGATGTGGGTCTCTTTGGACACTGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTG GGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCTATGTTGCCAGCACGTAATGGTGGGGACTCATAGGAGACTGCCGGGGT CAACTCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGCCCCTTATGTCCAGGGCTTCGAGCATGCTACAATGGCCGGTA CAAAGGGCTGCGAAGCCGTAAGGTGGAGCGAATCCCAAAAAGCCGGTCTCAGTTCGGATTGGGATCTGCAACTCGACCCCA TGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAACGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCAC GTCATGAAAGTCGGCAACACCCGAAGCCGGTGGCCTAACCCCTTGTGGGAGGGAGCTGTCGAAGGTGGGGCTGGCGATTGG GACGAAGTCGTAACAAGGTAACCGTA	Phylum: Actinobacteria Class: Actinobacteria Subclass: Actinobacteridae Order: Actinomycetales Suborder: Propionibacterineae Family: Nocardioidaceae (Sp. Nov.)
13	Streptomyces sp. strain EAG5	MF3 983 914	>AGAGTTTGATCCTGGCTCAGGATCGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGATGAACCGCTTTCGGGCGGG GATTAGTGGCGAACGGGTGAGTAACACGTGGGCAATCTGCCCTGCACTCTGGGACAAGCCCTGGAAACGGGGTCTAATACCGG ATATGACTGTCCATCGCATGGTGGATGGTGTAAAGCTCCGGCGGTGCAGGATGAGCCCGCGGCCTATCAGCTTGTTGGTGAGGT AGTGGCTCACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGCGACCGGCCACACTGGGACTGAGACACGGCCCAGACTCCT ACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGACGCCGCGTGAGGGATGACGGCCTTCGGG TTGTAAACCTCTTTCAGCAGGGAAGAAGCGAAAGTGACGGTACCTGCAGAAGAAGCGCCGGCTAACTACGTGCCAGCAGCCGC GGTAATACGTAGGGCGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCTCGTAGGCGGCTTGTCACGTCGGTTGTGAAAGC CCGGGGCTTAACCCCGGGTCTGCAGTCGATACGGGCAGGCTAGAGTTCGGTAGGGGAGATCGGAATTCCTGGTGTAGCGGTGA AATGCGCAGATATCAGGAGGAACACCGGTGGCGAAGGCGGATCTCTGGGCCGATACTGACGCTGAGGAGCGAAAGCGTGGGG GAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGCACTAGGTGTGGGCAACATCCCACGTTGTCCGTGC CGCAGCTAACGCATTAAGTGCCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAA GCGGCGGAGCATGTGGCTTAATTCGACGCAACGCGAAGAACCTTACCAAGGCTTGACATACACCGGAAACGTCTGGAGACAGG CGCCCCCTTGTGGTCGGTGTACAGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGC GCAACCCTTGTCCCGTGTTGCCAGCAGGCCCTTGTGGTGCTGGGGACTCACGGGAGACCGCCGGGGTCAACTCGGAGGAAGGT GGGGACGACGTCAAGTCATCATGCCCCTTATGTCTTGGGCTGCACACGTGCTACAATGGCCGGTACAATGAGCTGCGATACCGC GAGGTGGAGCGAATCTCAAAAAGCCGGTCTCAGTTCGGGTTGGGGTCTGCAACTCGACCCCATGAAGTCGGAGTCGCTAGTAA TCGCAGATCAGCATTGCTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCACGAAAGTCGGTAACACCCGA AGCCGGTGGCCCAACCCCTTGTGGGAGGGAGCTGTCGAAGGTGGGACTGGCGATTGGGACGAAGTCGTAACAAGGTAACCGTA	Phylum: Actinobacteria Class: Actinobacteria Order: Actinomycetales Family: Streptomycetaceae
From posterior gut				
14	M*ycobacterium omicsense* strain EPG1^T (=LMG 30313^T =JCM 32461^T =KCTC 49084^T)	MF3 983 92.1	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGGAAAGGCCCTTCGGGGTACTCG AGTGGCGAACGGGTGAGTAACACGTGGGTGATCTGCCCTGCACTTTGGGATAAGCCTGGGAAACTGGGTCTAATACCGAATAG GATCACTGGCTTCATGGTCGGTGGTGGAAAGCTTTTGCGGTGTGGGATGGGCCCGCGGCCTATCAGCTTGTTGGTGGGGTAATG GCCTACCAAGGCGACGACGGGTAGCCGGCCTGAGAGGGTGACCGGCCACACTGGGACTGAGATACGGCCCAGACTCCTACGGG AGGCAGCAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCGACGCCGCGTGAGGGATGACGGCCTTCGGGTTGTAA ACCTCTTTCGCCAGGGACGAAGCGCAAGTGACGGTACCTGGAGAAGAAGGACCGGCCAACTACGTGCCAGCAGCCGCGGTAAT ACGTAGGGTCCGAGCGTTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGGTGGTTTGTCGCGTTGTTCGTGAAAACTCACAGC TTAACTGTGGGCGTGCGGGCGATACGGGCAGACTAGAGTACTGCAGGGGAGACTGGAATTCCTGGTGTAGCGGTGGAATGCGC AGATATCAGGAGGAACACCGGTGGCGAAGGCGGGTCTCTGGGCAGTAACTGACGCTGAGGAGCGAAAGCGTGGGGAGCGAAC AGGATTAGATACCCTGGTAGTCCACGCCGTAAACGGTGGGTACTAGGTGTGGGTTTCCTTCCTTGGGATCCGTGCCGTAGCTAAC GCATTAAGTACCCCGCCTGGGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGCGGAGCA TGTGGATTAATTCGATGCAACGCGAAGAACCTTACCTGGGTTTGACATGCACAGGACGCCGGCAGAGATGTCGGTTCCCTTGTG GCCTGTGTGCAGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCT CATGTTGCCAGCACGTAATGGTGGGGACTCGTGGGAGACTGCCGGGGTCAACTCGGAGGAAGGTGGGGATGACGTCAAGTCAT CATGCCCCTTATGTCCAGGGCTTCACACATGCTACAATGGCCGGTACAAAGGGCTGCGATGCCGTGAGGTGGAGCGAATCCTTT CAAAGCCGGTCTCAGTTCGGATCGGGGTCTGCAACTCGACCCCGTGAAGTCGGAGTCGCTAGTAATCGCAGATCAGCAACGCTG CGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACGTCATGAAAGTCGGTAACACCCGAAGCCGGTGGCCTAACCCTT GTGGAGGGAGCCGTCGAAGGTGGGATCGGCGATTGGGACGAAGTCGTAACAAGGTAACCGTA	Phylum: Actinobacteria Class: Actinobacteria Subclass: Actinobacteridae Order: Actinomycetales Suborder: Corynebacterineae Family: Mycobacteriaceae (Sp. Nov.)
15	*Bacillus anthracis* strain EPG2	KX8 392 67.1	>AGCAGATAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGAATGGATTAAGAGCTTG CTCTTATGAAGTTAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCCATAAGACTGGGATAACTCCGGGAAACCGGGGCTA ATACCGGATAACATTTTGAACCGCATGGTTCGAAATTGAAAGGCGGCCTCGGCTGTCACTTATGGATGGACCCGCGTCGCATTAGC TAGTTGGTGAGGTAACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGC CCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGAACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAAGGC TTTCGGGTCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATAAGCTGGCACCTTGACGGTACCTAACCAGAAAGCCA CGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATTATTGGGCGTAAAGCGCGCGCAGGTGGT TTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTGAGTGCAGAAGAGGAAAGTGG AATTCCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGTCTGTAACTGACACTGAG GCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAGAGGGTTTCC GCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGGC CCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCTCTGACAACCCTAGAGA TAGGGCTTCTCCTTCGGGAGCAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAAC GAGCGCAACCCTTGATCTTAGTTGCCATCATTAAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATG ACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACAGTACAAGGAGCTGCAAGACCGCGAGGTGGAG CTAATCTCATAAAACCGTTCTCGGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCA TGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGGGGTAACC TTTTTGGAGCCAGCCGCCTAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAGGTAACCGTAACTTCTAGAAGTCTCCTACAA TTTTCCG	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae
16	*Bacillus cereus* strain EPG3	KX8 392 68.1	>GGGATAATTGTATGAGACTTCTAGAAGATAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAATACATGCAAGTC GAGCGAATGGATTAAGAGCTTGCTCTTATGAAGTTAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCCATAAGACTGGGAT AACTCCGGGAAACCGGGGCTAATACCGGATAACATTTTGAACCGCATGGTTCGAAATTGAAAGGCGGCTTCGGCTGTCACTTATGG ATGGACCCGCGTCGCATTAGCTAGTTGGTGAGGTAACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCC ACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAA CGCCGCGTGAGTGATGAAGGCTTTCGGGTCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATAAGCTGGCACCTTGA CGGTACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATTATTGGG CGTAAAGCGCGCGCAGGTGGTTTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTG AGTGCAGAAGAGGAAAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGCGACTTTC TGGTCTGTAACTGACACTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGT GCTAAGTGTTAGAGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAA CTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGA CATCCTCTGAAAACCCTAGAGATAGGGCTTCTCCTTCGGGAGCAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGA GATGTTGGGTTAAGTCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCATCATTAAGTTGGGCACTCTAAGGTGACTGCCGGTGAC AAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACGGTACAAAG AGCTGCAAGACCGCGAGGTGGAGCTAATCTCATAAAACCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGA ATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTA ACACCCGAAGTCGGTGGGGTAACCTTTTGGAGCCAGCCGCCTAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAGGTAACC GTA	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae
Facultative anaerobic or microaerofilic isolated from gut				
17	*Acinetobacter desnitrificae* EDN1 (=MCC	MF 398 396.	>AGAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAGCGGGGGAAGGTAGCTTGCTAC TGGACCTAGCGGCGGACGGGTGAGTAATGCTTAGGAATCTGCCTATTAGTGGGGGACAACATTCCGAAAGGAATGCTAATA CCGCATACGTCCTACGGGAGAAAGCAGGGGACCTTCGGGCCTTGCGCTAATAGATGAGCCTAAGTCGGATTAGCTAGTTGG TGGGGTAAAGGCCTACCAAGGCGACGATCTGTAGCGGGTCTGAGAGGATGATCCGCCACACTGGGACTGAGACACGGCCCA GACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGGGGAACCCTGATCCAGCCATGCCGCGTGTGTGAAGAAG	Phylum: Proteobacteria Class:

	3461, =KCTC 62305, =JCM 32455)	1	GCCTTATGGTTGTAAAGCACTTTAAGCGAGGAGGAGGCTACTGAGACTAATACTCTTGGATAGTGGACGTTACTCGCAGAAT AAGCACCGGCTAACTCTGTGCCAGCAGCCGCGGTAATACAGAGGGTGCGAGCGTTAATCGGATTTACTGGGCGTAAAGCGT GCGTAGGCGGCCATTTAAGTCAAATGTGAAATCCCCGAGCTTAACTTGGGAATTGCATTCGATACTGGATGGCTAGAGTATG GGAGAGGATGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGATGGCGAAGGCAGCCATCT GGCCTAATACTGACGCTGAGGTACGAAAGCATGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCATGCCGTAAACGATG TCTACTAGCCGTTGGGGCCTTTGAGGCTTTAGTGGCGCAGCTAACGCGATAAGTAGACCGCCTGGGGAGTACGGTCGCAAG ACTAAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGATGCAACGCGAAGAACCTTA CCTGGCCTTGACATACTAGAAACTTTCCAGAGATGGATTGGTGCCTTCGGGAATCTAGATACAGGTGCTGCATGGCTGTCGT CAGTTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTTTCCTTACTTGCCAGCATTTCGGATGGGAAC TTTAAGGATACTGCCAGTGACAAACTGGAGGAAGGCGGGGACGACGTCAAGTCATCATGGCCCTTACGGCCAGGGCTACAC ACGTGCTACAATGGTCGGTACAAAGGGTTGCTACCTAGCGATAGGATGCTAATCTCAAAAAGCCGATCGTAGTCCGGATTG GAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTAGTAATCGCGGATCAGAATGCCGCGGTGAATACGTTCCCGGGCCTT GTACACACCGCCCGTCACACCATGGGAGTTTGTTGCACCAGAAGTAGGTAGTCTAACCGCAAGGAGGACGCTTACCACGGT GTGGCCGATGACTGGGGTGAAGTCGTAACAAGGTAACCGTA	Gammaproteobacteria Order: Pseudomonadales Family: Moraxellaceae (Sp. Nov.)
18	*Klebsiella nitrificae* strain EN1	MF564193.1	>AGAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGGTAGCACAGAGAGCTTGCT CTCGGGTGACGAGCGGCGGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTACTGGAAACGGTAGC TAATACCGCATAACGTCGCAAGACCAAAGTGGGGGACCTTCGGGCCTCATGCCATCAGATGTGCCCAGATGGGATTAGCTG GTAGGTGGGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAGAGGATGACCAGCCACACTGGAACTGAGACACG GTCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCCATGCCGCGTGTGTGAA GAAGGCCTTCGGGTTGTAAAGCACTTTCAGCGGGGAGGAAGGCGGTGAGGTTAATAACCTCATCGATTGACGTTACCCGCA GAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAA GCGCACGCAGGCGGTCTGTCAAGTCGGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATTCAAAACTGGCAGGCTAGA GCTTTGTAAAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGGTGGCGAAGGCGGCC CCCTGGACAAAGACTGACGCTCAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCTGTAAAC GATGTCGATTTGGAGGTTGTCCCTTGAGGCGTGGCTTCCGGAGCTAACGCGTTAAATCGACCGCCTGGGGAGTACGGCCGC AAGGTTAAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGATGCAACGCGAAGAACC TTACCTGGTCTTGACATCCACAGAACTTTCCAGAGATGGATTGGTGCCTTCGGGAACTGTGAGACAGGTGCTGCATGGCTGT CGTCAGCTCGTGTTGTGAAATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATCCTTTGTTGCCAGCGGTTAGGCCGGG AACTCAAAGGAGACTGCCAGTGATAAACTGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGACCAGGGCTA CACACGTGCTACAATGGCATATACAAAGAGAAGCGACCTCGCGAGAGCAAGCGGACCTCATAAAGTATGTCGTAGTCCGGA TTGGAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTAGTAATCGTAGATCAGAATGCTACGGTGAATACGTTCCCGGGC CTTGTACACACCGCCCGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCGGGAGGGCGCTTACCAC TTTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAACCGTA	Phylum: Proteobacteria Class: Gammaproteobacteria Order: Enterobacteriales Family: Enterobacteriaceae (Sp. Nov.)
19	*Paracoccus rammohunii* strain ENF1	MF564191.1	>AGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGAGATCTTCGGATCTAGCGG CGGACGGGTGAGTAACGCGTGGGAATATGCCCTTCTCTACGGAATAGCCTCGGGAAACTGGGAGTAATACCGTATACGCCC TACGGGGAAAGATTTATCGGAGAAGGATTAGCCCGCGTTGGATTAGGTAGTTGGTGGGGTAATGGCCTACCAAGCCTACG ATCCATAGCTGGTTTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGG AATCTTAGACAATGGGGGAAACCCTGATCTAGCCATGCCGCGTGAGTGATGAAGGCCTTAGGGTTGTAAAGCTCTTTCAGCT GGGAAGATAATGACGGTACCAGCAGAAGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGGGCTAGC GTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGACCGGAAAGTTGGAGGTGAAATCCCAGGGCTCAACCTTGGAA CTGCCTTCAAAACTATCGGTCTGGAGTTCGAGAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGG AGGAACACCAGTGGCGAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAG ATACCCTGGTAGTCCACGCCGTAAACGATGAATGCCAGTCGTCGGGTAGCATGCTATTCGGTGACACACCTAACGGATTAA GCATTCCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTG GTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGACATTACAGGACATCCCCAGAGATGGGGCTTTCACTTCGGTG ACCTGTGGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTCGGTTAAGTCCGGCAACGAGCGCAACCCACA CTTCCAGTTGCCAGCATTCAGTTGGGCACTCTGGAAGAACTGCCGATGATAAGTCGGAGGAAGGTGTGGATGACGTCAAGT CCTCATGGCCCTTACGGGTTGGGCTACACACGTGCTACAATGGTGGTGACAGTGGGCAGCGAGACCGCGAGGTCGAG GGATTGGGGTCTGCAACTCGACCCCATGAAGTTGGAATCGCTAGTAATCGCGGAACAGCATGCCGCGGTGAATACGTTCCC GGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGGTCTACCCGACGGCCGTGCGCTAACCCTTACGGGAGGCAGCG GACCACGGTAGGCTCAGCGACTGGGGTGAAGTCGTAACAAGGTAACCGTA	Phylum: Proteobacteria Class: Alphaproteobacteria Order: Rhodobacterales Family: Rhodobacteraceae (Sp. Nov.)
20	*Ensifer (=**Sinorhizobium**)nitrofacere* strain ENF4	MF564192.1	>AGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAGCGCCCCGCAAGGGGAGCGGCA GACGGGTGAGTAACGCGTGGGAACGTACCCTTTGCTACGGAATAACTCCGGGAAACTGGAACTAATACCGTATGTGCCCTT CGGGGGAAAGATTTATCGGCAAAGGATCGGCCCCGCGTTGGATTAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCGACGAT CCATAGCTGGTCTGAGAGGATGATCAGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTGGGGA ATATTGGACAATGGGCGCAAGCCTGATCCAGCCATGCCGCGTGAGTGATGAAGGCCCTAGGGTTGTAAAGCTCTTTCACCG ATGAAGATAATGACGGTAGTCGGAGAAGAAGCCCCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGGCTAGC GTTGTTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGGTATTTAAGTCAGGGGTGAAATCCCGGAGCTCAACTCCGGAA CTGCCTTTGATACTGGGTACCTAGAGTATGGAAGAGGTAAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGG AGGAACACCAGTGGCGAAGGCGGCTTACTGGTCCATTACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAG ATACCCTGGTAGTCCACGCCGTAAACGATGAATGTTAGCCGTCGGGCAGTTTACTGTTCGGTGGCGCAGCTAACGCATTAAA CATTCCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGT TTAATTCGAAGCAACGCGCAGAACCTTACCAGCCCTTGACATCCCGATCGCGGACAGTGGAGACATTGTCCTTCAGTTCGGC TGGATCGGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGCGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCC TCGCCCTTAGTTGCCAGCATTCAGTTGGGCACTCTAAGGGACTGCCGGTGATAAGCCGAGAGGAAGGTGGGGATGACGTC AAGTCCTCATGGCCCTTACGGGCTGGGCTACACACGTGCTACAATGGTGGTGACAGTGGGCAGCGAGACCGCGAGGTCGAG CTAATCTCCAAAAGCCATCTCAGTTCGGATTGCACTCTGCAACTCGAGTGCATGAAGTTGGAATCGCTAGTAATCGCGGATC AGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGTTTTACCCGAAGGTAGTG TGCTAACCGCAAGGAGGCAGCTAACCACGGTAGGGTCAGCGACTGGGGTGAAGTCGTAACAAGGTAACCGTA	Phylum: Proteobacteria Class: Alphaproteobacteria Order: Rhizobiales Family: Rhizobiaceae (Sp. Nov.)
From coelomic fluid				
21	*Bacillus megaterium* strain Ah4	KU23052727	>AGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGAACTGATTAGAAGCTTGCTTCTAT GACGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCTGTAAGACTGGGATAACTTCGGGAAACCGAAGCTAATACCGG ATAGGATCTTCTCCTTCATGGGAGATGTTTGAAAGATGGTTTCGGCTATCACTTACAGATGGGCCCGCGGTGCATTAGCTAGTTGGT GAGGTAACGGCTCACCAAGGCAACGATGCATAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCAGACTC CTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAAGGCTTTCGGGT CGTAAAACTCTGTTGTTAGGGAAGAACAAGTACGAGAGTAACTGCTCGTACCTTGACGGTACCTAACCAGAAAGCCACGGCTAACT ACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTACCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTTCTTAAGT CTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGGAACTTGAGTGCAGAAGAGGAGAGTGGAATTCCACG TGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTTTTGGTCTGTAACTGACGCTGAGGTGCGAAA GCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAGAGGGTTTCCGCCCTTTAG TGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTCGCAAGACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAA GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCTCTGACAACTCTAGAGATAGAGCGT TCCCCTTCGGGGGACAGAGTGACAGGCGTTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG CAACCCTTGATCTTAGTTGCCAGCATTTAGTTGGGACTCTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTC AAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGTATGGTACAAAGGGCTGCAAGACCGCGAGGTCAAGCCAAT CCCATAAAACCATTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCC GCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGGAGTAACCGTAA GGAGCTAGCCGCCTAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAGGTAACCGTA	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae
22	*Bacillus pumilus* strain BP	KY616173.1	>AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGAACTGAAGGGAGCTTGCTCCCGGA AGTTAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGAGCTAATACCGGAT ATTTTCTTGAACCGCATGGTTCAAGAATGAAAGACGGTTTCGGCTGTCACTTACAGATGGACCCGCGGCGCATTAGCTAGTTGGTGA GGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCAGACTCCT ACGGGAGGCAGCAATAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAAGGTTTTCGGATCG TAAAGCTCTGTTGTTAGGGAAGAACAAGTACGAGAGTAACTGCTCGTACCTTGACGGTACCTAACCAGAAAGCCACGGCTAACTAC GTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGGGCTCGCAGGCGGTTTCTTAAGTCT GATGTGAAAGCCCCCGGCTCAACCGGGGAGGGTCATTGGAAACTGGGGAACTTGAGTGCAGAAGAGGAGAGTGGAATTCCACGTG TAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGACTCTCTGGTCTGTAACTGACGCTGAGGAGCGAAAGC GTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAGGGGGTTTCCGCCCCTTAGTG CTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTCGCAAGACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGC GGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCTCTGACAACCCTAGAGATAGGGCTTTC CCTTCGGGGACAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAAC CCTTGATCTTAGTTGCCAGCATTCAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAAT CATCATCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACAGAACAAAGGGCTGCGAGACCGCAAGGTTTAGCCAATCCCAC AAATCTGTTCTCAGTTCGGATCGCAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGT GAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGCAACACCCGAAGTCGGTGAGGTAACCTTTATGGAG CCAGCCGCCGAAGGTGGGGTAGATGATTGGGGTGAAGTCGTAACAAGGTAACCGTA	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae
23	*Bacillus cereus* strain BCR	MF3983	>AGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGCGTGCCTAACACATGCAAGTCGAGCGAATGGATTGAGAGCTTGCTCTCAA GAAGTTAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCCATAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGG	Phylum: Firmicutes

		97	ATAACATTTTGAACTGCATGGTTCGAAATTGAAAGGCGGCTTCGGCTGTCACTTATGGATGGACCCGCGTCGCATTAGCTAGTTGGT GAGGTAACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCAGACTC CTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAAGGCTTTCGGGT CGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATAAGCTGGCACCTTGACGGTACCTAACCAGAAAGCCACGGCTAAC TACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATTATTGGGCGTAAAGCGCGCGCAGGTGGTTTCTTAAG TCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTGAGTGCAGAAGAGGAAAGTGGAATTCCAT GTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGTCTGTAACTGTCACTGAGGCGCGAA AGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAGAGGGTTTCCGCCCTTTA GTGCTGAAGTTAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGGCCCGCACA AGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCTCTGAAAACCCTAGAGATAGGGCT TCTCCTTCGGGAGCAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGC AACCCTTGATCTTAGTTGCCATCATTAAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCA AATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACGGTACAAAGAGCTGCAAGACCGCGAGGTGGAGCTAATC TCATAAAACCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCG CGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGGGGTAACCTTTTG GAGCCAGCCGCCTAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAGGTAACCGTA	Class: Bacilli Order: Bacillales Family: Bacillaceae
24	*Bacillus thuringiensis* strain BT	MF5743 66	>TATGAAGTTAACGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCCATAGACTGGGATAACTCCGGGAAACCGGGGCTAATAC CGGATAATATTTTGAACTGCATGGTTCAAAATTGAAAGGCGGCTTCGGCTGTCACTTATGGATGGACCCGCGTCGCATTAGCTAGTT GGTGAGGTAACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCCAGA CTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCGTGAGTGATGAAGGCTTTCG GGTCGTAAAACTCTGTTGTTAGGGAAGAACAAGTGCTAGTTGAATAAGCTGGCACCTTGACGGTACCTAACCAGAAAGCCACGGCT AACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTATCCGGAATTATTGGGCGTAAAGCACGCGCAGGTGGTTTCTT AAGTCTGATGTGAAAGCCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGGGAGACTTGAGTGCAGAAGAGGAAAGTGGAATTC CATGTGTAGCGGTGAAATGCGTAGAGATATGGAGGAACACCAGTGGCGAAGGCGACTTTCTGGTCTGTAACTGACACTGAGGCGC GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAAGTGTTAGAGGGTTTCCGCCCT TTAGTGCTGAAGTTAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGGCCCGC ACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCTCTGAAAACCCTAGAGATAGG GCTTCTCCTTCGGGAGCAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAG CGCAACCCTTGATCTTAGTTGCCATCATTAAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACG TCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACGGTACAAAGAGCTGCAAGACCGCGAGGTGGAGCTA ATCTCATAAAACCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGC CGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGGGGTAACCTTT TTGGAGCCAGCCGCCTAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAGGTAACCGTAA	Phylum: Firmicutes Class: Bacilli Order: Bacillales Family: Bacillaceae
25	*Escherichia coli* strain AG2	MH236 118.1	>AGAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAACGGTAACAGAAAGCAGCTTGCTGCTT TGCTGACGAGTGGCGGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTACTGGAAACGGTAGCTAATACCG CATAACGTCGCAAGACCAAAGAGGGGGACCTTCGGGCCTCTTGCCATCGGATGTGCCCAGATGGGATTAGCTTGTTGGTGGGGTAA CGGCTCACCAAGGCGACGATCCCTAGCTGGTCTGAGAGGATGACCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGG AGGCAGCAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCCATGCCGCGTGTATGAAGAAGGCCTTCGGGTTGTAAAG TACTTTCAGCGGGGAGGAAGGGAGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAGAAGAAGCACCGGCTAACTCCGTGCC AGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATG TGAAATCCCCGGGCTCAACCTGGGAACTGCATCTGATACTGGCAAGCTTGAGTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCG GTGAAATGCGTAGAGATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGACGAAGACTGACGCTCAGGTGCGAAAGCGTGGG GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTTCCGGAG CTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGA GCATGTGGTTTAATTCGATGCAACGCGAAGAACCTTACCTGGTCTTGACATCCACAGAACTTTCCAGAGATGGATTGGTGCCTTCGG GAACTGTGAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTTGTGAAGTGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATC CTTTGTTGCCAGCGGTCCGGCCGGGAACTCAAAGGAGACTGCCAGTGATAAACTGGAGGAAGGTGGGGATGACGTCAAGTCATCA TGGCCCTTACGACCAGGGCTACACACGTGCTACAATGGCGCATACAAAGAGAAGCGACCTCGCGAGAGCAAGCGGACCTCATAAA GTGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTAGTAATCGTGGATCAGAATGCCACGGTGAA TACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCGGGAGGGCG CTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAACCGTA	Phylum: Proteobacteria Class: Gammaproteobacteria Order: Enterobacteriales Family: Enterobacteriaceae

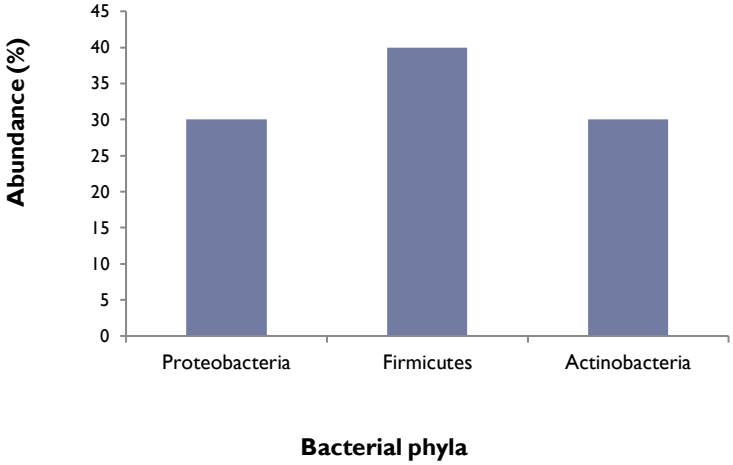


Fig 1.12. Major culturable bacteria in the gut of *E. fetida*

1.3.12. 1. Phylogenetic trese based on 16S rRNA gene sequences

Phylogenetic trees were constructed using MEGA vev7.0. For the strains ET03, EPG1, EAG2 and EAG3 detailed phylogenetic analysis have benn performed using the same protocol.

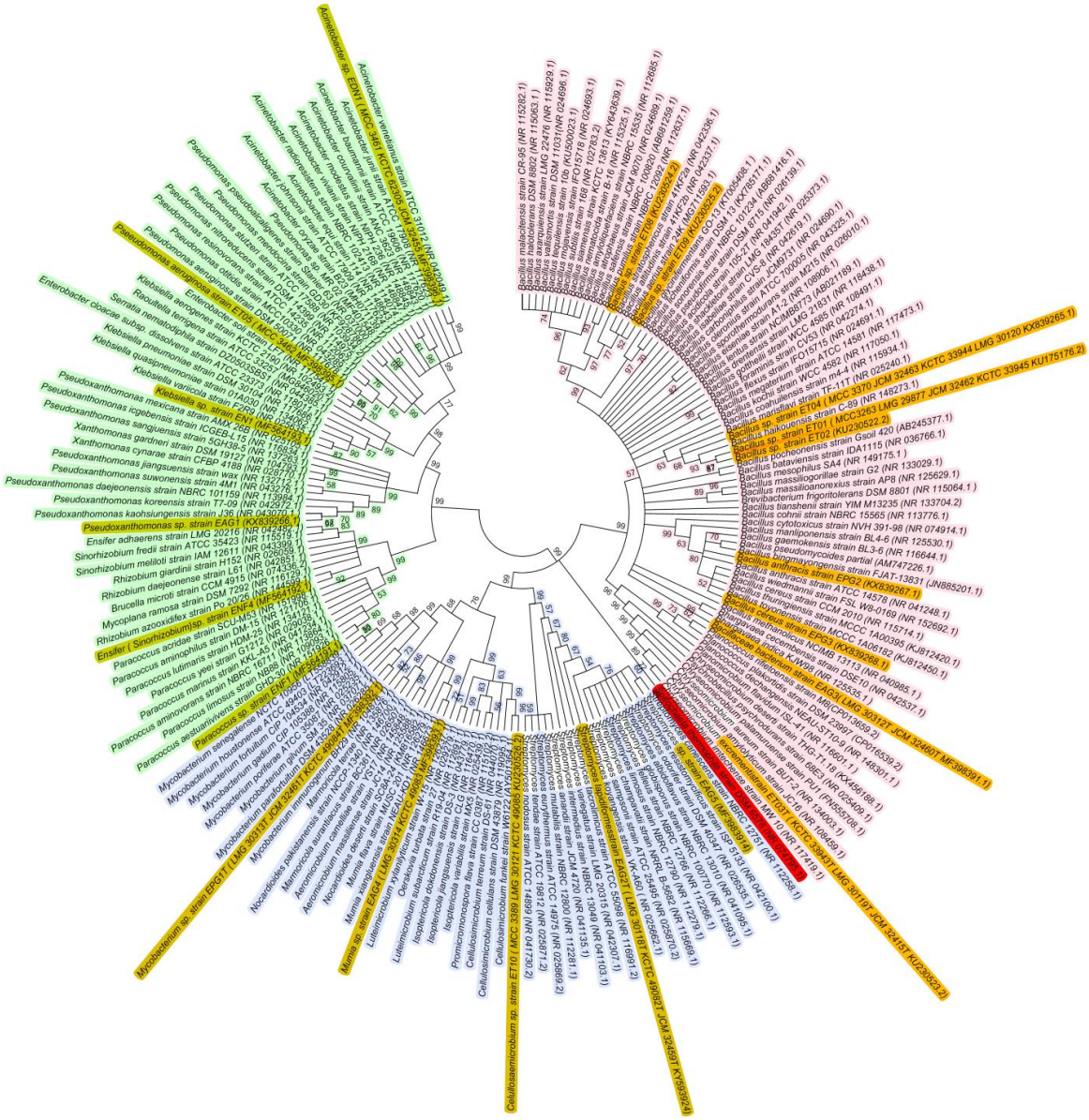


Figure.1.13- Phylogeny of the 20 isolates from the gut of *E. fetida* representing cultivable diversity from the phyla *Firmicutes* (pink), *Actinobacteria* (blue) and *Proteobacteria* (green).
The evolutionary history was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches less than 50% bootstrap replicates have been collapsed. The evolutionary distances were computed using the Kimura 2-parameter and are in the units of the number of base substitutions per site. This analysis involved 185 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1152 positions in the final dataset. This evolutionary analyses were conducted in MEGA7. Firmicutes [ ], Actinobacteria [ ] and proteobacteria [ ] are the major bacterial phyla. The highlighted (Yellow) lines represent the 20 isolates; the red line dictates spirochaetes which was not found in the cultivable colonies.

According to the comparison of 16S rRNA gene sequences (using NCBI, Blastn), the closest relative of strain ET03^T were *C. palamuruens* strain PU1^T (99.1%), *C. aureum* strain BUT-2 (99%), *Psychrobacillus psychrodurans* strain BAB-2243 (99%), *C. amylolyticum* strain ID4 (98.9%), *C. imtechense* strain HWG-A7 (98.4%), *C.deserti* strain THG-T1 (96.3%), *Planococcus rifietoensis* strain M8 (95.7%), *Planococcus plakortidis* strain DSM 23997 (95.6%) sequentially followed by other members of the Family *Planococcaceae*. A tree depicting the phylogenetic position of strain ET03^T within the genus *Chryseomicrobium* is shown in Fig. 1.14. Based on 16S rRNA gene sequence comparison, strain ET03^T forms a distinct subclade with *C. palamuruens* PU1^T and other members of the genus *Chryseomicrobium*.

According to the comparison of 16S rRNA gene sequences, the closest relatives of strain EPG1^T were *Mycobacterium parafortuitum* strain DSM 43528^T (99.3% similarity), *Mycobacterium houstonense* strain ATCC 49403^T (99.0%), *Mycobacterium gilvum* strain PYR-GCK (98.9%). A tree depicting the phylogenetic position of strain EPG1^T within the family *Mycobacteriaceae* is shown in Fig. 1.15. Based on 16S rRNA gene sequence comparison, strain EPG1^T forms a distinct clade parallel to *M. parafortuitum.*

According to the comparison of 16S rRNA gene sequences, the closest relative of strain EAG2^T was *Streptomyces koyangensis* VK-A60^T (=KCCM 10555^T=NBRC 100598^T), showing 99.7 % sequence similarity. A tree depicting the phylogenetic position of strain EAG2^T within the genus *Streptomyces* is shown in Fig. 1.16. The strain EAG2^T forms a phylogenic line (clade) distant from the three clusters of *S. intermedius*, *S. koyangensis*, and *S. odorifer* in the tree. The position of the strain EAG2^T in the phylogenetic tree was unaffected by either the tree-making algorithm or the outgroup strains used. These findings suggest that strain EAG2^T represents a novel species that is closely related to *S. koyangensis*, *S. intermedius*, and *S. odorifer*, having a high 16S rRNA gene sequence similarity. The designation of the strain EAG2^T as a separate genomic species was suggested by the bootstrap value (85%) in the N-J tree based on nearly complete 16S rRNA gene sequence data.

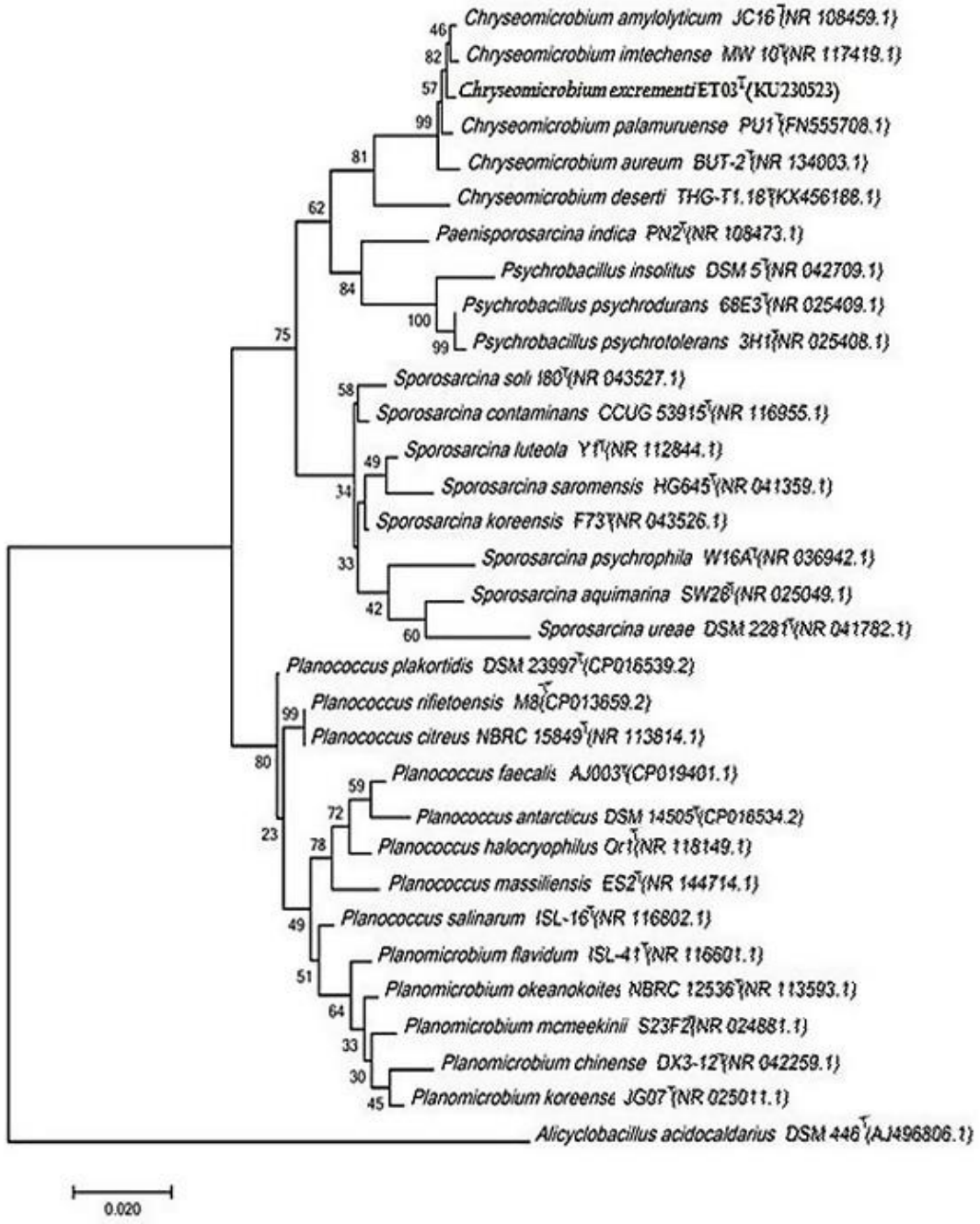


Fig. 1.14: Phylogenetic tree constructed by the neighbour-joining method based on 16S rRNA gene sequences showing the phylogenetic relationship between strain ET03^T and closely related species. Bootstrap percentages (based on 1000 replications) are shown at the nodes. The GenBank accession numbers for 16S rRNA gene sequences are shown in parentheses. *Alicyclobacillus acidocaldarius* strain DSM 446 (AJ496806.1) was used as the outgroup. Bar, 2 nt substitution per 100 nt.

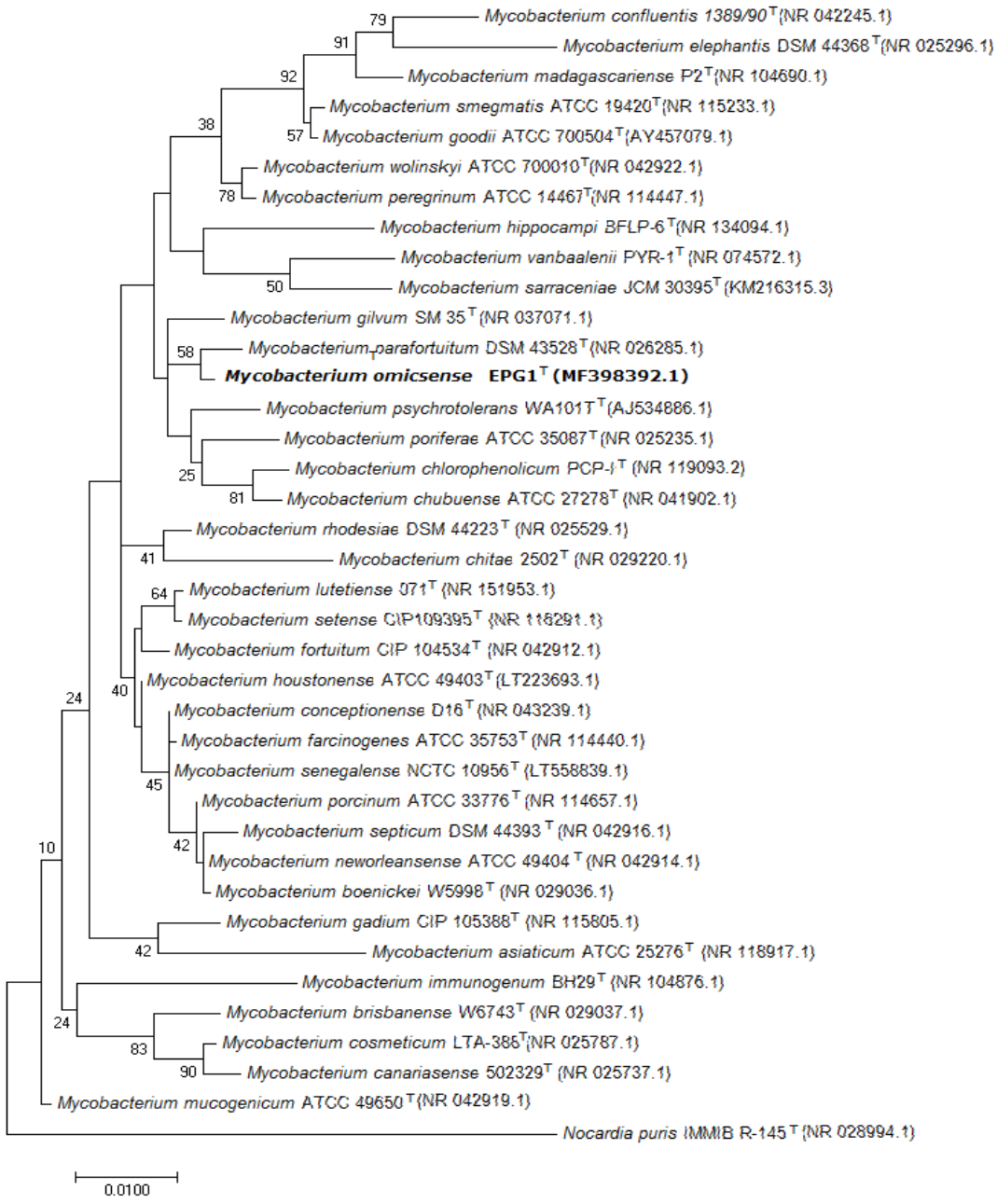


Fig. 1.15: Molecular Phylogenetic analysis of EPG1^T by Maximum Likelihood method

The evolutionary history was inferred by using the Maximum Likelihood method based on the Jukes-Cantor model. The tree with the highest log likelihood (-5467.4312) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 38 nucleotide sequences. There were a total of 1537 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

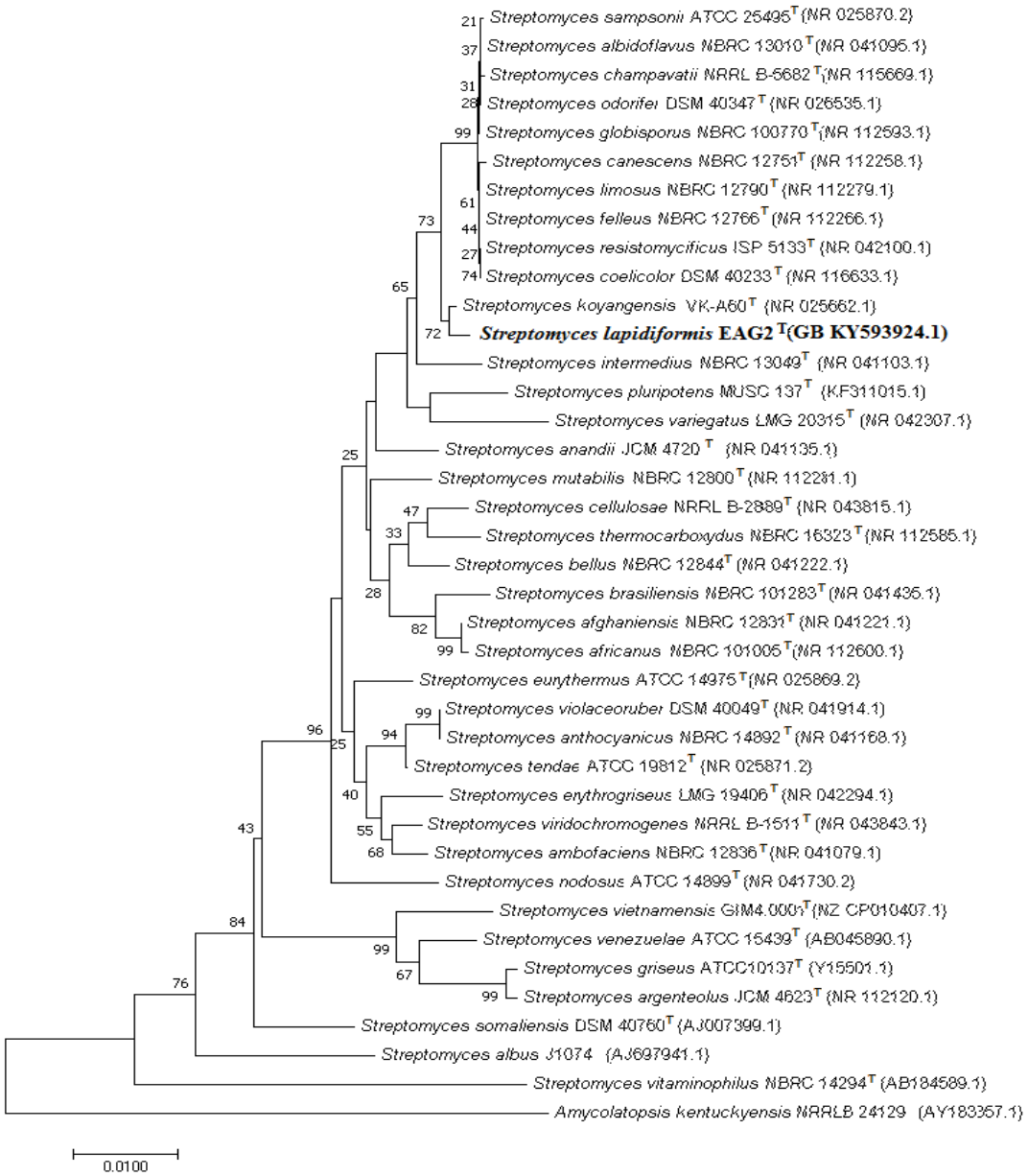


Fig. 1.16:. Evolutionary relationships of strain EAG2^{T}.

The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 0.31523288 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown (>20) next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. The analysis involved 39 nucleotide sequences. All ambiguous positions were removed for each sequence pair. There were a total of 1557 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

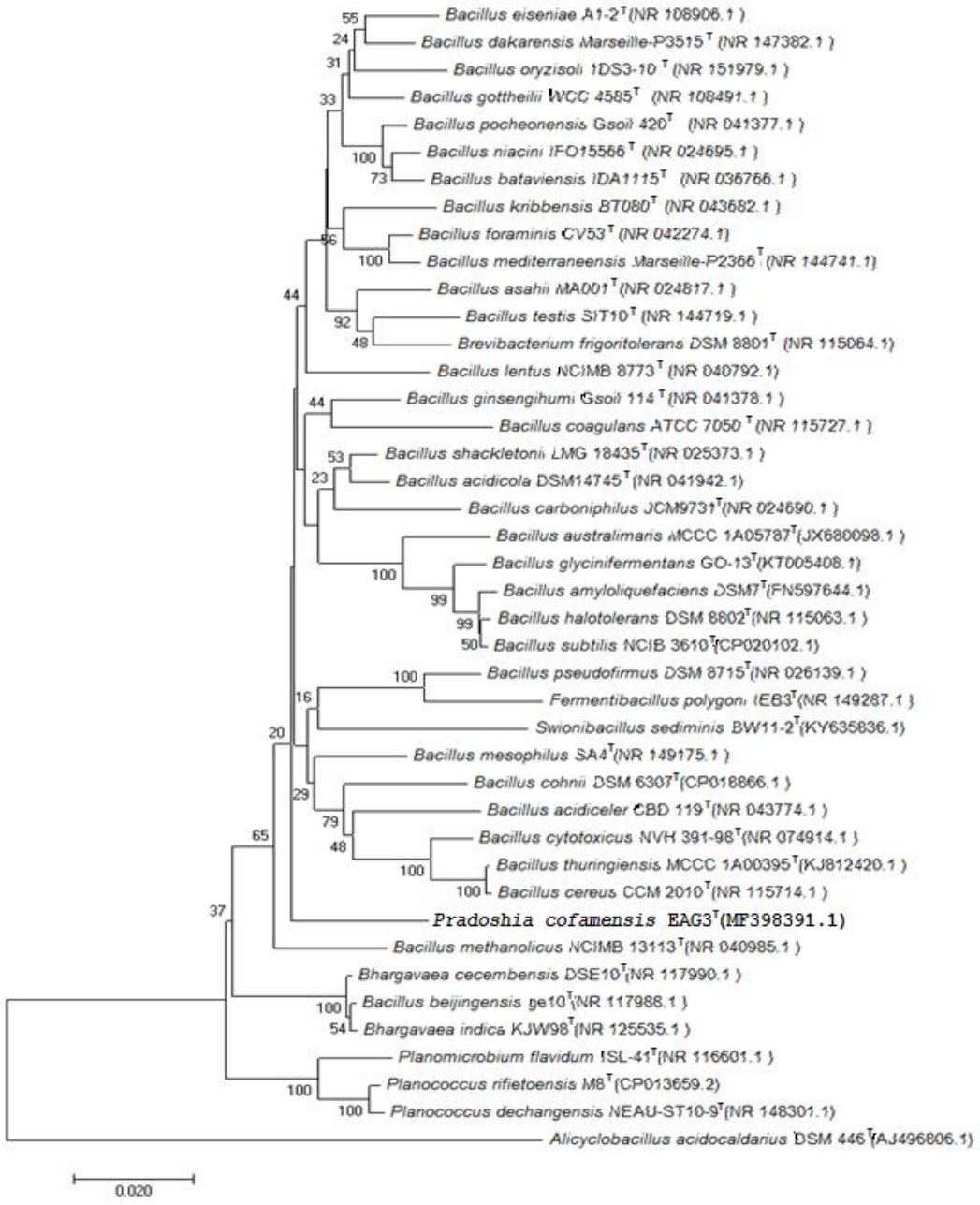


Fig. 1.17: Evolutionary relationships of strain EAG3^T. The evolutionary history was inferred using the Neighbor-Joining method . The optimal tree with the sum of branch length = 0.74518028 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Jukes-Cantor method and are in the units of the number of base substitutions per site. The analysis involved 42 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1414 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

According to the comparison of 16S rRNA gene sequences, the close relatives of strain EAG3^T from different genus of family *Bacillacae* were *Bacillus methanolicus* NCIMB 13113^T (95.7 % similarity), *Planococcus rifietoensis* M8^T (94.2% similarity), *Bhargavaea cecembensis* strain DSE10^T (94.2% similarity) *Planomicrobium flavidum* strain ISL-41^T (93.7% similarity) *and Fermentibacillus polygoni* strain IEB3^T (93.3% similarity). The 16S rRNA gene sequence comparisons showed sufficient differences so that the strain could be allocated to a separate species without the need for DNA–DNA hybridization experiments (Yarza *et al.*, 2014; Kim *et al*., 2014). A tree depicting the phylogenetic position of strain EAG3^T within the family *Bacillacae* is shown in Fig. 1.17. Based on 16S rRNA gene sequence comparison, strain EAG3^T forms a distinct uniramous clade.

1.3.12. 2. Biochemical characteristics of the novel bacterial strains

Strain ET03^T cells are Gram stain positive, non-motile rods, measuring 1.5±0.5µm in length and 0.5µm in width (Fig. 18-C). The physiological and biochemical test results have been enlisted in 'Table 1' in comparison to other close taxonomic neighbours. ET03^T contained MK-8 as the most predominant menaquinone; MK-7 and MK-6 were also detected (Table: 1.2). The cell-wall peptidoglycan of strain ET03^T contained L-Alanine, D-Glutamic acid, and 2,6,-Diaminopimelic acid in the peptide stem.

Strain EPG1^T cells were short spindle or rod-shaped measuring 1.2-2.2 µm in length x 0.6-0.9 µm in diameter as found in scanning electron microscope (Fig. 18-N). EPG1^T was negative for catalase, and β-Galactosidase activity, but showed nitrate reduction, arylsulfatase (3 d), acid phosphatase, urease, pyrazinamidase, α-esterase and β-esterase activity. The detailed physiological and biochemical test results have been enlisted in 'Table: 1.3' in comparison to other close taxonomic neighbours. EPG1^T contained MK-7, MK-8, and MK-9 as the most predominant menaquinones. TLC of methanolysates shows α-, α′-, and keto-mycolates, and wax esters.

Table.1.2**:** Characteristics that differentiate strain ET03^T from other members of the genus *Chryseomicrobium*.

Strains: 1, ET03^T; 2, *Chryseomicrobium imtechense* MW 10^T; 3, *Chryseomicrobium amylolyticum* JC16^T; 4, *Chryseomicrobium aureum*BUT-2^T; 5, *Chryseomicrobium palamuruense* PU1^T; 6, *Chryseomicrobium deserti* THG-T1.1B^T.

+, Positive; –, negative. AL, unidentified aminolipids; DPG, diphosphatidylglycerol; GL, unidentified glycolipid; PE, phosphatidyl ethanolamine; PG, Phosphatidyl-glycerol; phosphatidyl inositol (PI), phosphatidyl inositolmannoside (PIM) and PL, unidentified phospholipids; m-DAP-meso-diaminopimelic acid; ND, Not detected.

		Strains					
Sl. No.	**Characters**	**1**	**2**	**3**	**4**	**5**	**6**
1	Cell shape and size (µm) (length x diameter)	Rods, 1.5-2.2 x 0.5-0.6	Rod, 1.7–2.9 x 0.3–0.7	Rods, 2.0–3.0 x 1.0	Rods, 1.5–2.0 x 0.5 – 0.86	Rods, 1.6–2.0 x 0.6–0.7	Rods, 2.4-2.7 x 0.5-0.7
2	Motility	non-motile	non-motile	non-motile	non-motile	motile	non-motile
3	Growth temp. range (optimum)	20–40 ℃ (35–37)	4–45 ℃	25–40 ℃ (30–37)	20-35℃	18 - 40 ℃	20–35 ℃ (28–30)
4	pH range (optimum)	6-9	6–9	7–11 (7–8)	7-10	7 to 10	5-7 (7)
5	NaCl tolerance limit (%, w/v)	8	6	5	7	9	3
6	Catalase reaction	-	-	+	-	+	+
7	nitrate reduction	+	-	-	-	-	-
8	VP reaction	-	+	-	-	-	-
9	Urease activity	-	-	-	+	+	+
10	Oxidase	-	-	-	-	+	-
	Hydrolysis of:						
11	Starch	-	-	+	-	+	-
12	Gelatin	+	-	-	+	-	- ?
	Organic substrates utilized for growth:						
13	Citrate	-	+	-	-	+	+
14	Glycerol	-	+	+	-	+	-
	Acid production from various carbohydrates						
15	Glucose	+	+	-	-	+	-
16	Salicin	+	+	-	+	-	-
17	Mannose	-	-	+	-	+	+
18	Fructose	+	+	-	-	-	-
19	Maltose	+	+	-	-	+	+
20	Sucrose	-	+	-	-	-	+
21	Inulin	-	+	-	+	-	-
22	Trehalose	-	-	+	-	-	+
23	Melibiose	-	+	-	-	+	+
24	Cellobiose	-	+	-	-	+	-
25	**Menaquinones** (Descending abundance)	MK-8, 6, 7	MK7, 8, 7_{H2}, 6	MK7, 8, 6	MK7, 6, 8	MK8	MK-7, 8, 6
26	**Polar lipids**	PE, PI, PIM , PG	DPG, PG, PE, PC, GL	DPG, PG, PE, AL, PL,	DPG, PG, PE, PL	DPG, PG, PE	DPG, PE, PG, GL, AL
27	**Peptidoglycan types**	L-Ala-D-Glu	L-Lys-D-Asp	L-Orn-D-Glu	L-Orn-D-Glu	m-DAP	L-Orn-D-Glu
28	**DNA G+C content** (mol%)	42.9	53.4	57.6	48.5	48.5	50.4

Table 1.3: Characteristics that distinguish strain EPG1^T from other closely related Mycobacteria.
Strains: 1, EPG1^T; 2, *M. parafortuitum* DSM 43528^T*; 3, M. gilvum* PYR-GCK*; 4, M. houstonense* ATCC 49403^T and *5, M. fortuitum* CT6.
+, Positive; –, negative, NK- Not Known. aSymbols: +, 90% or more of strains are positive; −, 90% or more of strains are negative; d, 11–89% of strains are positive; v, variability of reaction; w, weak reaction; N, non- chromogenic; S, scotochromogenic; P, photochromogenic. nd, data not available.

No.	Characters	1	2	3	4	5
Cell morphology-						
1	Cell shape, size (µm) (length x diameter)	Short spindle or rods 1.2-2.2 x 0.6-0.9	Short - medium Rods, 2-3 x 1	Cocci 4 x 3.5	Long filamentous	Cocci/ rod 0.4-1.6 x 0.4-0.8
2	Pigmentation	P	P	S	N	N
Culture condition-						
3	Growth at 42°C	−	v	nd	+	v
4	Growth on 5% (w/v) NaCl	−	nd	nd	+	+
Enzyme activity-						
5	Nitrate reduction	+	d	+	+	+
6	Arylsulfatase (3 d)	+	−	+	+	+
7	Acid phosphatase	+	−	nd	Nd	+
8	Catalase	-	nd	+	+	nd
9	Urease	+	+	nd	+	+
10	Pyrazinamidase	+	nd	nd	Nd	d
11	α-Esterase	+	−	nd	Nd	d
12	β-Esterase	+	v	nd	Nd	d
13	β-Galactosidase	-	nd	−	Nd	−
Utilization of :						
14	Acetamide	V	nd	nd	+	d
15	Citrate	+	+	−	−	−
16	Xylose	+	v	nd	-	−
17	Mannitol	-	+	+	+	d
18	Sorbitol	-	nd	nd	+	-
19	Trehalose	-	nd	nd	+	d
Acid from:						
20	l-Arabinose	-	+	nd	-	nd
21	d-Mannitol	-	+	+	+	−
22	l-Rhamnose	-	v	−	-	−
23	d-Sorbitol	-	−	+	+	−
24	**Menaquinones** (Descending abundance)	MK-7, 8, 9	MK7, 8, 7_{H2}, 6	MK7, 8, 6	MK7, 6, 8	MK8
25	**Polar lipids**	PG, DPG, PE	DPG, PG, PE, PC, GL	DPG, PG, PE, AL, PL,	DPG, PG, PE, PL	DPG, PG, PE
26	Mycolates	α-, α′-, keto and wax ester	α-, α′-, keto and wax ester	α-, α′-, keto and wax ester	54-60 carbon mycolic acid	α-, α′-, and epoxy
27	**DNA G+C content** mol% (method)	68.3(WGS)	68.5(WGS)	67.9 (WGS)	64 (T_m)	66.2(WGS)

Table 1.4: Characteristics that distinguish strain EAG2^T from other closely related species of the genus *Streptomyces*.
Strains: 1, EAG2^T; 2, *Streptomyces koyangensis* VK-A60^T; 3, *Streptomyces sampsonii* ATCC25495^T; 4, *Streptomyces albus* ATCC 25426^T. +, Positive; –, negative, NK- Not Known.

Sl. No.	Characters	1	2	3	4
Colony morphology-					
1	Aerial mass colour	White	White/Yellow	Yellow/Grey	White/Yellow
2	Colour of substrate mycellium	Grey	Brown	Yellow/Orange/ Brown	colorless/pale yellow
3	Spore chain morphology	Closed spiral	Rectiflexibile	Rectiflexibile	Spiral
4	Number of spores per chain	10-30	10–50 or more	>50	10-50
5	NaCl tolerance (%, w/v)	7	10	7	7
6	Melanin production	-	+	-	-
7	nitrate reduction	+	+	ND	-
8	Urease activity	+	ND	ND	+
Hydrolysis of:					
9	Starch	+	+	+	+
10	Aesculine	Weakly +	+	+	+
11	Gelatin	+	+	+	+
Organic substrates utilized for growth:					
12	D-Glucose	+	+	+	+
13	Arabinose	+	+	+	Variable
14	Fructose	+	-	+	+
15	Rhamnose	+	-	-	-
16	Sucrose	-	-	-	-
17	DNA G+C content (mol%)	73.1	67.8	NK	NK
18	Fatty acid	$C_{14:0}$ 0.93 C14:0 iso 1.80 $C_{15:0}$ iso 13.38 C15:0 anteiso 51.83 C16:0 3.89 C16:0 iso 1.10 C17:0 iso 1.29 C17:0 anteiso 3.76	$C_{14:0}$ 1.33 C14:0 iso 8.84 C15:0 iso 7.02 C15:0 anteiso 16.54 C16:0 11.60 C16:0 iso 28.77 C17:0 iso 1.94 C17:0 anteiso 9.01	NK	NK

Table 1.5: Characteristics that distinguish strain EAG3^T from other closely related species of the genus Fam. *Bacillaceae*.

Strains: 1, EAG3^T; 2, *Bacillus methanolicus* NCIMB 13113^T; 3, *Planococcus rifietoensis* M8^T; 4, *Bhargavaea cecembensis* DSE10^T; 5, *Planomicrobium flavidum* ISL-41^T and 6, *Fermentibacillus polygoni* IEB3^T. +, Positive; –, negative, NK- Not Known.

		Strains					
Sl. No.	**Characters**	**1**	**2**	**3**	**4**	**5**	**6**
Cell morphology-							
1	Cell shape, size (µm) (length x diameter)	short rods 1.2-1.7 x 0.6-0.8	Rods 4-6 x >1	Cocci 4 x 3.5	Rods 2 x 0.6	Cocci/ rod 0.4-1.6 x 0.4-0.8	Rods 0.6–1.0 x 1.3–4.5
2	Motility	Motile	Motile	Non-motile	Non-motile	Motile	Motile
3	Sporulation	spore forming	Spore forming	No spore	No spore	No spore	Spore forming
Culture condition-							
4	Growth temp(°C) (optimum)	**4-42** (28)	35–60 (55)	30-42 (37)	15-55 (37)	4-37 (30)	12-40 (32)
5	pH range (optimum)	7-10 (7-8)	7-10	6-10	(7-7.5)	6-9 (7-8)	7.5-12 (9-10)
6	NaCl tolerance (%, w/v)	4	2	15	6	13	5
Enzyme activity-							
7	Catalase	+	+	+	+	+	+
8	Oxidase	-	+	+	+	+	+
9	nitrate reduction	+	-	-	+	-	-
10	Urease	-	-	-	+	-	-
Fermentation of-							
11	Fructose	+	-	-	-	+	-
12	Inulin	+	-	-	-	NK	-
13	Lactose	-	-	-	-	-	-
14	Maltose	+	+	-	-	-	-
15	Mannitol	-	+	-	-	-	-
16	Raffinose	-	+	-	-	-	-
17	Ribose	-	+	NK	-	-	+
18	Salicin	+	-	-	-	NK	NK
19	Sorbitol	-	+	-	-	-	-
20	Sucrose	+	-	-	-	-	-
21	Trehalose	-	-	-	-	-	-
Hydrolysis of-							
22	Esculine	-	+	-	+	-	+
23	Casein	-	-	-	+	+	-
24	Gelatin	-	ND	+	-	-	+
25	Starch	-	-	-	+	-	-
26	**G+C (mol%)**	41.82	48-50	**NK**	59.5	45.9	49.1
27	**Major isoprenoid quinones**	MK-7, MK-8	MK-7	Mk-7, Mk-8	MK-6, MK-8	MK-6, MK7, MK8	MK-7

Table 1.6: Cellular fatty acid profile (indicated numerically as percentages) of novel isolate EAG3^T in comparison to other closely related genus of the Fam. *Bacillaceae*.
Strains: 1, EAG3^T; 2, *Bacillus methanolicus* NCIMB 13113^T; 3, *Planococcus rifietoensis* M8^T; 4, *Bhargavaea cecembensis* DSE10^T; 5, *Planomicrobium flavidum* ISL-41^T and 6, *Fermentibacillus polygoni* IEB3^T.
ND- Not detected; '-' indicate 'data unavailable/not reported'.

*Summed feature 4 represents iso-$C_{17:1}$ I and/or anteiso/iso-$C_{17:1}$ B, which could not be separated by GC with the MIDI-Sherlock
Identification system.

Fatty acid	**Bacterial strains**					
	1	**2**	**3**	**4**	**5**	**6**
Branched chain						
iso-C14 : 0	7.5	-	2.8	8.5	8	10.2
iso-C15 : 0	30.7	27	6.9	12.8	1.8	12.2
anteiso-C15 : 0	12.9	16	37	31.2	39	47.1
iso-C16 : 0	11.5	12	7.6	10.7	11.5	7.8
iso-C17 : 0	2.6	4	4.2	-	2.8	2.9
anteiso-C17 : 0	5	14	8.6	6.2	11.3	2.4
Straight chain						
C14 : 0	0.6	-	0.6	-	-	1.8
C15 : 0	ND	-	2.5	-	-	8.1
C16 : 0	1.5	-	6.8	4.7	-	4.9
C17 : 0	0.3	-	4.4	-	-	0.8
C18 : 0	0.6	-	5.9	-	-	-
Unsaturated						
C16 : 1ω7c OH	13.7	-	4	4.5	11	-
C16 : 1 ω 11c	1.6	-	0.8	-	1	-
Summed feature 4*	4.2	-	1.9	-	8.4	-

The growth of EAG2^T colonies on LA is lichenoid, hard (stony), densely textured, chalky, raised, adhering to the medium and formed clear concentric rings with time. In liquid media, especially in shaken culture, the growth of EAG2^T is in the form of spherical growths or puffballs. Strain EAG2^T can be readily differentiated from the closest relatives, *Streptomyces koyangensis* VK- A60^T *Streptomyces albus* ATCC 25426^T and *Streptomyces sampsonii* ATCC25495^T with reference to physiological and biochemical characteristics including spore pattern, production of melanin, cellular

fatty acids and DNA G+C content (Table: 1.4). EAG2^T contained MK-8(60%) and MK-9(40%) as the most predominant menaquinones.

Strain EAG3^T cells are short bacilli as found in scanning electron microscope (Fig. 1.18-K). The physiological and biochemical test results have been enlisted in 'Table 5' in comparison to other close taxonomic neighbours. EAG3^T contained MK-7 and MK-8 as the major isoprenoid quinones. MK-6 and MK-7H_2 are also present in small amount.

1.3.12. 3. Determination of Polar lipid, FAME and GC mol% of the strains

Phosphatidyl ethanolamine (PE), phosphatidyl inositol (PI) and phosphatidyl inositolmannoside (PIM) were identified as the phosplolipids present in strain ET03^T along with at least three spots of unknown lipids. Phosphatidyl glycerol (PG) was characteristically found in strain ET03^T as the primary glycolipid.

Major cellular fatty acids were 13-Methyltetradecanoic acid or iso-15:0 (45%), (9Z)-9-Hexadecenoic acid or 16:1 ω7c alcohol (13%), 14-Methylpentadecanoic acid or iso-16:0 (11.8%) and 12-Methyltridecanoic acid or iso-14:0 (6.2%). Trace amount of 12-Methyltetradecanoic acid or anteiso-15: 0 (3.4%), 15-Methylhexadecanoic acid or iso-17: 0 (2.2%) and 14-Methylhexadecanoic acid or anteiso-17: 0 (1.6%) were present which is typical of members of the genus *Chryseomicrobium*, but the proportions differed from those reported for other members of the genus.

The predominant polar lipids were phosphatidylglycerol (PG), diphosphatidylglycerol (DPG), phosphatidylinnocitol (PI) and phosphatidylinnocitol mannoside (PIM). Major cellular fatty acids were hexadecanoic acid or $C_{16:0}$ (22.9%), (9Z)-9-octadecenoic acid or $C_{18:1}$ ω9c (21.1%) and 10-methyloctadecanoic acid or $C_{18:0}$ 10-methyl (11.5%). Small amount of tetradecanoic acid or $C_{14:0}$ (5.4%) and Octadecanoic acid or $C_{18:0}$ (1.9%) were also present. Fatty acids of strain EPG1^T were typical of other members of the Genus *Mycobacterium*, but the proportions differed.

Phosphatidyl-ethanolamine (PE), and phosphatidyl-inositol (PI) were identified as the major phosplolipids present in strain EAG2^T. Trace amount of phosphatidylinositol mannoside (PIM), and unidentified phospholipid (PL) were also detected. Major cellular fatty acids were 12-methyltetradecanoic acid (C15:0 anteiso), 13-Methyltetradecanoic acid (C15:0 iso), Hexadecanoic acid (C16:0) and 14-Methylhexadecanoic acid (C17:0

anteiso). The major fatty acid groups found in EAG2^T are enlisted in Table1 which differed from those reported for other members of the genus and showed best similarity index of 0.386 supporting claim of novel species

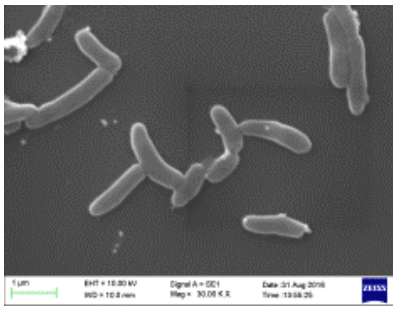
The predominant polar lipids present in strain EAG3^T were phosphatidylglycerol (PG), diphosphatidylglycerol (DPG) and phosphatidylethanolamine (PE). Major cellular fatty acids were 12-Methyltridecanoic acid ($C_{14:0}$ iso 7.5%), 13-Methyltetradecanoic acid ($C_{15:0}$ iso 30.7%), 12-Methyltetradecanoic acid (C15:0 anteiso 12.9%), 14-Methylpentadecanoic acid (C16:0iso 11.5%), (9Z)-9-Hexadecen-1-ol (16:1 ω7c alcohol 13.7%), 14-Methylhexadecanoic acid (C17:0 anteiso 5%) and 15-Methylhexadecanoic acid (C17:0 iso 2.6%). Fatty acids of strain EAG3^T were typical of other members of the family *Bacillacae*, but the proportions differed (Table: 1.6).

1.3.12 4. Scanning Electron Microscopy (SEM) of the strains.

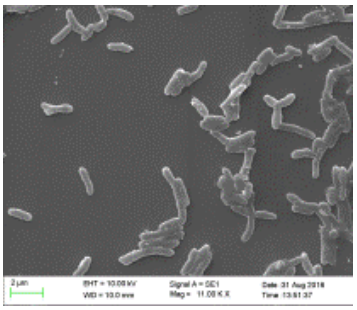
Details of the cell shape of the strains were ascertained with help of a scanning electron microscope. Figure 1.18-A-Y depicts the images.

1.3.12.5. Accession number & submission certificates from 'Type' culture collection centres

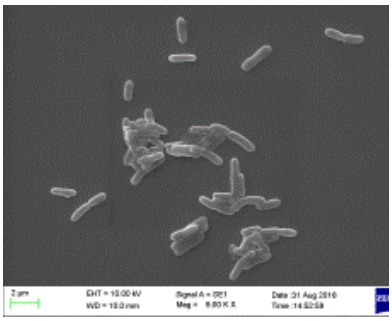
Accession numbers acquired for some of the unique bacterial strains from different internationally recognized Type culture collection centres are mentioned in Table 1.1 within parenthesis. Culture submission certificates describing their availability from these centres authenticate the submission.



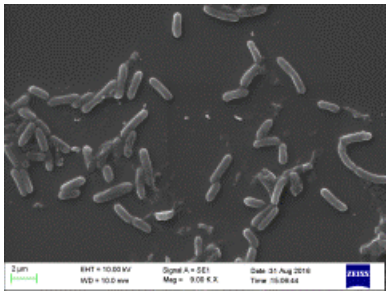
A.*Bacillus efetidiens* strain ET01



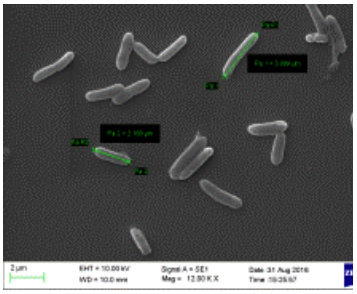
B. *Bacillus* sp. strain ET02



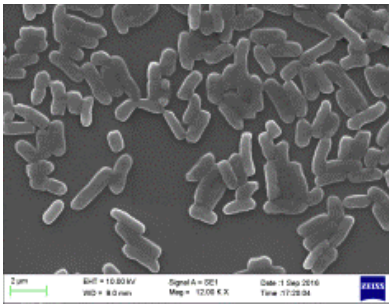
C.*Chryseomicrobium excrementi* strain ET03



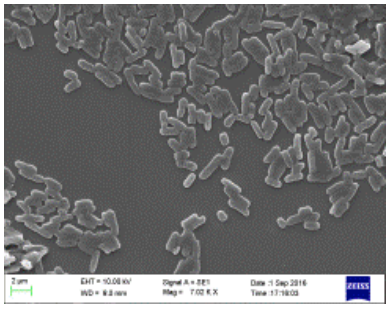
D. *Bacillus eisenifilia* strain ET04



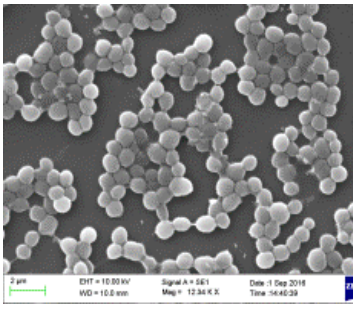
E. *Pseeudomonas aeruginosa* strain ET05



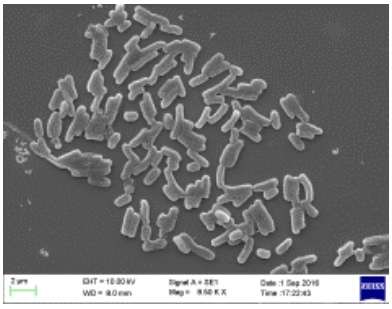
F. *Bacillus* sp. strain ET08



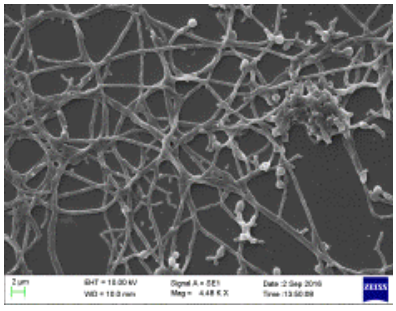
G. *Bacillus* sp. strain ET09



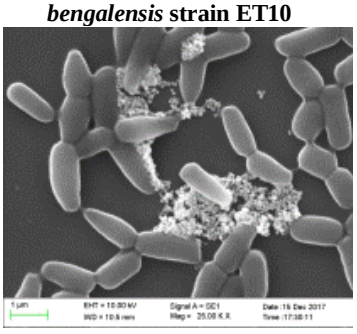
H. *Celullosaemicrobium bengalensis* strain ET10



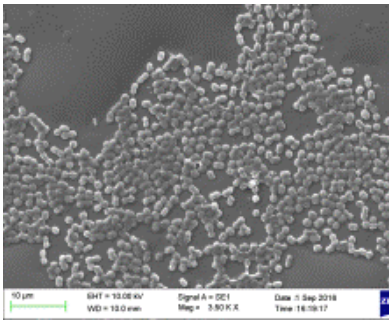
I. *Pseudoxanthomonas* sp. strain EAG1



J.*Streptomyces lapidiformes* strain EAG2

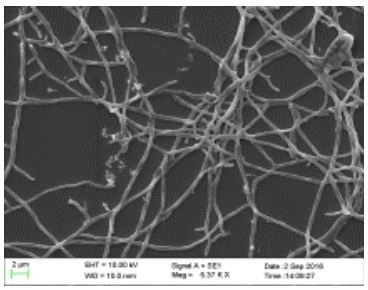


K. *Pradoshia cofami* strain EAG3

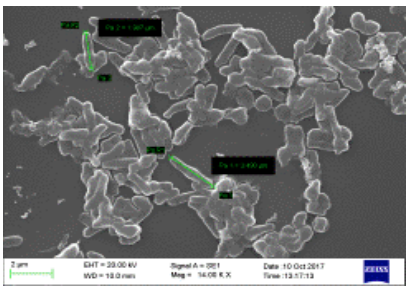


L. *Mumia enteroni* strain EAG4

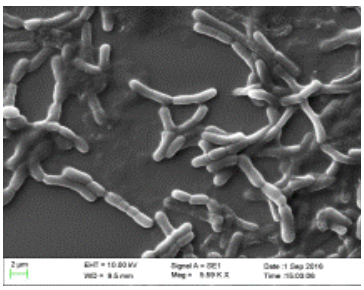
Cont….



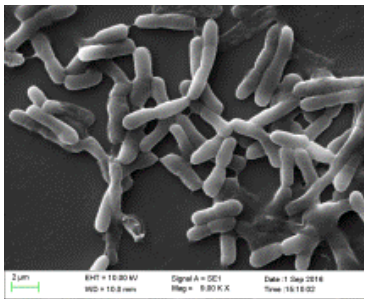
M. *Streptomyces* sp. strain EAG5



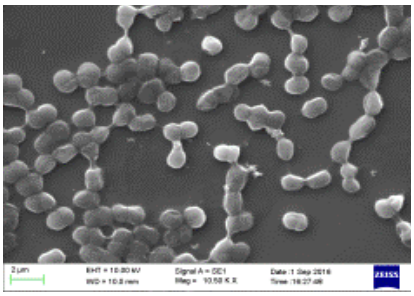
N. *Mycobacterium omicsense* strain EPG1



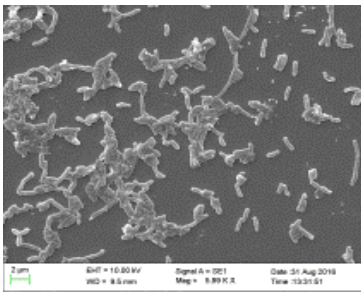
O. *Bacillus anthracis* strain EPG2



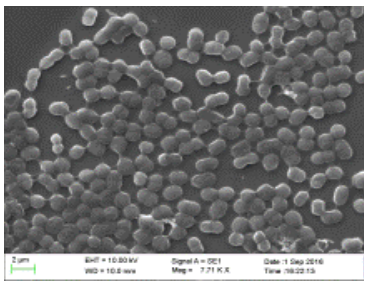
P *Bacillus cereus* strain EPG3



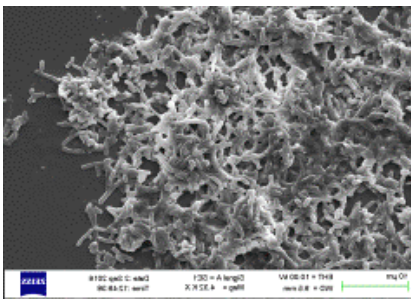
Q.*Acinetobacter desnitrificae* stain EDN1



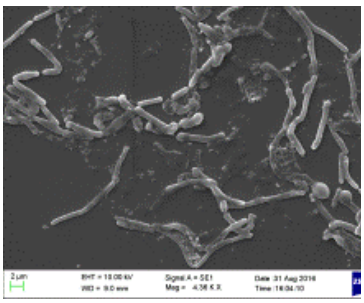
R. *Klebsiella nitrificae* strain EN1



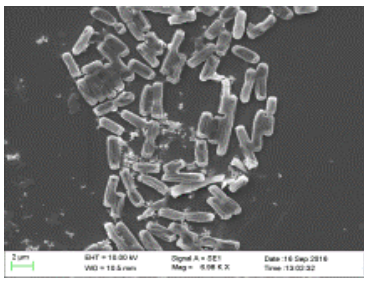
S. *Paracoccus rammohunii* strain ENF1



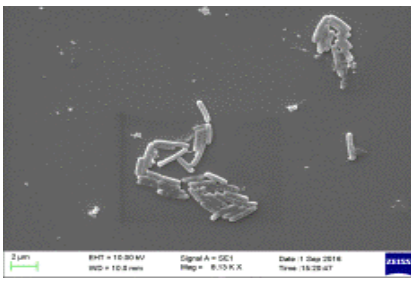
T. *Ensifer nitrofacere* strain ENF4



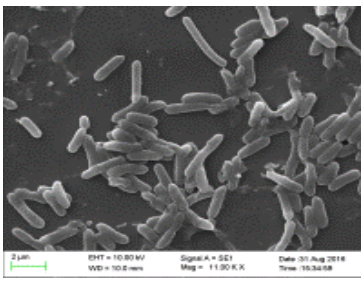
U. *Bacillus megaterium* strain Ah4



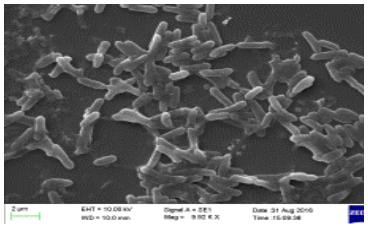
V. *Bacillus pumilus* strain BP



W. *Bacillus cereus* strain BCR



X. *Bacillus thuringiensis* strain BT



Y. *Escherichia coli* strain AG2

Fig. 1.18: (A-Y). Electron micrographs of log phase cultures of the bacterial strains isolated from the gut (A-T) and coelomic fluid (U-Y) parts of *E. fetida.*

CHAPTER - 2

WHOLE METAGENOMIC STUDY OF THE GUT CONTENT OF *EISENIA FETIDA* & WGS OF FEW GUT ASSOCIATED STRAINS

2.1 Introduction

All multicellular organisms are intricately associated with microbes by means of metabolic and physiological interactions. Traditional microbiological techniques have often failed to describe these interactions and may therefore be inadequate in detecting perturbations within soil microbial communities because 99% of soil microorganisms are not cultivable (Schwieger and Tebbe 1997). Microbes, as communities, are key players in maintaining environmental stability within and around the metazoans. Study of only cultivable microbial components, therefore lags behind in the understanding of the complex communities and role of their individual members. Several culture-independent methods have been developed for the assessment of microbial community structure and identification of species within the community. Most common are methods that rely on extraction of DNA from soil and subsequent high throughput sequencing.

In principle, the concept behind NGS technology is similar to capillary electrophoresis sequencing. Fluorescently labeled deoxyribonucleotide triphosphates (dNTPs) are incorporated into a DNA template strand during sequential cycles of DNA synthesis catalyzed by the DNA polymerase. Just at the point of incorporation of nucleotides during each DNA synthesis cycle, the individual nucleotides are identified by their unique fluorophore excitation pattern. The same process of DNA synthesis and nucleotide identification is parallelly extended for the millions of fragments generated from the whole metagenome or genome in NGS. This sequencing by synthesis (SBS) chemistry on Illumina platform (>90% of the world's sequencing data) is the most widely used sequencing technology. This is advantageous because of its high accuracy, error-free reads, and a high percentage of base calls above Q30. Illumina NGS workflows have the following four basic steps-

1. Library Preparation— The DNA or cDNA sample is randomly fragmentated to produce the sequencing library. These fragments are ligated with 5′and 3′adapters. Alternatively, these two steps may be combined into a single step “tagmentation” incorporating both fragmentation and ligation reactions and thus increasing its efficiency. Adapter-ligated fragments are then amplified by PCR and gel purified.

2. Cluster Generation— The library is loaded into a flow cell for cluster generation. The fragments are captured on a lawn of surface-bound oligos which are complementary to the

adapters. Each fragment bound on the surface is then amplified into distinct, clonal clusters.

3. Sequencing—Illumina SBS technology is capable of detecting single bases as they are incorporated into DNA template strands. The four types of reversible terminator–bound dNTPs are present during each sequencing cycle, thus natural competition minimizes incorporation bias and reduces error rates (Ross *et al.* 2013; Bentley *et al.* 2008). This highly accurate base-by-base result is the essence of NGS sequencing that virtually eliminates sequence context–specific errors, even within repetitive sequence regions and homopolymers.

4. Data Analysis—The sequence reads are aligned to a reference genome. Following alignment, various types of analysis can be done. Single nucleotide polymorphism (SNP) or insertion-deletion (indel) identification, read counting for RNA study, phylogenetic or metagenomic analysis are done during data analysis.

A major advantage in modern NGS technology is its paired-end (PE) sequencing which involves sequencing both ends of each DNA fragments in a library and then aligning the forward and reverse reads for that particular fragment. It not only doubles the number of reads for the same time and effort in library preparation, but also makes the aligned sequences more accurate (Nakazato *et al.* 2013). Differential read-pair spacing also allows removal of PCR duplicates, which is a common artifact produced during library preparation from PCR amplification. PE sequencing also produces a higher number of SNV calls following read-pair alignment **(**Nakazato *et al.* 2013**)**. These advancements of the paired-end approach of NGS has made it researchers choice. Although some methods such as small RNA sequencing is best served by single-read sequencing. Metagenomics is a combination of *research techniques*, comprising many related approaches and methods.

Meta (Greek) means "transcendent." Bypassing the shortcomings of classical Microbiology by which most microbes due to unculturability cannot be studied, Metagenomics can reveal genomic diversity of any environmental sample. Although only a decade old, this science of metagenomics has enabled us to investigate microbes in their natural environments. Complexity of the communities along with their functional attributes can also be revealed. Thus metagenomics has brought about a paradigm shift in biology, medicine, ecology, and biotechnology (Gilbert, 2011). We can uncover patterns of

ecological interactions like lateral gene transfer and metabolic complementation through the lens of metagenomics. Whole metgenome analysis in our study has been carried out on Illumina MiSeq platform.

Whole Genome Sequencing (WGS) using Next Generation Sequencing (NGS) is a useful tool in determining the DNA sequence, information which is valuable in furthering our understanding of biological processes. WGS is a blooming technology of preference because of its universality and affordability. This technology enables us to assemble genomes either *de novo* or compare the genome of any organism to a reference genome. In this study, WGS of the notably unique *E. fetida* associated bacterial strains ET03, EPG1, EAG2 and EAG3 has been performed using Illumina NextSeq 500 NGS platform.

2.2 Materials and methods

2.2.1. Isolation of Gut-Metagenomic DNA

Genomic DNA extraction from gut sample was carried out for identification of the culture independent microbial community by following methods described by Zhou *et al.* (1996). Quality was checked on 1% agarose gel (loaded 5 µl) for the Single intact band. The gel was run at 110 V for 30 mins. 1 µl of the sample was loaded in Nanodrop 8000 for determining $A_{260/280}$ ratio. 1 µl of the sample was used for determining concentration using Qubit® 2.0 Fluorometer.

2.2.2. Preparation of 2 x300 MiSeq library

The paired-end sequencing library was prepared using illumina TruSeq DNA LT Library Preparation Kit. Library preparation was started with gDNA fragmentation of 1µg DNA, followed by paired-end adapter ligation. The ligated product will be purified using 1X Ampure beads. The purified ligated product was subjected to size-selection at ~650-1000 bp. The size-selected product was PCR amplified as per manufacture's protocol.

2.2.3. Cluster Generation and Sequencing

QC passed library with a mean peak size (Bioanalyser profile) was loaded onto MiSeq platform for cluster generation and sequencing. Paired-End sequencing allowed the template fragments to be sequenced from both the forward and reverse ends. The kit reagents were used in binding of samples to complementary adapter oligos on paired-end

flow cell. The adapters were designed to allow selective cleavage of the forward strands after re-synthesis of the reverse strand during sequencing. The copied reverse strand was then used to sequence from the opposite end of the fragment

2.2.4. Whole metagenome Assembly

High quality reads was assembled using CLC workbench with default parameters (mismatch cost 2, insertion cost 3, deletion cost 3, length fraction 0.5, similarity fraction 0.8 etc.).

2.2.5. MEGAN analysis

2.2.5.1. Annotation of contigs

Assembled contigs were then subjected to annotation. The annotation will be performed by aligning the contigs to non-redundant database of NCBI using BLASTX. The output of the BLASTX was uploaded into MEGAN. Phylogenetic diversity of the Contigs computed by MEGAN using LCA algorithm, based on a BLASTX comparison of the contigs against the NCBI-NR database

2.2.5.2. Taxonomic analysis

BLASTx comparison against the NCBI-NR database was used to reveal the Phylogenetic diversity of the 216738 contigs by MEGAN using LCA algorithm.

2.2.5.3. Functional analysis using the SEED classification:

MEGAN maps each contig to a SEED functional role, using the highest scoring BLAST match to a known protein sequence. The SEED classification, depicted as a rooted tree represents the different subsystems and the functional roles. The contigs sequences are assigned to major functional groups like carbohydrates metabolism, Protein metabolism, DNA metabolism etc.

2.2.5.4. Functional analysis using the KEGG classification

MEGAN matches each contig to a KEGG orthology (KO) accession number, using the best hit to a known reference sequence. This information is used for assigning the contigs to enzymes and pathways. The KEGG classification is represented by a rooted tree to show different pathways like metabolism, genetic information processing etc.

2.2.5.5. Functional analysis using the COG classification

Gene function in the COG classification encompasses a collection of biologically defined clusters of orthologous groups. The COG classification is represented as a tree containing nodes and edges. MEGAN maps each read onto a gene with a known COG or NOG. The contigs sequences are assigned to major COG categories such as metabolism, Information, storage and processing etc.

2.2.5.6. Rarefaction Analysis

Rarefaction allows the calculation of species richness for a given number of individual samples, based on the construction of so called rarefaction graphs. This graph can be used to estimate (roughly) how many additional species are likely to be discovered if one were to increase number of reads by factor of two. This can be addressed by plotting the discovery rate of dataset, which is obtained by repeatedly selecting random subsample of the dataset at 10,20, .., 90 % of the original size and then plotting the number of leaves predicted by LCA algorithm.

2.2.6. MG-RAST Analysis

Dataset metagenome was uploaded on MG-RAST (MGRAST ID: 4584728.3) which contains 216,738 sequences (a total of 132,421,787 basepairs with an average length of 610 bps). The pie chart in the figure 7 breaks down the uploaded sequences into 5 distinct categories i.e. Annotated protein, unknown protein, ribosomal RNA sequences, unknown sequences and failed QC sequences.

2.2.6.1. Kmer Profiles

The kmer abundance spectra are tools to summarize the redundancy (repetitiveness) of sequence datasets by counting the number of occurrences of 15 and 6 bp sequences. The kmer spectrum plots the number of distinct N-bp sequences as a function of coverage level, placing low-coverage (rare) sequences at left and high-coverage, repetitive sequences at right. The kmer rank abundance graph plots the kmer coverage as a function of abundance rank, with the most abundant sequences at left. The ranked kmer consumed graph shows the fraction of the dataset that is explained by the most abundant kmers, as a function of the number of kmers used.

2.2.6.2. Source Hits Distribution

The predicted protein features were annotated by the different databases. These include protein databases, protein databases with functional hierarchy information, and ribosomal RNA databases. The annotated reads are colored by e-value range. Different databases have different numbers of hits, but can also have different types of annotation data.

2.2.6.3. Functional Category Hits Distribution

COG (Clusters of Orthologous Groups), KO [KEGG (Kyoto Encyclopedia of Genes and Genomes) orthology] and NOG (Non-supervised Orthologous Groups) functional Category Hits Distribution were performed in the MG-RAST along with subsystem analysis

2.2.6.4. Taxonomic Hits Distribution

Taxonomic Hits illustrating the distribution of taxonomic domains, phyla, class, order, family and genus for the annotations were represented by the pie charts, where each slice indicates the percentage of reads with predicted proteins and rRNA genes annotated to the indicated taxonomic level using all source databases.

2.2.6.5. Rank Abundance Plot

The species abundances ordered from the most abundant to least abundant has been plotted. Only the top 50 most abundant were taken into account. The y-axis plots the abundances of annotations in each species on a log scale. The rank abundance curve is a tool for visually representing taxonomic richness and evenness.

2.2.6.6. Rarefaction Curve

Rarefaction curve is a plot of the total number of distinct species annotated as a function of the number of sequences that were sampled. Sampling curve generally rises sharply at first and then level off as fewer new species are found later. These curves are calculated in respect to species abundance. These rarefaction curves represent the average number of different species annotated for subsamples of the complete dataset.

2.2.6.7. Alpha Diversity

Alpha diversity indicates the diversity of organisms in a sample. The alpha diversity of the

samples is calculated from the species-level annotations. Species richness is the number of distinct species that can be annotated in the dataset. Shannon diversity is the abundance-weighted average of the logarithm of relative abundances of these annotated species.

2.2.6.8. Sequence Length Histogram and GC Distribution

The Sequence Length Histogram and GC Distribution are plotted on the basis of distribution of sequence lengths in basepairs and the respective GC percentages for these amplicons. The data used in these graphs are based on raw upload and post QC sequences.

2.2.7. Whole genome sequencing (WGS) and Metabolic analysis of four novel bacterial strains isolated from the gut of *E. fetida*: A necessary corroboration with the metagenomic analyses

WGS begins with library preparation, as explained in the metagenomic Experimental Design section. Sequencing was performed on Illumina NextSeq 500 NGS platform.

2.2.7.1. *Isolation of genomic DNA*

Genomic DNA extraction, purification and quantification from each of the four strains ET03, EPG1, EAG2 and EAG3 were carried out by following methods described by Furlong *et al.* in 2002 with modifications.

2.2.7.2 *Sequencing, Assembly and Gene Prediction.*

The genome of strain EPG1^T was sequenced using NextSeq 500. Briefly, approximately 200 ng DNA was fragmented by covaris M220 to generate~400bp segments. End-repaired products were size selected by AMPure XP beads, PCR amplified with index primers and analyzed in 4200 Tape Station system (Agilent Technologies). After obtaining Qubit concentration, PE Illumina libraries were loaded onto NextSeq 500 for cluster generation and sequencing. The copied reverse strands were then used to sequence from the opposite end of the fragment. Thus adapter free data of 1.1 Gb was generated which is required for the genome to be used for a taxonomic purpose (Chun *et al.*, 2018). The high- quality reads were then *de-novo* assembled using SPAdes genome assembler. Prokka (Seemann, 2014) was used to predict the genes from final scaffolds.

2.2.7.3.Phylogenetic analysis (AAF), Average nucleotide identity (ANI) and Genome-to-genome direct comparison (GGDC) analyses.

An assembly and alignment-free (AAF) method (Fan *et al.,* 2015) was used to construct phylogeny from next-generation sequencing data. The calculation of the BLAST based average nucleotide identity (ANI) score was done using the JSpeciesWS program with the default parameters (Richter *et al.,* 2015). Genome-to-genome direct comparison (GGDC) analyses were performed using all three equations in the GGDC program, version 2.1 (Meier-Kolthoff *et al.,* 2014).

2.2.7.4. Functional analysis by gene annotation (BLASTx) and Gene ontology (Blast2Go)

Functional annotation of the genes was performed using BLASTx program of NCBI blast-2.3.0+ standalone tool to find out homologous sequences against NR (non-redundant) protein database. BLASTx accession IDs were used to retrieve UniProt IDs using PIRs including PSD, UniProt, SwissProt, TrEMBL, RefSeq, GenPept and PDB databases. Gene ontology (GO) annotations of the genes were determined by the Blast2GO platform. GO mapping was carried out in order to retrieve GO terms for all BLASTx functional annotated genes and to classify the functions of the predicted genes. GO mapping also provides the ontology of defined terms representing gene product properties grouped into three main domains: Biological process, Molecular function and Cellular component.

2.3. Results and discussion

2.3.1. Qualitative and quantitative analysis of Gut-Metagenomic DNA

1% agarose gel showed the distinct band of isolated genomic DNA (Fig. 2.1). The quantification on nanodrop analysis is represented in table: 2.1.

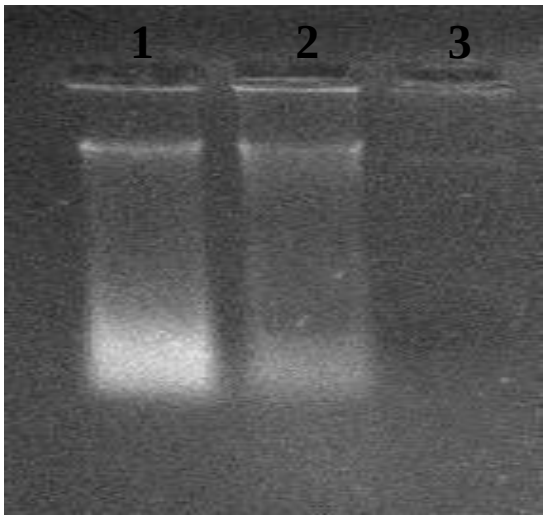


Fig. 2.1 : QC of DNA in 1% agarose gel

Table.2.1: Quantification of gDNA on nanodrop

S. No.	Sample ID	Volume	Conc. (ng/µl)	Yield (µg)
1	1	150	31. 5	4.73
2	2	60	9.7	0.58
3	3	30	0.85	0.025

2.3.2. Preparation of 2 x300 MiSeq library

The library was prepared using illumina TruSeq DNA LT Library Preparation Kit (Fig. 2.2). The mean size of the library is 936 bp (Table: 2.2). The library was sequenced on Miseq using 2 x 300 bp chemistry kit.

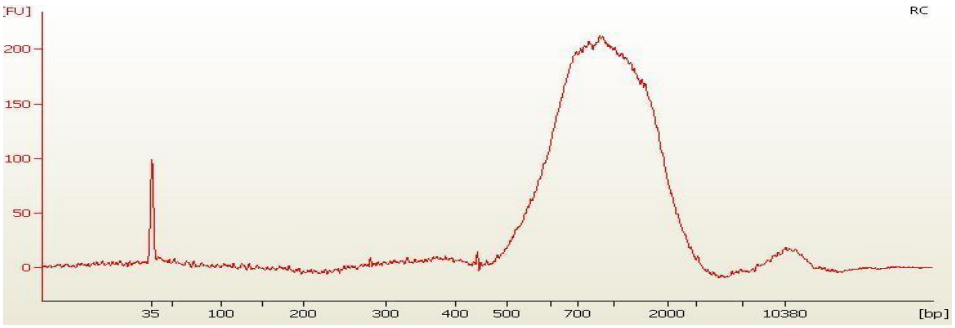


Fig. 2.2: Bioanalyzer profile of Library on HS DNA chip

Table 2.2: MiSeq library information sheet

From [bp]	To [bp]	Corr. Area	% of Total	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/µl]	Molarity [pmol/l]
549	1666	2430.1	78	936	32.8	19552.5	35480.1

2.3.3. Cluster Generation and Sequencing

A total of 89,94,027 high-quality reads (4.02GB) were assembled using CLC workbench with default parameters (mismatch cost 2, insertion cost 3, deletion cost 3, length fraction 0.5, similarity fraction 0.8 etc.)

2.3.4. Whole metagenome Assembly

A total of 89,94,027 high quality reads were assembled using CLC workbench with default parameters (mismatch cost 2, insertion cost 3, deletion cost 3, length fraction 0.5, similarity fraction 0.8 etc.).

Total genome length with 132315394 reads distributed in 216738 No. of contigs was created. The Max. contig size was 403603.

2.3.5. MEGAN analysis

2.3.5.1. Annotation of contigs

Phylogenetic diversity of the metagenomic Contigs computed by MEGAN using LCA algorithm, based on a BLASTX comparison of the 216738contigs against the NCBI-NR database. Each circle represents a taxon in the NCBI taxonomy and is labeled by its name and the number of contigs that are assigned either directly to the taxon, or indirectly via one of its subtaxa. The size of the circle is scaled logarithmically to represent the number of contigs assigned directly to the taxon.

2.3.5.2. Taxonomic analysis

At Phylum level:

A total number of 112045 contigs assigned to bacteria, out of these contigs, 21697contigs were assigned to *Proteobacteria*, 14583 contigs to *Firmicutes*, 15383 contigs to *Bacteroidetes* and 11523 contigs assigned to *Actinobacteria* (Fig. 2.3).

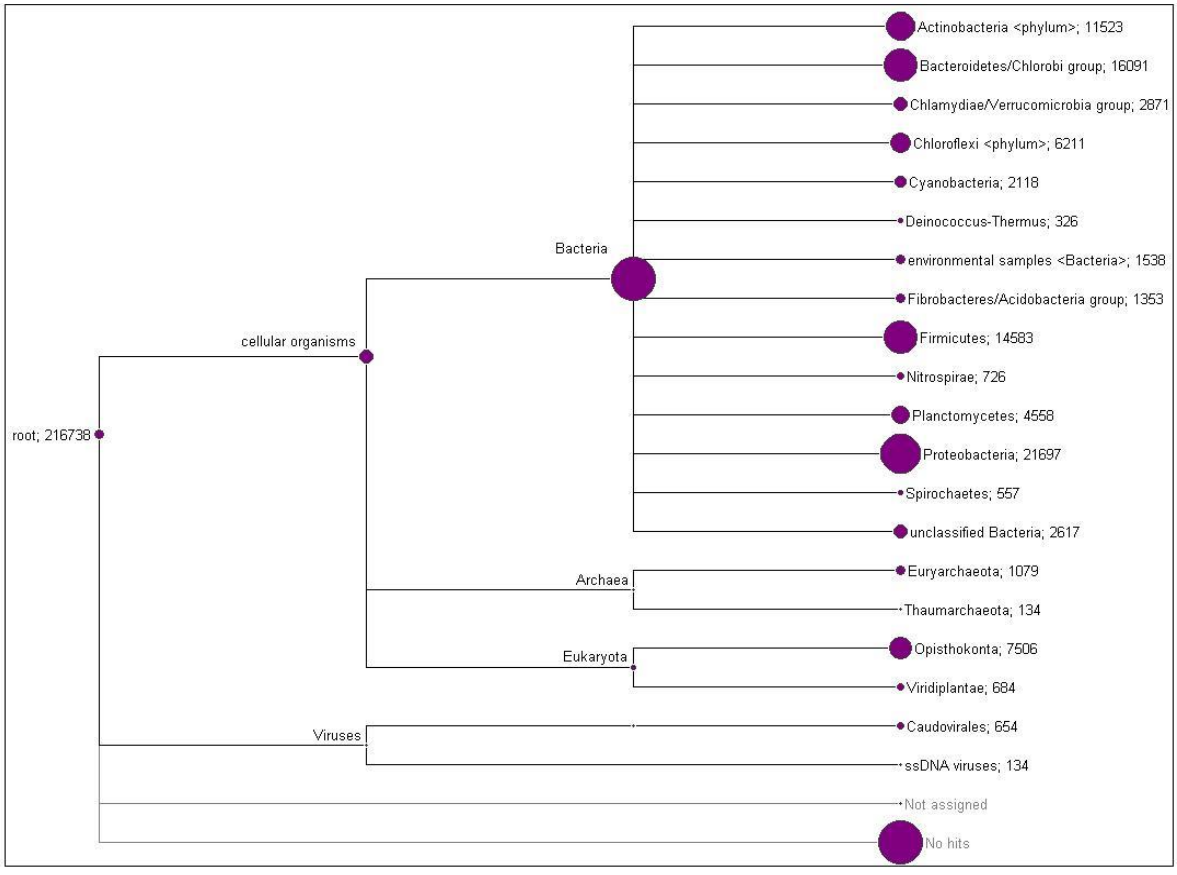


Fig. 2.3: Taxonomic View of the Metagenome

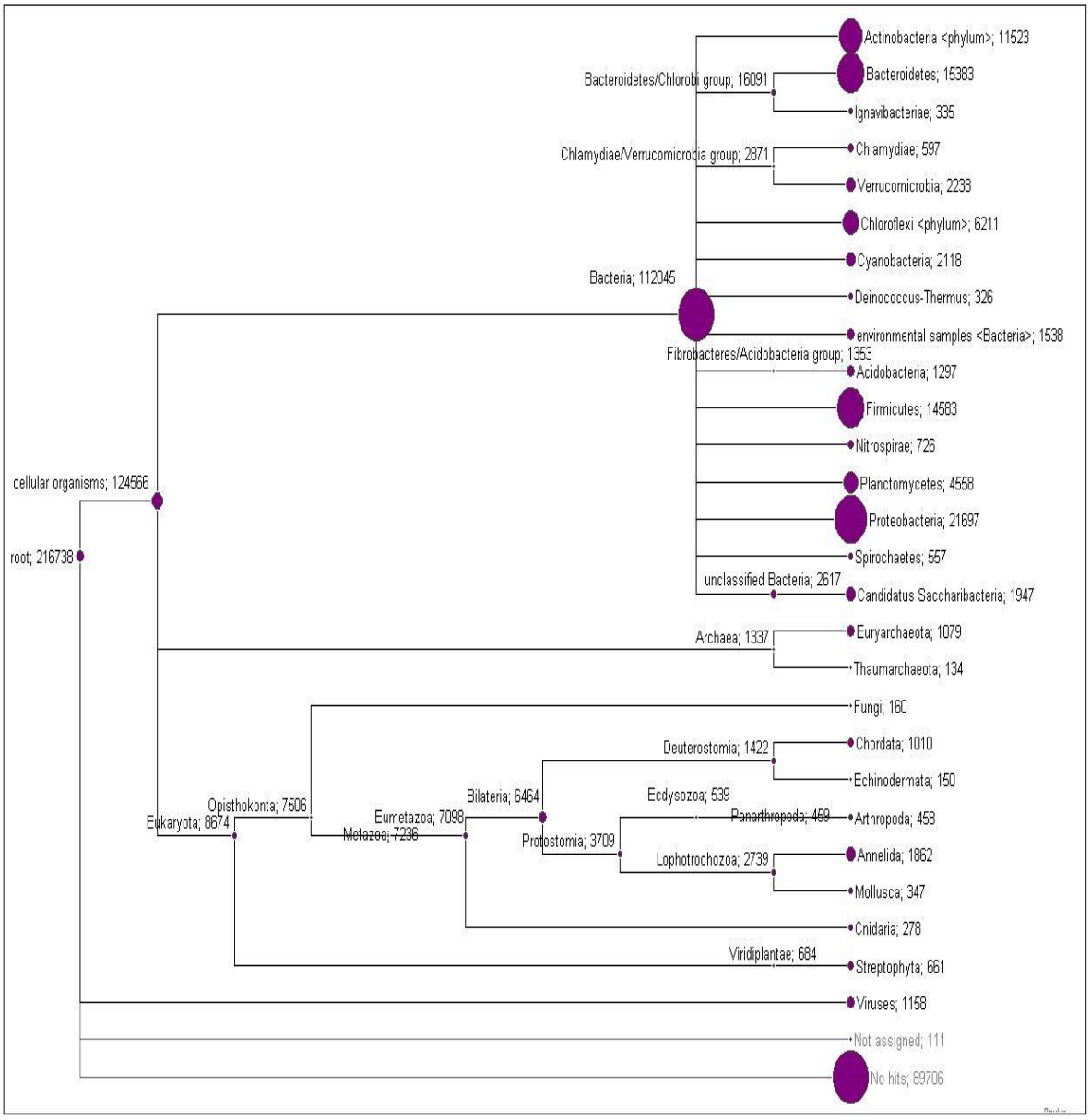


Figure 2.4: Taxonomic view of the metagenome at phylum level. The size of the circle isscaled logarithmically to represent the number of contigs assigned directly to the phylum.

As shown in Table 2.3: maximum contigs assigned on *Caldilinea aerophila* (1962 contigs), *Flavobacterium* (1774 Contigs), Cytophagaceae (938 contigs), *Sorangium cellulosum* (843 contigs), *Mycobacterium* (809 contigs) etc (Fig. 2.4).

Table 2.3: Species level phylogenetic diversity of the metagenome computed by MEGAN

Species Name	No. of Contigs assigned	Species Name	No. of Contigs assigned	Species Name	No. of Contigs assigned
Caldilinea aerophila	1962	*Coriobacteriaceae*	287	*Chloracidobacterium*	
Flavobacterium	1774	*Microbacterium*	286	*Thermophilum*	178
uncultured bacterium	1338	*Endopterygota*	273	*Schlesneria paludicola*	177
		Candidatus Solibacter			
Cytophagaceae	938	*usitatus*	270	*Micrococcaceae*	177
Sorangium cellulosum	843	*Methanobacterium*	269	*Singulisphaera acidiphila*	176
Mycobacterium	809	*Paenibacillus*	266	*Eubacterium sp. CAG:146*	175
Capitella teleta	763	*Pirellula staleyi*	265	*Chromatiaceae*	172
Ilumatobacter				*Candidatus Microthrix*	
coccineus	739	*Methanobrevibacter*	261	*parvicella*	172
Helobdella robusta	716	*Lactobacillales*	260	*Rhodococcus*	169
Candidatus Nitrospira		*Herpetosiphon*			
defluvii	700	*aurantiacus*	260	*Chloroflexus*	165
Anaerolinea					
thermophila	679	*Pseudonocardiaceae*	259	*Opitutus terrae*	165
Mesorhizobium	677	*Legionella*	256	*Desulfobacteraceae*	164
Chlamydiales	597	*Nannocystineae*	245	*Conexibacter woesei*	164
Streptomyces	583	*Chthoniobacter flavus*	245	*Fungi*	160
Butyrivibrio	517	*unclassified Siphoviridae*	242	*Methylococcaceae*	160
Fulvivirga imtechensis	480	*Saprospiraceae*	241	*Gastropoda*	158
Rhodobacteraceae	469	*Ktedonobacter racemifer*	236	*Leptospira*	157
		Candidatus			
Rhodocyclaceae	467	*Saccharibacteria*		*Treponema*	156
		bacterium			
Comamonadaceae	443	*RAAC3_TM7_1*	232	*Thermoanaerobacterales*	153
		Algoriphagus			
environmental samples		*machipongonensis*	232	*Planctomyces maris*	150
<clostridial				*Strongylocentrotus*	
firmicutes>	441	*Ruminococcus*	229	*purpuratus*	149
Pseudomonas	439	*Bacteroides*	228	*Halobacteriaceae*	148
Niastella koreensis	422	*Peptococcaceae*	225	*Thermaceae*	147
Bacillus	415	*Nocardioides*	220	*Deinococcales*	145
Mogibacterium sp.					
CM50	413	*Frankia*	220	*Myxococcus*	144
Verrucomicrobium					
spinosum	413	*Burkholderia*	218	*Geobacter*	144
Sphingobacteriaceae	407	*Acinetobacter*	213	*Oxalobacteraceae*	143
		Clostridiales bacterium			
Xanthomonadaceae	393	*NK3B98*	212	*Niabella*	140
Sphingomonadaceae	381	*Myoviridae*	210	*Podoviridae*	138
		unclassified			
Micromonosporaceae	381	*Lachnospiraceae*	205	*Sphaerobacteraceae*	137
		Candidatus			
		Saccharimonas			
Segetibacter koreensis	373	*aalborgensis*	197	*Porphyromonadaceae*	137
Rhodopirellula	355	*Nostocales*	197	*Blastopirellula marina*	136
Chroococcales	339	*Veillonellaceae*	196	*Methanosarcinales*	135
Oscillatoriales	334	*Hydra vulgaris*	194	*Desulfovibrionales*	135
Pedosphaera parvula	332	*Clupeocephala*	190	*ssDNA viruses*	134
Gemmata					
obscuriglobus	331	*Rhizobium*	190	*Papilionoideae*	134
		unclassified			
Zavarzinella formosa	301	*Acidobacteriaceae*	189	*Thaumarchaeota*	134
				Burkholderiales Genera	
Rhodospirillaceae	299	*Methylobacteriaceae*	187	*incertae*	
Bacteriovorax	297	*Ignavibacterium album*	185	*Sedis*	131
Bradyrhizobium	295	*Streptosporangineae*	184	*Ectothiorhodospiraceae*	129
Alteromonadales	294	*environmental samples*		*Euarchontoglires*	128
Enterobacteriaceae	290	*<Clostridium>*	182	*OMG group*	127
Roseiflexus	289	*Bdellovibrio*	179	*Liliopsida*	126
		Candidatus		*Verrucomicrobia*	126

2.3.5.3. Functional analysis using the SEED classification:

MEGAN maps each contig to a SEED functional role, using the highest scoring BLAST match to a known protein sequence. The SEED classification is shown as a rooted tree in which internal nodes represent the different subsystems and leaves represent the functionality. The different leaves may represent the same functional role, if it occurs in different types of subsystems. The contigs sequences assigned to major functional groups such as carbohydrates (61 contigs), Protein metabolism (69 contigs), DNA metabolism (38 contigs) etc (Fig. 2.5).

Fig. 2.5: SEED analysis of contigs formed by metagenome sample. The size of the circle is scaled logarithmically to represent the number of contigs assigned directly to the SEED category.

2.3.5.4. Functional analysis using the KEGG classification

In KEGG analysis, MEGAN attempts to match each contig to a KEGG orthology (KO) accession number, using the best hit to a reference sequence for which a KO accession number is known. This information is used to assign contigs to enzymes and pathways. The KEGG classification is shown by a rooted tree with many nodes, whose leaves represent the pathways. The contigs sequences are assigned to major pathways such as metabolism (435 contigs), genetic information processing (120 contigs) etc (Fig. 2.6).

Fig.2.6: KEGG analysis of metagenome Contigs. The size of the circle is scaled logarithmically to represent the number of reads assigned directly to the KEGG pathway.

2.3.5.5. Functional analysis using the COG classification

The COG classification of gene function made up of a defined clusters of orthologous groups. The COG classification is shown as a tree containing many nodes and edges. Genes are mapped onto COGs and NOGs . MEGAN attempts to map each read onto a gene that has a known COG or NOG. The contigs sequences assigned to major COG

categories such as metabolism (184 contigs), Information, storage and processing (149 contigs) etc (Fig. 2.7).

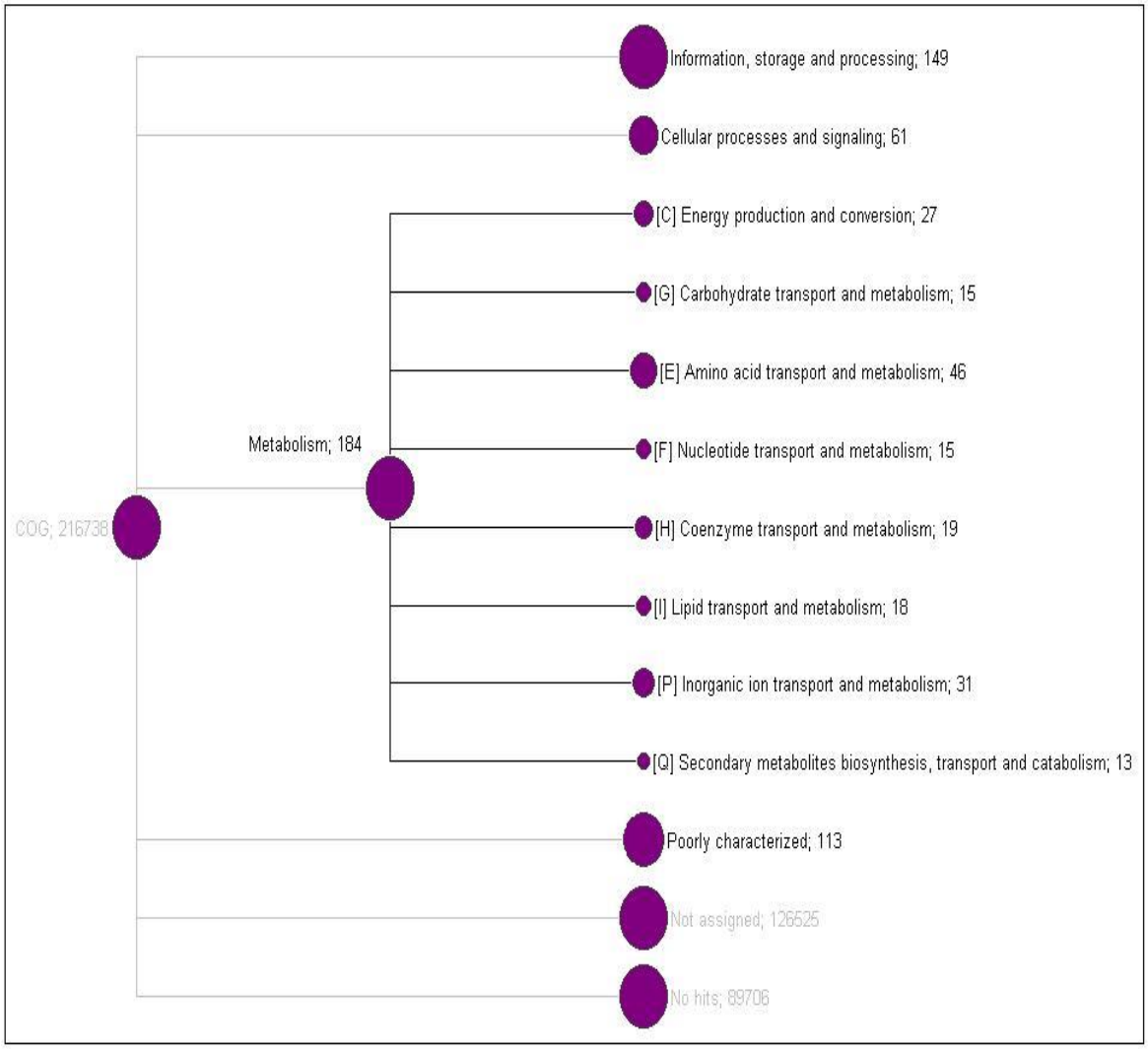


Fig. 2.7: COG analysis of metagenome Contigs. The size of the circle is scaled logarithmically to represent the number of contigs assigned directly to the COG ortholog group.

2.3.5.6. Rarefaction Analysis

Rarefaction allows the calculation of species richness for a given number of individual samples, based on the construction of so-called rarefaction graphs (Fig: 2.8). This graph can be used to estimate (roughly) how many additional species are likely to be discovered if one were to increase the number of reads by a factor of two. This can be addressed by

plotting the discovery rate of the dataset, which is obtained by repeatedly selecting a random subsample of the dataset at 10,20, .., 90 % of the original size and then plotting the number of leaves predicted by LCA algorithm.

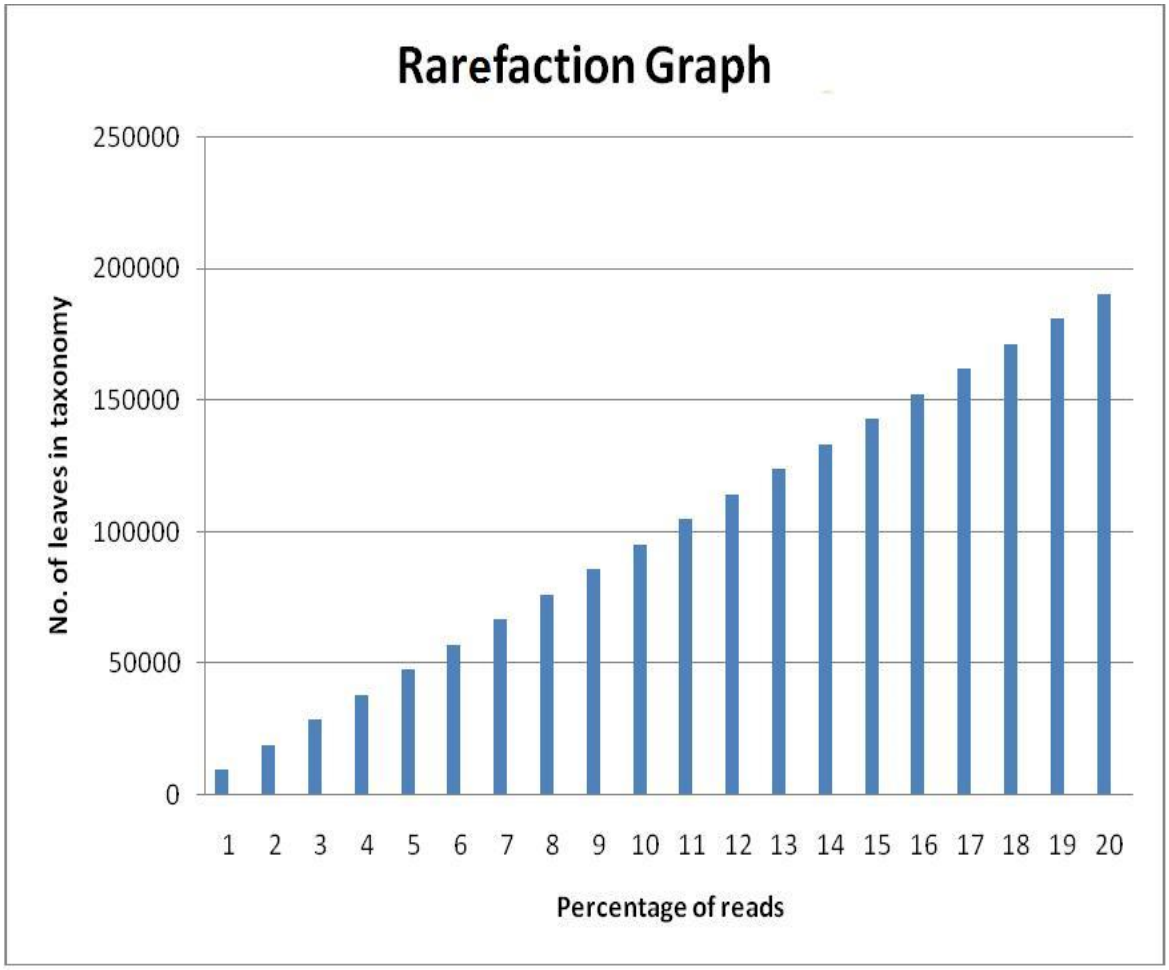


Fig: 2.8: A Rarefaction graph computed by MEGAN 5 for the metagenome Contigs dataset. The x-axis represents the percentage of reads subsamples from the total dataset and the y-axis represents the number of strong leaves computed by LCA algorithm, approximately the number of identified species.

2.3.6. MG-RAST Analysis

Dataset metagenome was uploaded on MG-RAST (ID: 4584728.3) which contains 216,738 sequences (a total of 132,421,787 basepairs with an average length of 610 bps). The pie chart in the figure 7 breaks down the uploaded sequences into 5 distinct categories i.e. Annotated protein, unknown protein, ribosomal RNA sequences, unknown sequences and

failed QC sequences. A total of 3,113 sequences failed to pass the QC pipeline. Of the sequences that passed QC, 97,132 sequences contain predicted proteins.

A total of 3,113 sequences failed quality control. Of those, dereplication identified 117 sequences as artificial duplicate reads. Of the 213,625 sequences (totaling 123,282,788 bps) that passed quality control, 252,680 produced a total of 251,613 predicted protein-coding regions. Of these 251,613 predicted protein features, 96,072 have been assigned an annotation using at least one of our protein databases (M5NR) and 155,541 have no significant similarities to the protein database (orfans). 71,852 features were assigned to functional categories (Fig.2.9).

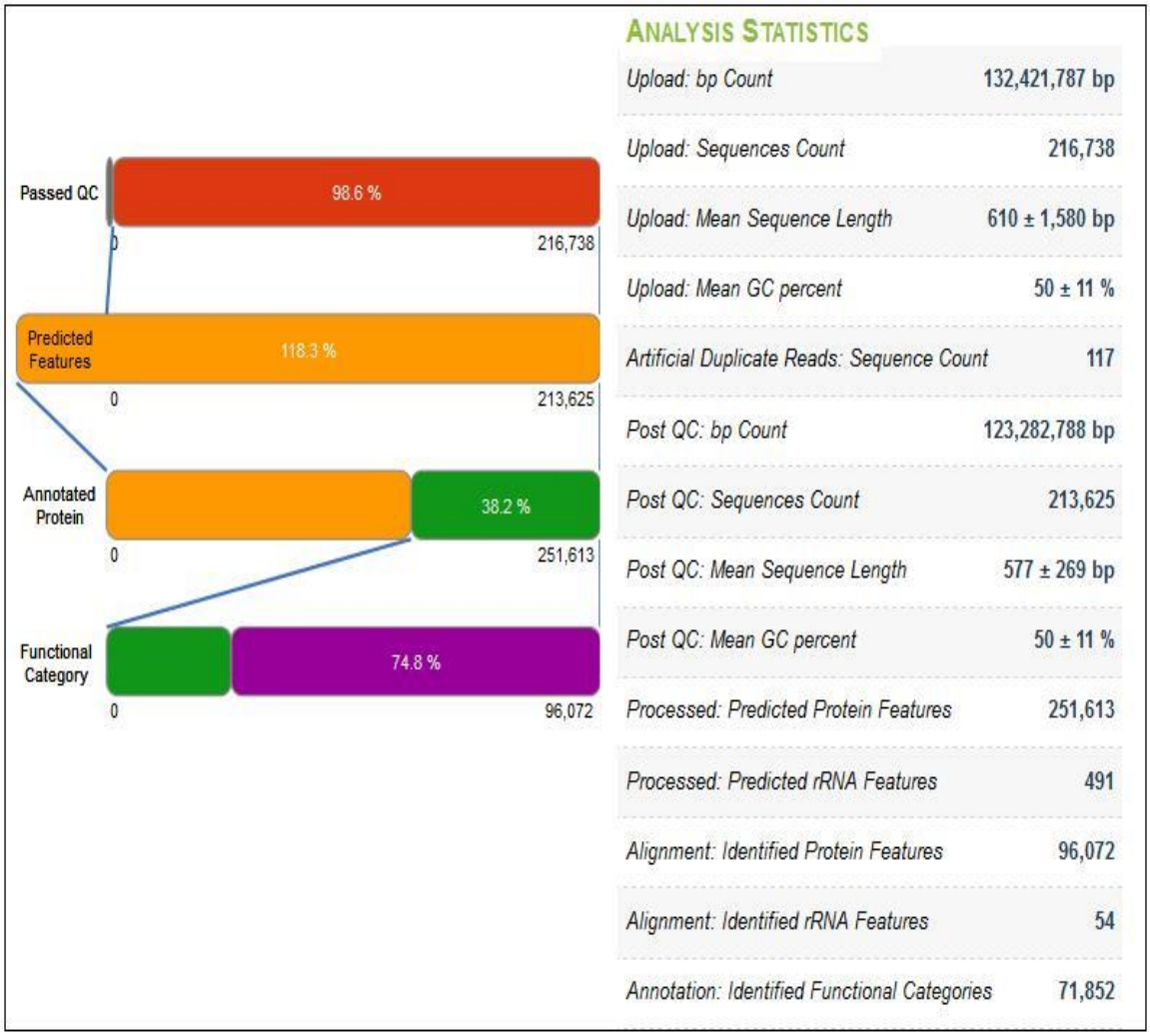


Fig.2.9: Analysis Flowchart of metagenome.

2.3.6.1. Kmer Profiles

The kmer abundance spectra are tools to summarize the redundancy (repetitiveness) of sequence datasets by counting the number of occurrences of 15 and 6 bp sequences.

The kmer spectrum plots the number of distinct N-bp sequences as a function of coverage level, placing low-coverage (rare) sequences at left and high-coverage, repetitive sequences at right. The kmer rank abundance graph plots the kmer coverage as a function of abundance rank, with the most abundant sequences at left. The ranked kmer consumed graph shows the fraction of the dataset that is explained by the most abundant kmers, as a function of the number of kmers used (Fig. 2.10).

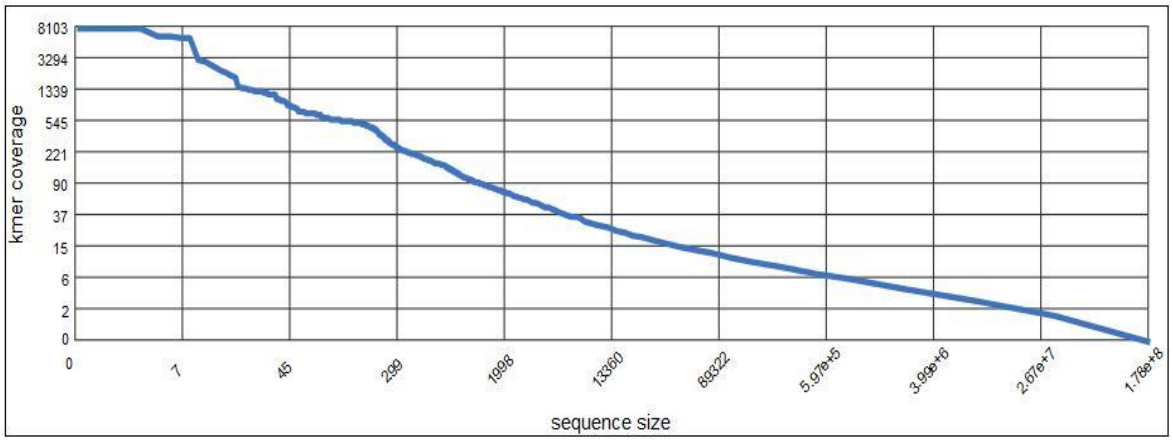


Fig. 2.10: Kmer profiles of metagenome taking sequence size in X-axis and Kmer coverage in Y-axis.

2.3.6.2. Source Hits Distribution

A total of 96,072 of the predicted protein features could be annotated with similarity to a protein of known function. 71,852 of these annotated features could be placed in a functional hierarchy. 96 of reads had similarity to ribosomal RNA genes. The graph below displays the number of features in this dataset that were annotated by the different databases below. These include protein databases, protein databases with functional hierarchy information, and ribosomal RNA databases. The bars representing annotated reads are coloured by e-value range. Different databases have different numbers of hits, but can also have different types of annotation data. There are 15,945,780 sequences in the M5NR protein database and 309,342 sequences in the M5RNA ribosomal database (Fig.2.11). The M5NR protein database contains all the unique sequences from the below

protein databases and the M5RNA ribosomal database contains all the unique sequences from the ribosomal RNA databases.

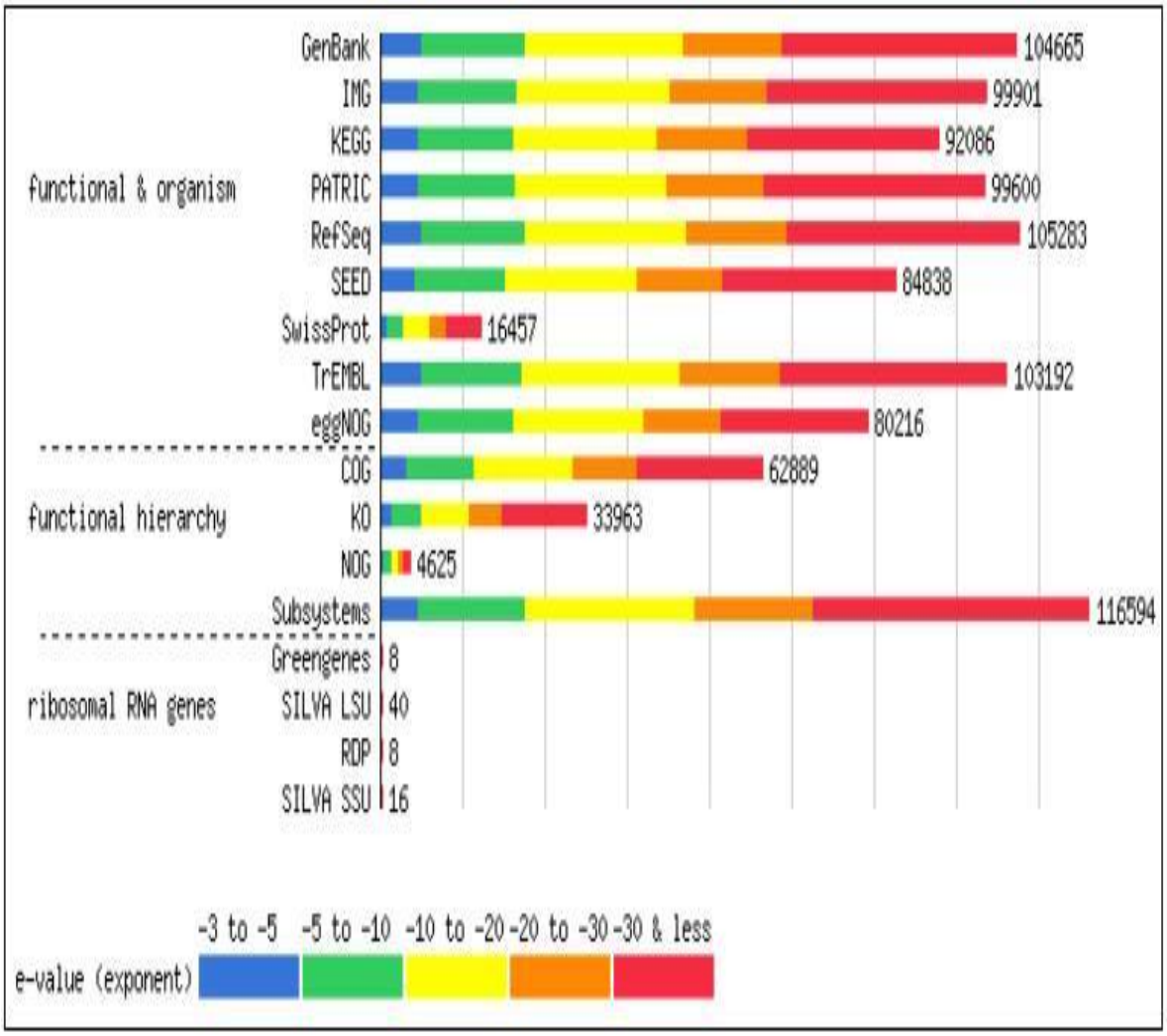


Fig.2.11: Source hit distribution of metagenome. The bar represent the e-value range from -3 to -30 or less, indicating that sequences with lower e-value are having significant matches with the functional databases.

2.3.6.3. Functional Category Hits Distribution

The pie charts below illustrate the distribution of functional categories for at the highest level supported by these functional hierarchies. Each slice indicates the percentage of reads with predicted protein functions annotated to the category for the given source.

COG Hits Distribution

COG functional category of metagenome metagenome has 61,017 predicted functions (Fig. 2.12) distributed in three main categories.

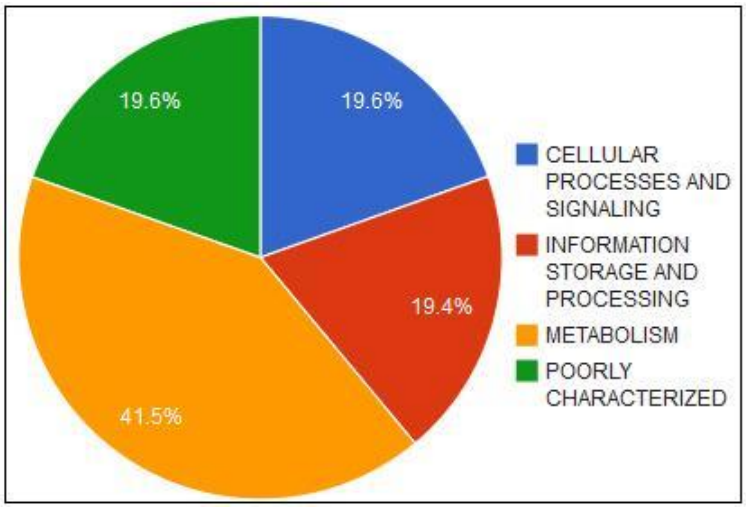


Fig. 2.12: COG Functional Category Hits Distribution

KO Hits Distribution

KO functional category of metagenome metagenome has 36,850 predicted functions (Fig. 2.13) distributed in six main categories.

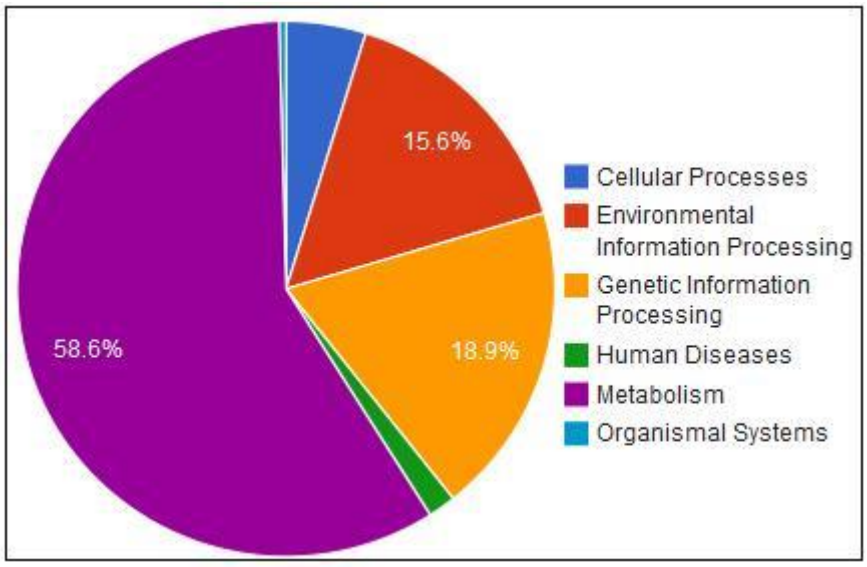


Fig. 2.13: KO Functional Category Hits Distribution

NOG Hits Distribution

NOG functional category of metagenome metagenome has 4,677 predicted functions (Fig. 2.14) distributed in four main categories.

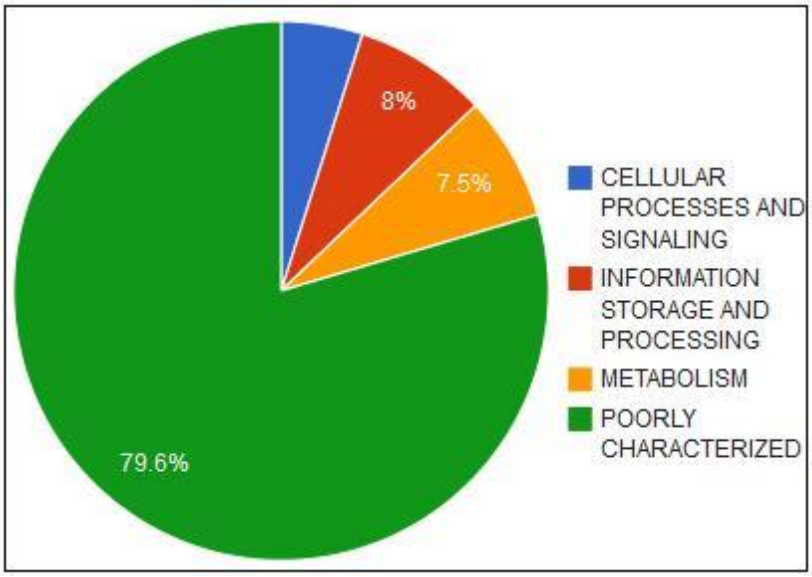


Fig. 2.14: NOG Functional Category Hits Distribution

Subsystems Hits Distribution

COG functional category of metagenome has 87,355 predicted functions in subsystem (Fig 2.15).

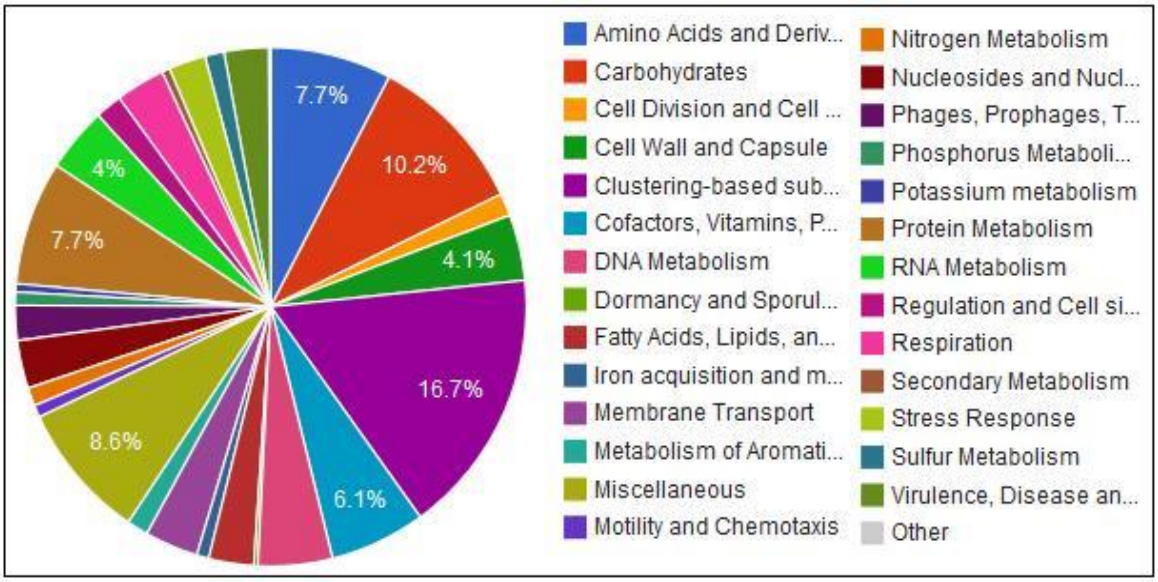


Fig 2.15: Subsystems Functional Category Hits Distribution

2.3.6.4. Taxonomic Hits Distribution

The pie charts below illustrate the distribution of taxonomic domains, phyla, and orders for the annotations. Each slice indicates the percentage of reads with predicted proteins and ribosomal RNA genes annotated to the indicated taxonomic level. This information is based on all the annotation source databases used by MG-RAST.

At Domain Level

Taxonomic hits distribution at domain level shows that metagenome has 92.8 % Bacteria, 5% Eukaryota as represented in Fig. 2.16.

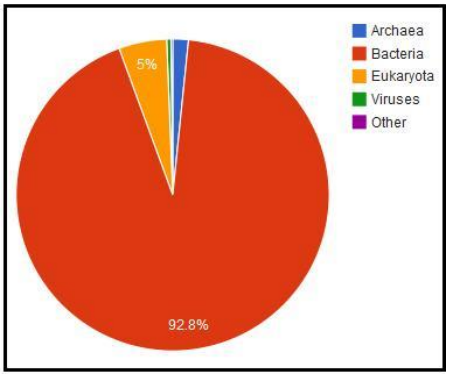


Fig. 2.16: Taxonomic hits distribution at domain level of metagenome that contains maximum hits from bacteria followed by eukaryota.

At Phylum Level

Taxonomic hit distribution at phylum level shows that metagenome has 28.5% *Proteobacteria*, 15.2 % *Firmicutes*, 13.1% *Actinobacteria* etc. as represented in Fig. 2.17.

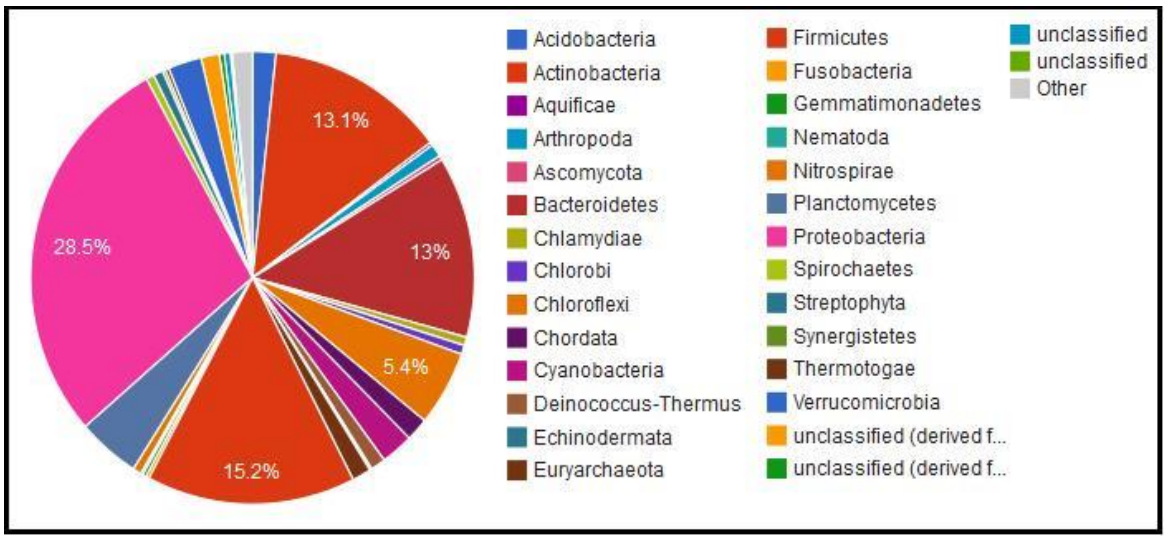


Fig. 2.17: Taxonomic hits distribution at phylum level of metagenome.

At Class Level

Taxonomic hit distribution at class level shows that metagenome has 13.1 % *Actinobacteria* (class), 10.5% *Clostridia*, 9% *α-Prteobacteria* etc. as represented in Fig. 2.18.

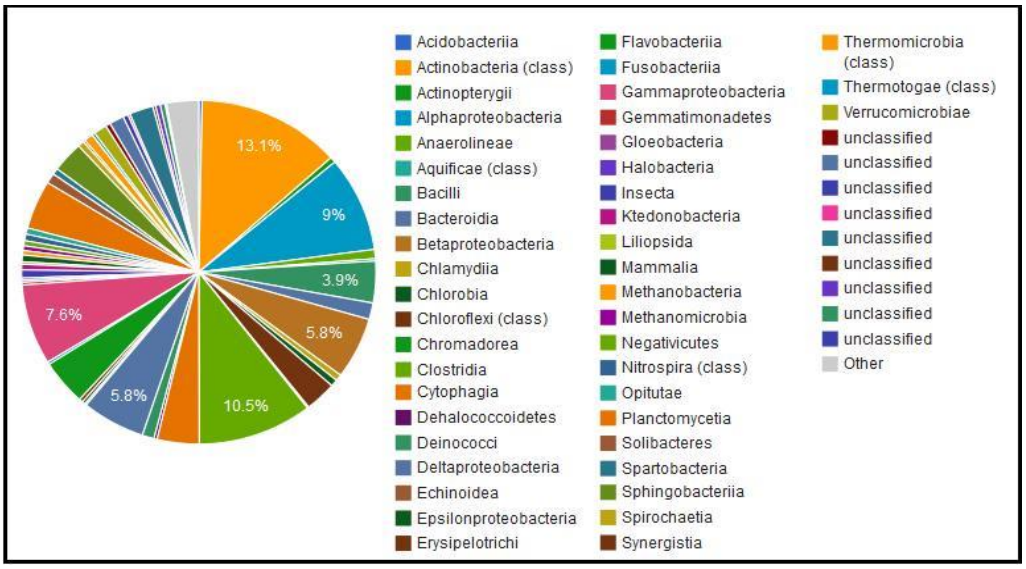


Fig. 2.18: Taxonomic hits distribution at class level of metagenome.

At Order Level

Taxonomic hit distribution at order level shows that metagenome has 11.5 % *Actinomycetales*, 9.4% *Clostridiales*, 5.1% *Rhizobiales* etc. as represented in Fig. 2.19**.**

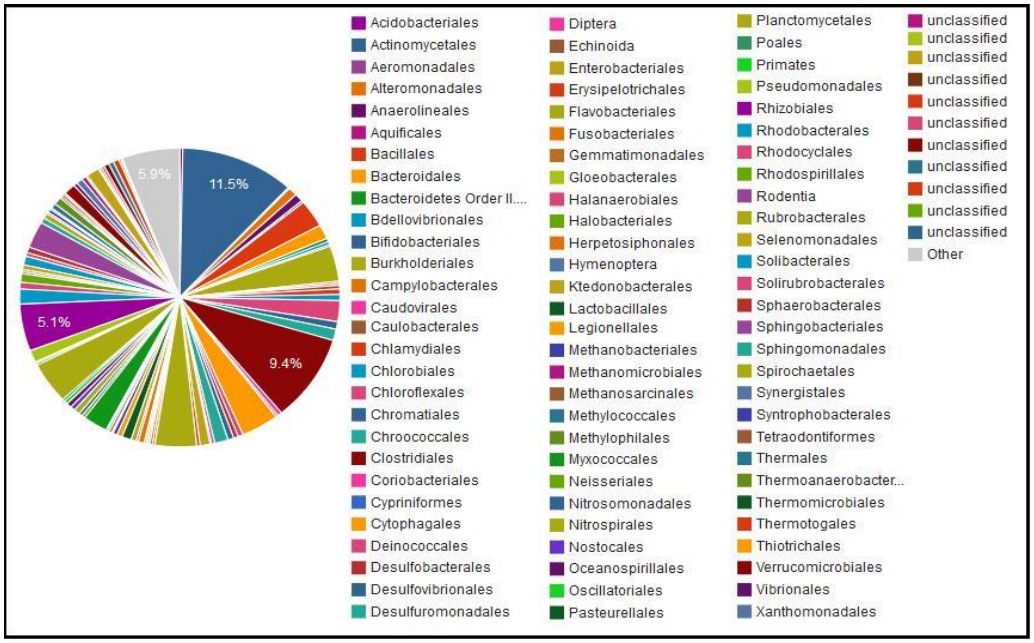


Fig 2.19: Taxonomic hits distribution at order level of metagenome.

At Family Level

Taxonomic hit distribution at family level shows that metagenome has 4.5% *Planctomycetaceae*, 3.7% *Flavobacteriaceae* etc. as represented in Fig. 2.20.

Fig.2.20: Taxonomic hits distribution at family level of metagenome.

At Genus Level

Taxonomic hit distribution at genus level shows that metagenome has 2.4% *Clostridium*, 1.8% *Eubacterium* etc as represented in (Fig. 2.21)**.**

Fig. 2.21:Taxonomic hits distribution at genus level of metagenome

2.3.6.5. Rank Abundance Plot

The plot below shows the species abundances ordered from the most abundant to least abundant. Only the top 50 most abundant are shown. The y-axis plots the abundances of annotations in each species on a log scale (Fig.2.22). Rank abundance curve is a tool for visually representing taxonomic richness and evenness.

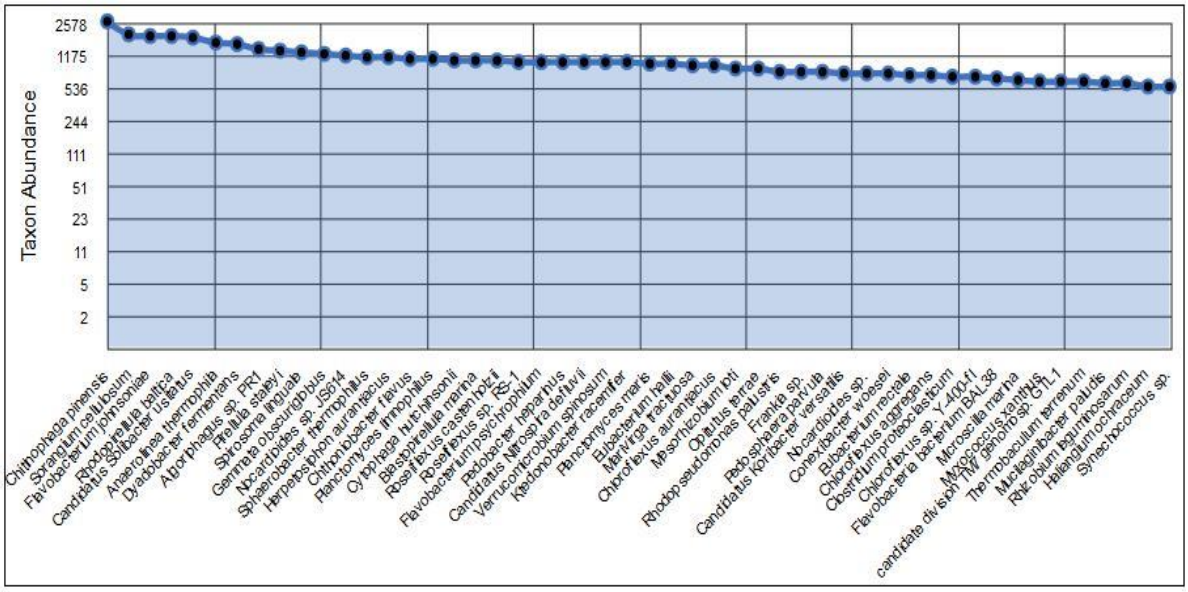


Fig.2.22: Rank abundance plot of metagenome. This figure provides a rank-ordered list of taxonomic units at a user-defined taxonomic level, ordered by their abundance in the annotations. The species Chitinophaga pinensi, sorangium cellulosum, and Actinobacteria showed highest abundance whereas Haliangium ochraceum and Synechococcus sp. showed the lowest abundance.

2.3.6.6. Rarefaction Curve

The plot below shows the rarefaction curve of annotated species richness. This curve is a plot of the total number of distinct species annotations as a function of the number of sequences sampled. On the left, a steep slope indicates that a large fraction of the species diversity remains to be discovered. If the curve becomes flatter to the right, a reasonable number of individuals is sampled: more intensive sampling is likely to yield only few additional species. Sampling curves generally rise very quickly at first and then level off towards an asymptote as fewer new species are found per unit of individuals collected. These rarefaction curves are calculated from the table of species abundance (Fig.2.23). The curves represent the average number of different species annotations for subsamples of the complete dataset.

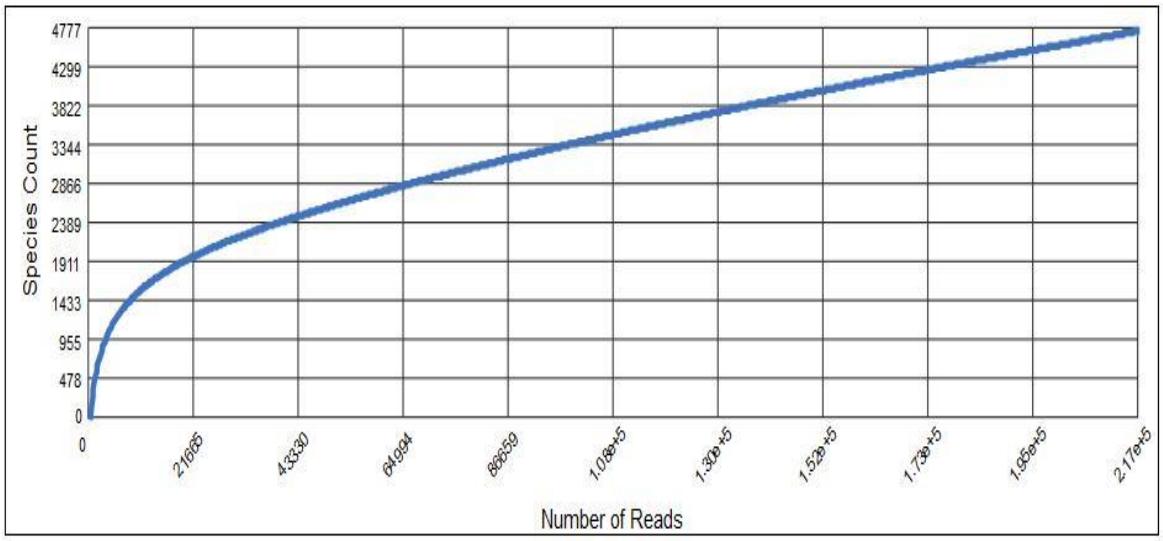


Fig. 2.23: Rarefaction curve of metagenome. Rarefaction plot shows a curve of annotated species richness. This curve is a plot of the total number of distinct species annotations as a function of the number of sequences sampled.

2.3.6.7. Alpha Diversity

The above image shows the range of α-diversity values in sample metagenome. The min, max, and mean values are shown, with the standard deviation ranges (σ and 2σ) in different shades. The α-diversity of this metagenome is shown in red.

Alpha diversity summarizes the diversity of organisms in a sample with a single number. The alpha diversity of annotated samples can be estimated from the distribution of the species-level annotations. Annotated species richness is the number of distinct species annotations in the combined MG-RAST dataset. Shannon diversity is an abundance-weighted average of the logarithm of the relative abundances of annotated species. The species-level annotations are from all the annotation source databases used by MG-RAST.

2.3.6.8. Sequence Length Histogram and GC Distribution

Sequence Length Histogram

The histograms below show the distribution of sequence lengths in basepairs for this amplicon. Each position represents the number of sequences within a length bp range. The data used in these graphs are based on raw upload (Fig. 2.24) and post QC sequences (Fig. 2.25).

a) Raw upload

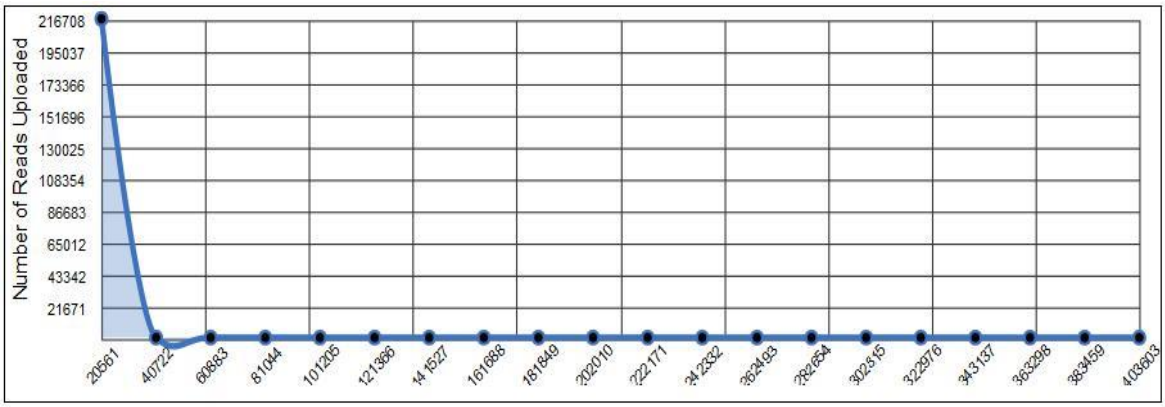


Fig. 2.24: metagenome-Sequence length histogram in base pairs of raw data.

b) Post QC

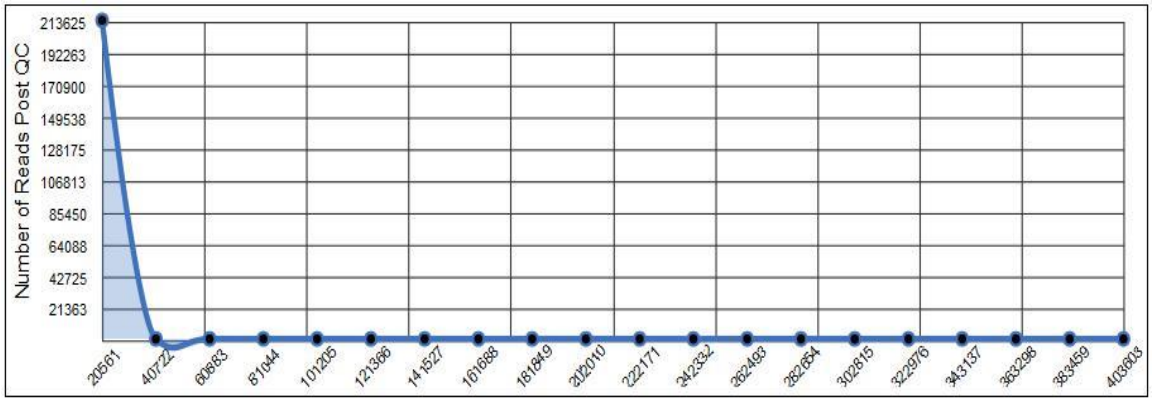


Fig.2.25: metagenome-Sequence length histogram in base pairs of QC filtered data.

Sequence GC Distribution

The histograms below show the distribution of the GC percentage for this amplicon dataset. Each position represents the number of sequences within a GC percentage range. The data used in these graphs is based on raw upload (Fig. 2.26) and post QC sequences (Fig. 2.27).

a) Raw Upload

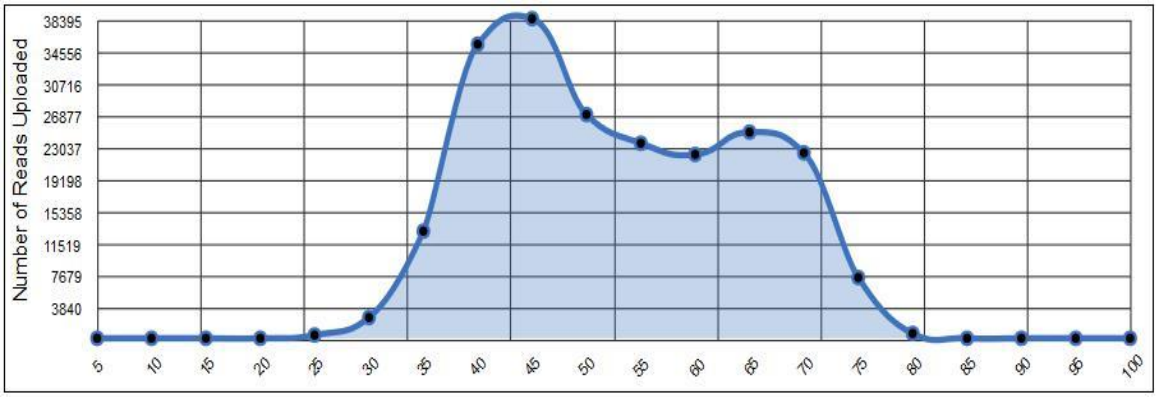


Fig. 2.26: metagenome-Sequence GC distribution of raw data.

b) Post QC

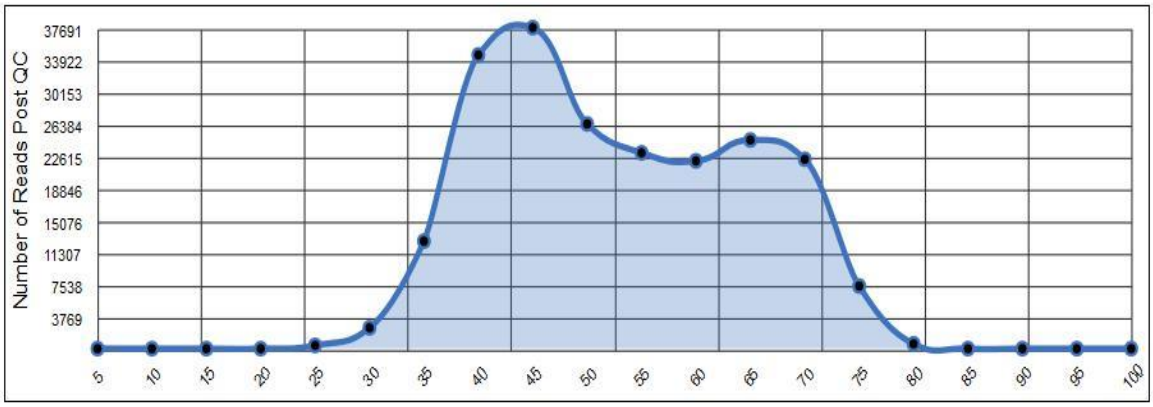


Fig. 2.27: metagenome-Sequence GC distribution of QC filtered data.

2.3.7. Whole genome sequencing (WGS) and metabolic analysis of four novel bacterial strains isolated from gut of *Eisenia fetida*: A necessary corroboration with the metagenomic analyses

2.3.7.1. Isolation of genomic DNA

Genomic DNA extraction, purification and quantification from each each of the four strains ET03^T, EPG1^T, EAG2^T and EAG3^T were carried out by following methods described by Furlong *et al.,* (2002) with modifications. Qualitative and quantitative analysis of the gDNA of the four strains was performed on agarose gel (Fig 2.28) and nanodrop respectively (Fig 2.28).

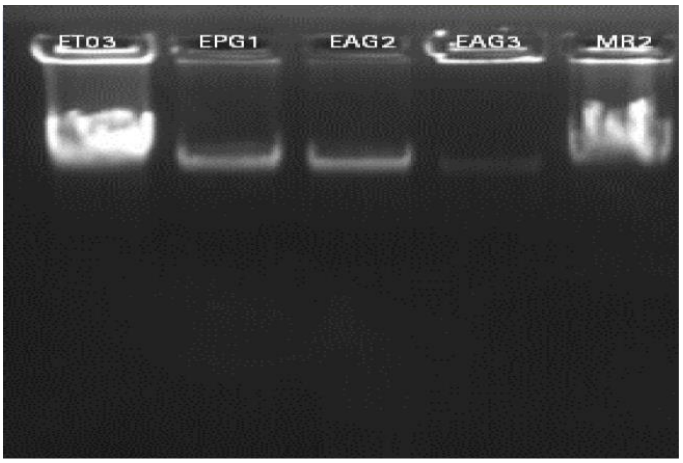


Fig 2.28: Qualitative analysis of genomic DNA of the four strains on agarose gel

Table 2.4: Quantitative analysis of genomic DNA of the four strains on nanodrop

Sr. No.	Sample ID	Concentration (ng/µl)	NanoDrop OD $A_{260/280}$	NanoDrop OD $A_{260/230}$	Remark
1	ET03	524	1.81	2.1	QC Pass
2	EPG1	250	1.77	2.0	QC Pass
3	EAG2	275	1.67	1.97	QC Pass
4	EAG3	114	1.97	1.89	QC Pass

2.3.7.2 Sequencing, Assembly and Gene prediction.

The QC passed samples were used to prepare paired-end (PE) library using Illumina TrueSeq Nano DNA library preparation kit generating mean fragment leanth of 436bp, 430bp, 439bp and 468bp for the isolates ET03, EPG1, EAG2 and EAG3 respectively. Sequencing of the libraries on NextSeq 500 on 2 x 150bp chemistry generated raw data

which were then processed using Trimmomatic v0.35 generating high-quality clean reads as mentioned in Table.2.5.

Table.2.5: High-quality clean reads of the four strains.

Sl. No.	Sample	No. of reads	Total no. of bases	Total data in GB
1	EAG2	3,398,633	1,003,598,780	1.0
2	EAG3	3,580,745	1,061,770,079	1.1
3	EPG1	3,744,675	1,118,323,081	1.1
4	ET03	3,974,770	1,176,381,672	1.1

de-novo assembly of the high- quality reads using SPAdes genome assembler arranged the genomes in scaffolds as mentioned in the Table.2.6.

Table.2.6: Scaffolds of the four strains.

Sl. No.	Sample	Scaffolds	Size of assembly	Average size of scaffolds	N50 (bp)
1	EAG2	190	7,180,376	37,791	84,238
2	EAG3	57	3,821,314	67,041,	266,364
3	EPG1	43	6,025,698	140,133	389,527
4	ET03	21	2,648,915	126,139	894,072

Gene prediction by Prokka from final scaffolds deciferred 2,670, 5,671, 6182 and 3,722 genes of the strains ET03^{T}, EPG1^{T}, EAG2^{T} and EAG3^{T} respectively.

2.3.7.3.Phylogenetic analysis (AAF), Average nucleotide identity (ANI) and Genome-to-genome direct comparison (GGDC).

An assembly and alignment-free (AAF) method (Fan *et al.*, 2015) was used to construct phylogeny from next-generation sequencing data. The calculation of the BLAST-based average nucleotide identity (ANI) score was done using the JSpeciesWS program with the default parameters (Richter *et al*., 2015). Genome-to-genome direct comparison (GGDC) analyses were performed using all three equations in the GGDC program, version 2.1 (Meier-Kolthoff *et al.*, 2014).

The draft genome of strain ET03^T includes 2644068 nucleotide bases distributed in 14 scaffolds (N_{50} 894072). Based on whole- genome data, the DNA G+C content of ET03^T was calculated as 42.9 mol%. The difference between G + C content of the strain ET03^T and any of the five previously reported species was > 5%. Hence there remains no doubt about the novelty of the species and the reason for this inference has been detailed in the method. A total of 2728 genes were predicted (using NCBI pipeline) of which, 2664 were protein-coding genes including 121*de-novo* genes with no blast hit. Of the protein-coding genes, there were at least six genes coding for several dioxygenases [2-nitropropane dioxygenase (Ac. No. PJK17992), biphenyl 2,3-dioxygenase (Ac. No. PJK16435), glyoxalase/bleomycin resistance/estradiol dioxygenase (Ac. No. PJK16272) and three ring-cleaving dioxygenases (Ac. No.s PKJ15803, PKJ15871, and PJK15941]. These microbial genes are responsible for degradation of biphenyl like PCB and several xenobiotic aromatic compound in the environment. The phylogenetic tree constructed from next-generation sequencing data using AAF method depicts distinction of ET03^T from its taxonomic neighbours. Average nucleotide identity (ANI) scores generated during global comparisons of the genome sequence (Kim *et al.*, 2014) of ET03^{T}with previously deposited whole genome sequences in databases indicates sufficient distance from *Bhargavaea cecembensis* T14, *Jeotgalibacillus malaysiensis* D5, *Planomicrobium glaciei* UCD-HAM, *Solibacillus silvestris* StLB046, *Paenisporosarcin*a sp. HGH0030, *Planococcus kocurii* ATCC 43650, *Planococcus maritimus* MKU009, *Planococcus plakortidis* DSM 2399739, *Planococcus rifietoensis* M8 and *Sporosarcina psychrophila* DSM 6497which supports the findings of 16S rRNA gene phylogeny. The ANI score for strain comparisons between EAG3^T and *Paenisporosarcin*a sp. HGH0030 was the highest 68.63% *(coverage= 38.82%)* which is far below 95- 96 % cut-off value for novel species determination by this approach. The GGDC calculations with BLAST+ for strain EAG3^T and other taxonomic neighbours Family *Planococcaceae* gave DNA–DNA homology values of <30% by any of the three models used, assuring sufficient distance among the genomes taken for the analysis.

The draft genome of strain EAG2^T includes 7180290 bases distributed in 190 scaffolds (N50 84238). Based on whole- genome data, the DNA G+C content of EAG2^T was calculated as 73.1 mol %. A total of 6454 genes were predicted (using NCBI pipeline)

Table 2.7: Average nucleotide identity (ANI) matrix analyses of genomic DNA of ET03^T and ten other available whole genomes of Family *Planococcaceae*. Values in parentheses indicate the percentage (%) of alignment.

	ET 03	***Bhargavaea cecembensis* T14**	***Jeotgalibacillus malaysiensis* D5**	***Planomicrobium glaciei* UCD-HAM**	***Solibacillus silvestris* StLB046**	***Paenisporosarcina* sp. HGH0030**	***Planococcus kocurii* ATCC 43650**	***Planococcus maritimus* MKU009**	***Planococcus plakortidis* DSM 23997**	***Planococcus rifietoensis* M8**	***Sporosarcina psychrophila* DSM 6497**
ET03	*	66.33 *(30.59)*	66.90 *(31.23)*	67.44 *(39.29)*	67.71 *(33.84)*	68.63 *(38.82)*	67.48 *(39.30)*	67.46 *(37.26)*	67.37 *(36.99)*	67.24 *(37.84)*	67.06 *(35.97)*
***Bhargavaea cecembensis* T14**	66.36 *(25.75)*	*	66.13 *(23.35)*	67.40 *(36.82)*	65.04 *(27.02)*	65.74 *(29.38)*	65.55 *(29.88)*	67.03 *(32.81)*	67.86 *(34.83)*	67.47 *(34.76)*	66.33 *(35.43)*
***Jeotgalibacillus malaysiensis* D5**	67.70 *(21.48)*	67.05 *(19.12)*	*	67.39 *(24.41)*	67.30 *(20.47)*	67.26 *(22.11)*	67.27 *(23.68)*	67.18 *(24.05)*	67.14 *(23.58)*	67.17 *(24.42)*	67.05 *(22.60)*
***Planomicrobium glaciei* UCD-HAM**	67.44 *(27.85)*	67.47 *(30.72)*	66.63 *(25.38)*	*	66.99 *(26.61)*	68.65 *(32.22)*	72.94 *(51.99)*	72.19 *(49.03)*	72.57 *(49.22)*	72.70 *(50.79)*	67.65 *(32.45)*
***Solibacillus silvestris* StLB046**	68.53 *(23.55)*	66.21 *(22.17)*	67.41 *(21.17)*	67.74 *(26.35)*	*	69.04 *(27.78)*	68.30 *(26.39)*	67.74 *(23.78)*	67.55 *(23.34)*	67.60 *(24.03)*	67.72 *(30.61)*
***Paenisporosarcina* sp. HGH0030**	68.61 *(29.37)*	66.00 *(26.87)*	66.95 *(24.77)*	68.74 *(35.13)*	68.60 *(30.30)*	*	69.09 *(35.58)*	68.21 *(32.50)*	67.86 *(31.16)*	68.07 *(32.24)*	68.89 *(37.57)*
***Planococcus kocurii* ATCC 43650**	68.13 *(31.10)*	66.33 *(28.08)*	67.15 *(27.92)*	73.22 *(58.04)*	68.11 *(29.39)*	69.29 *(36.42)*	*	72.23 *(53.89)*	71.58 *(52.84)*	72.03 *(54.40)*	68.37 *(35.12)*
***Planococcus maritimus* MKU009**	67.22 *(31.54)*	66.94 *(32.25)*	66.45 *(28.82)*	72.20 *(57.81)*	66.74 *(27.84)*	67.94 *(35.31)*	71.82 *(57.28)*	*	83.62 *(80.23)*	84.60 *(83.88)*	67.02 *(33.71)*
***Planococcus plakortidis* DSM 2399739**	67.93 *(31.49)*	68.28 *(35.51)*	67.20 *(29.33)*	72.94 *(58.71)*	67.21 *(28.75)*	68.12 *(35.04)*	71.66 *(56.92)*	83.79 *(81.12)*	*	87.63 *(85.45)*	67.49 *(34.00)*
***Planococcus rifietoensis* M8**	67.75 *(29.89)*	67.01 *(28.46)*	72.96 *(56.31)*	67.18 *(27.12)*	68.13 *(33.27)*	72.01 *(54.51)*	84.62 *(78.91)*	87.36 *(80.25)*	67.60 *(32.73)*	*	68.03 *(32.30)*
***Sporosarcina psychrophila* DSM 6497**	67.70 *(21.80)*	67.14 *(25.24)*	66.75 *(20.73)*	68.13 *(27.65)*	67.52 *(26.68)*	69.17 *(29.57)*	68.08 *(27.61)*	67.62 *(24.58)*	67.66 *(23.86)*	67.56 *(25.32)*	*

Table 2.8: Genome-to-Genome Distance between *strain* ET03^T and other taxonomic neighbours of Family *Planococcaceae* calculated using GGDC 2.1.

*The reference genomes used in this study were of the bacteria **1.** *Bhargavaea cecembensis* T14, **2.** *Jeotgalibacillus malaysiensis* D5, **3.** *Planomicrobium glaciei* UCD-HAM, **4.** *Solibacillus silvestris* StLB046, **5.** *Paenisporosarcin*a sp. HGH0030, **6.** *Planococcus kocurii* ATCC 43650, **7.** *Planococcus maritimus* MKU009, **8.** *Planococcus plakortidis* DSM 2399739, **9.** *Planococcus rifietoensis* M8 and **10.** *Sporosarcina psychrophila* DSM 6497**.**

No	Reference genome*	Formula 1				Formula 2				Formula 3				G+C difference
		DDH	Model C.I.	Distance	Prob. DDH ≥70%	DDH	Model C.I.	Distance	Prob. DDH ≥70%	DDH	Model C.I.	Distance	Prob DDH ≥70%	
1	**LQNT00000000.1**	12.5	[9.9 - 15.8%]	0.9977	0	25.6	[23.2 - 28%]	0.17	0.01	12.9	[10.6 - 15.7%]	0.9981	0	9.84
2	**CP009416.1**	12.9	[10.2 - 16.2%]	0.9766	0	25	[22.7 - 27.5%]	0.1738	0.01	13.3	[11 - 16.1%]	0.9806	0	0.95
3	**LGAF00000000.1**	12.6	[9.9 - 15.8%]	0.9951	0	14.5	[12.5 - 16.7%]	0.3011	0	13	[10.6 - 15.7%]	0.9966	0	4.1
4	**AP012157.1**	13	[10.3 - 16.2%]	0.9746	0	27.8	[25.4 - 30.3%]	0.1552	0.04	13.4	[11 - 16.1%]	0.9785	0	5.06
5	**AGEQ00000000.1**	13.1	[10.4 - 16.4%]	0.9658	0	20.2	[18 - 22.6%]	0.2178	0	13.5	[11.1 - 16.2%]	0.9732	0	4.41
6	**NZ_CP013661.2**	13.1	[10.4 - 16.4%]	0.9687	0	21.5	[19.2 - 23.9%]	0.2044	0	13.4	[11.1 - 16.2%]	0.9751	0	2.74
7	**NZ_LTZG00000000.1**	12.8	[10.1 - 16.1%]	0.9822	0	19.3	[17.1 - 21.7%]	0.2281	0	13.2	[10.8 - 15.9%]	0.9863	0	3.55
8	**NZ_CP016539.2**	13	[10.3 - 16.3%]	0.9743	0	28.7	[26.4 - 31.2%]	0.1493	0.06	13.4	[11 - 16.1%]	0.9781	0	6.32
9	**CP013659.2**	12.9	[10.2 - 16.2%]	0.9792	0	26.1	[23.7 - 28.5%]	0.1664	0.02	13.3	[10.9 - 16%]	0.9827	0	4.79
10	**NZ_CP014616.1**	12.9	[10.2 - 16.2%]	0.9776	0	23.5	[21.3 - 26%]	0.1857	0	13.3	[10.9 - 16%]	0.9818	0	3.32

The draft genome of strain EPG1^T is 5.7Mb, distributed in 37 scaffolds. A total of 5710 genes were predicted (using NCBI pipeline) including 5580 protein coding genes, 53 RNA coding genes and 77 pseudogenes. 60 CDS coded for proteins with no blast hit. The DNA G+C content of strain EPG1^T is 68.4 mol %. The phylogenetic tree constructed from next-generation sequencing data using AAF method depicts distinction of EPG1^T from its taxonomic neighbours. Average nucleotide identity (ANI) scores generated during global comparisons of the genome sequence of EPG1^T with previously deposited whole genome sequences of *M. asiaticum* 1245335. *M. chubuense* DSM 44219, *M. conceptionense* ACS1953, *M. cosmeticum* DSM 44829, *M. fortuitum* CT6, *M. gilvum* PYR-GCK, *M. goodii* X7B, *M. mucogenicum* 1127319.6, *M. immunogenum* H076, *M. rufum* JS14, *M. vanbaalenii* PYR-1 and *M. parafortuitum* DSM43528 indicates sufficient distance and supports the findings of 16S rRNA gene phylogeny. The ANI score for strain comparisons between EPG1^T and *M. gilvum* PYR-GCK was the highest (82.42%) which was below 95-96 % cut-off value for novel species determination by this approach. The GGDC calculations with BLAST+ for strain EPG1^T and other taxonomic neighbours of genus *Mycobacterium* gave highest DNA–DNA homology values of 56.7% by model 2, and with *M. parafortuitum* DSM43528 which was far below cut-off value (70 %) for novel species delineation.

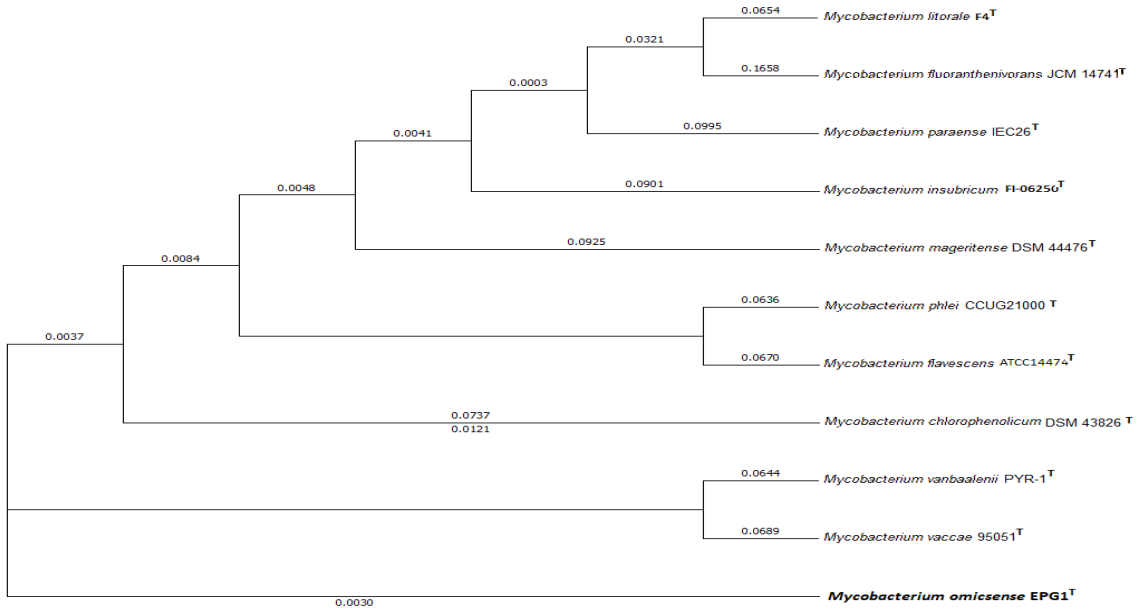


Fig. 2.29: Phylogenetic relationship of strain EPG1^T (*Mycobacterium omicsense*, sp. nov.) and closely related genomes. An assembly and alignment-free (AAF) method was used to construct phylogeny from next-generation sequencing data.

Table 2.9: Average nucleotide identity (ANI) matrix analyses of genomic DNA of EPG1^T and eleven other available whole genomes of Mycobacteria. Values in parentheses indicate the percentage (%) of alignment.

	***M. omicsense* EPG1**	***M. asiaticum* 1245335.1**	***M. chubuense* DSM 44219**	***M. conceptionense* ACS1953**	***M. cosmeticum* DSM 44829**	***M. fortuitum* CT6**	***M. gilvum* PYR-GCK**	***M. goodii* X7B**	***M. mucogenicum* 1127319.6**	***M. immunogenum* H076**	***M. rufum* JS14**	***M. vanbaaleni* PYR-1**
***M. omicsense* EPG1**	*	73.68 *(43.11)*	79.08 *(62.40)*	75.64 *(53.05)*	75.78 *(53.30)*	75.35 *(52.17)*	82.42 *(70.98)*	76.30 *(54.64)*	74.77 *(45.06)*	71.49 *(33.62)*	79.05 *(62.32)*	80.94 *(67.09)*
***M. asiaticum* 1245335.1**	73.58 *(45.29)*	*	73.49 *(45.20)*	73.61 *(45.73)*	73.67 *(44.12)*	73.41 *(45.94)*	73.26 *(43.58)*	73.70 *(43.97)*	73.05 *(42.67)*	70.95 *(32.58)*	73.62 *(44.79)*	73.47 *(44.51)*
***M. chubuense* DSM 44219**	79.29 *(62.38)*	73.98 *(42.18)*	*	75.91 *(52.23)*	76.38 *(55.61)*	75.65 *(51.50)*	79.25 *(60.67)*	76.50 *(54.63)*	75.15 *(45.60)*	71.72 *(33.61)*	88.20 *(83.28)*	80.02 *(63.94)*
***M. conceptionense* ACS1953**	75.42 *(51.82)*	73.66 *(42.16)*	75.49 *(51.58)*	*	76.73 *(55.32)*	85.82 *(76.75)*	75.33 *(51.00)*	78.59 *(60.58)*	75.34 *(49.02)*	72.08 *(36.37)*	75.83 *(51.57)*	75.72 *(51.80)*
***M. cosmeticum* DSM 44829**	75.77 *(50.57)*	73.98 *(39.30)*	76.02 *(53.35)*	76.83 *(53.65)*	*	76.30 *(51.91)*	75.68 *(47.91)*	76.51 *(55.28)*	76.06 *(49.24)*	72.14 *(34.95)*	76.37 *(53.52)*	75.89 *(50.72)*
***M. fortuitum* CT6**	75.24 *(52.10)*	73.46 *(43.67)*	75.34 *(51.61)*	85.77 *(78.78)*	76.10 *(54.84)*	*	75.14 *(50.22)*	78.30 *(61.15)*	74.96 *(51.37)*	71.41 *(37.29)*	75.39 *(51.11)*	75.35 *(52.09)*
***M. gilvum* PYR-GCK**	82.69 *(70.75)*	73.56 *(41.44)*	79.28 *(61.22)*	75.63 *(53.06)*	76.05 *(51.63)*	75.45 *(51.32)*	*	76.16 *(50.61)*	74.75 *(43.46)*	71.64 *(32.67)*	79.87 *(62.73)*	82.25 *(71.29)*
***M. goodii* X7B**	75.89 *(47.22)*	73.64 *(36.26)*	75.84 *(48.22)*	78.37 *(54.35)*	76.09 *(50.92)*	78.18 *(52.99)*	75.75 *(43.23)*	*	74.91 *(40.55)*	71.55 *(31.24)*	75.93 *(48.00)*	76.02 *(46.39)*
***M. mucogenicum* 1127319.6**	74.70 *(45.12)*	73.35 *(40.11)*	74.95 *(46.25)*	75.61 *(49.79)*	76.25 *(51.54)*	75.27 *(50.71)*	74.59 *(43.58)*	75.18 *(46.78)*	*	71.78 *(36.85)*	74.92 *(45.67)*	74.84 *(44.77)*
***M. immunogenum* H076**	71.30 *(35.01)*	70.87 *(33.15)*	71.41 *(34.93)*	72.02 *(38.91)*	72.15 *(37.57)*	71.39 *(38.46)*	71.46 *(33.41)*	71.58 *(37.35)*	71.49 *(38.12)*	*	71.99 *(35.38)*	71.42 *(34.31)*
***M. rufum* JS14**	79.34 *(60.11)*	74.04 *(41.02)*	88.18 *(80.06)*	76.32 *(50.83)*	76.69 *(54.19)*	75.80 *(49.25)*	79.78 *(59.30)*	76.64 *(52.90)*	75.16 *(43.80)*	72.38 *(33.06)*	*	80.16 *(61.78)*
***M. vanbaaleni* PYR-1**	80.86 *(63.25)*	73.73 *(39.36)*	79.68 *(60.71)*	75.78 *(50.16)*	75.95 *(50.51)*	75.55 *(49.27)*	81.63 *(66.19)*	76.27 *(50.35)*	74.64 *(42.60)*	71.63 *(30.99)*	79.77 *(60.74)*	*

Table 2.10: Average nucleotide identity (ANI) matrix analyses of whole genome of EAG2^T and nine other available whole genomes of genus *Streptomyces*. Values in parentheses indicate the percentage (%) of alignment.

*The reference genomes used in this study were of the bacteria 1. *M. asiaticum* 1245335.1, 2. *M. chubuense* DSM 44219, *3. M. conceptionense* ACS1953, 4. *M. cosmeticum* DSM 44829, 5. *M. fortuitum* CT6, 6. *M. gilvum* PYR-GCK, 7. *M. goodii* X7B, 8. *M. mucogenicum* 1127319.6, 9. *M. immunogenum* H076, 10. *M. rufum* JS14, 11. *M. vanbaalenii* PYR-1 and 12. *M. parafortuitum* DSM43528.

No		Formula 1				Formula 2				Formula 3				G+C difference
	Reference genome*	DDH	Model C.I.	Distance	Prob DDH ≥ 70 %	DDH	Model C.I.	Distance	Prob DDH ≥ 70 %	DDH	Model C.I.	Distance	Prob DDH ≥ 70 %	
1	NZ_LZLR00000000.1	12.6	[9.9 - 15.8%]	0.9963	0	17.7	[15.6 - 20.1 %]	0.2481	0	13	[10.6 - 15.7 %]	0.9972	0	1.69
2	NZ_MVHO00000000.1	22.4	[19.1 - 26%]	0.6523	0	22.8	[20.5 - 25.2 %]	0.1923	0	21.4	[18.7 - 24.5 %]	0.7192	0	1.18
3	LZHX00000000.1	13.2	[10.5 - 16.5%]	0.9632	0	20.1	[17.9 - 22.5 %]	0.2189	0	13.5	[11.1 - 16.3 %]	0.9713	0	1.81
4	CCBB010000001.1	19.5	[16.4 - 23.1%]	0.7247	0	20.9	[18.7 - 23.3 %]	0.2102	0	18.9	[16.2 - 21.9 %]	0.7826	0	0.06
5	CP011269.1	18.9	[15.8 - 22.5%]	0.7419	0	20.3	[18.1 - 22.7 %]	0.2163	0	18.4	[15.7 - 21.4 %]	0.7977	0	2.11
6	HG964481.1	17.2	[14.2 - 20.8%]	0.7955	0	20.4	[18.2 - 22.8 %]	0.2155	0	17	[14.4 - 20%]	0.8395	0	1.79
7	NZ_LZLC00000000.1	12.7	[10 - 16%]	0.9866	0	20.9	[18.7 - 23.3 %]	0.2103	0	13.1	[10.8 - 15.9 %]	0.9894	0	1.06
8	CP000656.1	40.5	[37.1 - 44%]	0.3874	1.91	26.5	[24.1 - 28.9 %]	0.1638	0.02	36	[33 - 39%]	0.4877	0.03	0.41
9	CP012150.1	20.5	[17.3 - 24.1%]	0.6989	0	20.7	[18.5 - 23.1 %]	0.2123	0	19.7	[16.9 - 22.7 %]	0.7628	0	0.76
10	JROA00000000.1	28.9	[25.5 - 32.5%]	0.5302	0.06	23	[20.7 - 25.4 %]	0.1905	0	26.5	[23.6 - 29.6 %]	0.6196	0	1.04
11	CP000511.1	32	[28.7 - 35.6%]	0.4842	0.18	24.8	[22.5 - 27.3 %]	0.1756	0.01	29.3	[26.4 - 32.4 %]	0.5748	0	0.53
12	MVID00000000.1	44.3	[40.9 - 47.7%]	0.3525	4.38	56.7	[54 - 59.5 %]	0.0575	41.11	46	[43 - 49%]	0.3898	0.93	0.13

including 5991 protein coding genes, 80 RNA coding genes and 383 pseudogenes. 77 CDS coded for proteins with no blast hit. Genome sequence of EAG2^T when compared globally with the whole genome sequences of *S. sampsonii* strain K140, *S. pleuripotens* strain MUSC135 and *S. variegatus* strain NRRL B16380, the average nucleotide identity (ANI) scores generated from it has enabled to calculate the evolutionary distances. The phylogenetic tree thus drawn has further validated the position of EAG2^T with respect to the three species in the 16S rRNA gene phylogeny. As genome sequence of the phylogenetically closest member was not available, we tried to look for the genes in EAG2^T corresponding to the distinctive phenotype displayed by *S. koyangensis*. Since the phylogenetically nearest member of EAG2^T, *Streptomyces koyangensis* VK-A60^T, forms melanin (Lee *et al*., 2005; El-Naggar *et al*., 2017), genome of EAG2^T was searched for any genes for the expression of melanin like *melC1* and *melC2* or *melC* operon under positive control by a multifunctional regulatory protein, AdpA1, as were reported in some species of *Streptomyces* (Undabarrena *et al*., 2017). The genome of EAG2^T is devoid of any such tyrosinase genes (*melC2*) for melanin formation. Again, *S. koyangensis* could not utilize cellulose, but EAGT^2 genome contains three genes related to cellulose catabolic process including cellulose 1,4-beta-cellobiosidase (EEC. 3.2.1.911) that catalyzes hydrolysis of (1→4)-beta-D-glucosidic linkage in cellulose releasing cellobiose from the non-reducing ends of the chains. The strain *Streptomyces koyangensis* VK-A60^T was sensitive to tellurite (could not grow in 0.001% potassium tellurite) (Lee *et al*., 2005) while EAGT^2 possessed three genes coding for tellurite resistance. A total number of 21, 5, 3, and 3 genes were identified in the genome of EAG2^T for urease, nitrate reductase, nitrite reductase, and carbonic anhydrase respectively. Genes encoding NasA and NirBD reductases for reduction of nitrite to ammonia are also present. It has been found in quite a few bacteria that the key enzymes involved in calcification were urease and carbonic anhydrase (Achal and Pan, 2011). In the genome of EAG2^T, there are three genes for cation: proton antiporters, two genes for calcium binding protein, and two genes coding for ammonia channel protein/ ammonia transporters. It has been speculated by the earlier authors that in *Streptomyces*, the genes predisposed for carbonatogenesis include those that are responsible for reduction of nitrate to ammonia, hydrolysis of urea, active transport of calcium ion, and reversible hydration of carbon di-oxide (Maciejewska *et al*., 2017).

Table 2.11: Genome-to-Genome Distance between *strain* EAG2^T and other taxonomic neighbours of genus *Streptomyces* calculated using GGDC 2.1.

EAG2^T has 94.80% *(alignment=75.95%) as the* highest identity (<95% cut off for species delineation) with S. *albus* SM 254.

	EAG2	***S. afghaniensis* 772**	***S. albus* SM254**	***S. ambofaciens* ATCC 23877**	***S. coelicolor* A3(2)**	***S. griseus* S4-7**	***S. nodosus* ATCC 14899**	***S. pluripotens* MUSC 135**	***S. venezuelae* ATCC 15439**	***S. vitaminophilus* ATCC 31673**
EAG2	*	77.44 *(40.29)*	94.65 *(76.20)*	78.10 *(43.35)*	77.81 *(43.57)*	77.97 *(41.66)*	76.91 *(39.60)*	76.45 *(36.83)*	77.32 *(40.11)*	74.00 *(29.74)*
***S. afghaniensis* 772**	77.04 *(35.80)*	*	77.00 *(36.01)*	82.34 *(50.50)*	81.83 *(51.07)*	77.27 *(38.56)*	80.07 *(42.04)*	80.07 *(41.66)*	77.35 *(40.75)*	74.03 *(28.03)*
***S. albus* SM254**	94.80 *(75.95)*	77.42 *(40.56)*	*	77.91 *(43.30)*	77.80 *(43.83)*	77.95 *(43.06)*	77.02 *(39.30)*	76.39 *(36.19)*	77.39 *(39.95)*	74.11 *(30.33)*
***S. ambofaciens* ATCC 23877**	77.47 *(40.58)*	82.64 *(52.94)*	77.27 *(40.59)*	*	86.85 *(62.39)*	77.38 *(41.67)*	79.71 *(44.20)*	79.81 *(42.08)*	77.23 *(42.28)*	73.68 *(30.66)*
***S. coelicolor* A3(2)**	77.25 *(37.64)*	81.95 *(49.89)*	77.41 *(37.78)*	86.77 *(57.28)*	*	77.28 *(39.68)*	79.86 *(42.09)*	79.68 *(40.35)*	77.35 *(40.00)*	77.25 *(37.64)*
***S. griseus* S4-7**	77.48 *(42.10)*	77.54 *(42.80)*	77.60 *(43.18)*	77.67 *(44.24)*	77.72 *(44.92)*	*	76.91 *(39.24)*	76.34 *(37.71)*	78.54 *(46.63)*	77.48 *(42.10)*
***S. nodosus* ATCC 14899**	76.45 *(38.95)*	80.25 *(47.55)*	76.46 *(39.43)*	79.96 *(47.35)*	79.95 *(49.11)*	76.75 *(39.51)*	*	78.94 *(43.40)*	76.61 *(40.60)*	76.45 *(38.95)*
***S. pluripotens* MUSC 135**	76.16 *(37.89)*	80.48 *(50.80)*	76.09 *(37.58)*	80.14 *(48.09)*	79.92 *(49.53)*	76.33 *(39.60)*	79.17 *(45.44)*	*	76.42 *(41.99)*	73.08 *(29.23)*
***S. venezuelae* ATCC 15439**	76.64 *(33.86)*	77.14 *(39.31)*	76.64 *(34.10)*	77.21 *(38.47)*	77.19 *(39.48)*	78.16 *(40.20)*	76.46 *(34.77)*	76.20 *(33.90)*	*	73.33 *(26.92)*
***S. vitaminophilus* ATCC 31673**	74.04 *(33.03)*	74.08 *(34.81)*	74.06 *(33.10)*	74.22 *(35.26)*	74.13 *(36.52)*	74.04 *(34.18)*	74.09 *(33.13)*	73.43 *(30.93)*	73.79 *(34.79)*	*

Table.2.12: Genome-to-Genome Distance between *strain* EAG2^T and other taxonomic neighbours of genus *Streptomyces* calculated using GGDC 2.1.

*The reference genomes (from Sl. No. 1 to **9**) used in this study were of the bacteria *S. pluripotens strain* MUSC 135, *Streptomyces.* sp. Tu6071, *S. ambofaciens strain* DSM 40697, *S. coelicolor* A3(2), *S. variegatus strain* NRRL B-16380, *S. vietnamensis strain* GIM4.0001**,** *S. vitaminophilus* ATCC 31673, *S. sampsonii* KJ40 and *S. griseus* S4-7 respectively**.**

No	Reference genome*	Formula 1				Formula 2				Formula 3				G+C difference
		DDH	Model C.I.	Distance	Prob. DDH >= 70 %	DDH	Model C.I.	Distance	Prob. DDH >= 70 %	DDH	Model C.I.	Distance	Prob. DDH >= 70 %	
1	CP021080.1	18.6	[15.4 - 22.1 %]	0.7528	0	22	[19.7 - 24.4 %]	0.1996	0	18.2	[15.6 - 21.2 %]	0.8022	0	3.26
2	CM001165.1	18.9	[15.7 - 22.4 %]	0.7438	0	22.9	[20.6 - 25.3 %]	0.1914	0	18.6	[15.9 - 21.6 %]	0.7929	0	0.08
3	CP012949.1	21.9	[18.6 - 25.5 %]	0.6639	0	22.5	[20.2 - 24.9 %]	0.195	0	21	[18.2 - 24.1 %]	0.7295	0	0.87
4	AL645882.2	21.2	[17.9 - 24.8 %]	0.6817	0	22.6	[20.4 - 25.1 %]	0.1934	0	20.4	[17.7 - 23.5 %]	0.7433	0	1.08
5	JYJH010000021.1	12.9	[10.2 - 16.2 %]	0.9767	0	27	[24.7 - 29.5 %]	0.16	0.03	13.3	[11 - 16.1 %]	0.9804	0	3.07
6	CP010407.1	19.5	[16.3 - 23.1 %]	0.7264	0	23	[20.7 - 25.5 %]	0.1901	0	19.1	[16.4 - 22.1 %]	0.7784	0	1.11
7	LLZU01000038.1	13	[10.3 - 16.3 %]	0.9737	0	21.8	[19.5 - 24.2 %]	0.2015	0	13.3	[11 - 16.1 %]	0.979	0	1.32
8	KV861362.1	78.9	[74.9 - 82.4 %]	0.1378	89.87	62.8	[59.9 - 65.7 %]	0.0468	61.32	78.5	[75 - 81.5 %]	0.1782	93.9	0.23
9	JYBE010000014.1	13	[10.3 - 16.3 %]	0.9739	0	22.3	[20 - 24.7 %]	0.197	0	13.3	[11 - 16.1 %]	0.9791	0	1.7

The realistic species concept for Bacteria is in due course based on DNA-DNA hybridization (DDH). Published DDH results have shown that the three closest species *of S. koyangensis* VK-A60^T were *S. griseus* IFO12875^T (68.5%), *S. limosus* DSM40131^T (66.8%), and *S. sampsonii* ATCC25495^T (64.0%) (Lee *et al*., 2005). Hybridization of *S. limosus* DSM40131^T DNA with DNA of *S. griseus* IFO12875^T or *S. canescens* DSM40001^T or *S. coelicolor* DSM 40233^T or *S. felleus* DSM40130^T or *S. somaliensis* DSM40267 was below 70%, but with DNA of *S. sampsonii* ATCC25495^T or *S. odorifer* DSM40347^T was above 90%. Hence, from this DDH results, it may be concluded that *S. limosus* DSM40131^T, *S. sampsonii* ATCC25495^T, and *S. odorifer* DSM40347^T are synonymous. Again, hybridization of *S. sampsonii* ATCC25495^T DNA with DNA of *S. griseus* IFO12875^T or *S. canescens* DSM40001^T or *S. coelicolor* DSM 40233^T or *S. odorifer* DSM40347^T or *S. felleus* DSM40130^T or *S. somaliensis* DSM40267 was below 70%, but DDH with *S. limosus* DSM40131^T DNA was 95.1%. The DDH results taken together, leaves no doubt about the synonymy of *S. sampsonii* ATCC25495^T and *S. limosus* DSM40131^T. On the other hand high DDH scores (> 70 %) were obtained for the pairs: (i) *S. odorifer* DSM40347^T and *S. canescens* DSM40001^T (70.4%) or *S. coelicolor* DSM 40233^T (80.0%); (ii) *S. felleus* DSM40130^T and *S. odorifer* DSM40347^T (85.4%) or *S. limosus* DSM40131^T (81.9%) or *S. coelicolor* DSM 40233^T (77%); (iii) *S. somaliensis* DSM40267 and *S. limosus* DSM40131^T (89%) or *S. odorifer* DSM40347^T (80.0%). It was clear from the DDH results that VK-A60^T has a distinct status of a species of the genus *Streptomyces* (*S. koyangensis*) (Labeda *et al*., 2017). In the present study, the strain EAG2^T is set out as a different clade with closest relationship with *S. koyangensis* KCCM 10555^T in the 16S phylogenetic tree. In order to obtain an estimate of the overall similarity between genomes of the two different strains, representing two different species, we have taken the advantage of the recent technological progress in the area of genome sequencing and bioinformatics methods to replace the wet-lab DDH by in-silico genome-to-genome comparison via web service (http://ggdc.dsmz.de/) for genome-based species delineation . Results of genome-to-genome distance between EAG2^T and nine *Streptomyces* species (including *S. griseus* which was found closest to *S. koyangensis* KCCM 10555^T) calculated using formula 2 (identities/ HSP length) were presented. DDH values obtained for EAG^T

and *S. griseus* S4-7 (22.3%) or *S. pluripotens* MUSC 135 (22 %) or *Streptomyces* sp. Tu 6071 (22.9%) or *S. ambofaciens* DSM 40697 (22.5 %) or *S. coelicolor* A3(2) (22.6 %) or *S. variegatus* NRRLB-16380 (27 %) or *S. vietnamensis* GIM 4.0001 (23 %) or *S. vitaminophilus* ATCC 31673^T (21.8 %) have ruled out the possibility of EAG2^T being a synonymous species. The closest genome distance of EAG2^T was 0.0468 (DDH = 62.8 %) with *S. sampsonii* KJ40 has established EAG^T as a novel species. Due to unavailability of the genome of any of the *S. koyangensis* strains, we relied on the comparison of the reported DDH values of *S. koyangensis* VK-A60^T (with its neighboring strains) with the DDH values derived from genome- to- genome distance values of EAG2^T (with its neighboring *Streptomyces* strains, common to the ones taken for *S. koyangensis* VK-A60^T). DDH between EAG2^T and *S. griseus* S4-7 is 22.3% while it is 68.5% between *S. koyangensis* VK-A60^T and *S. griseus* IFO12875^T. Since both EAG2^T and *S. koyangensis* VK-A60^T have demonstrated comparable DDH values with (i) *S. sampsonii* KJ40 (62.8 %) or *S. sampsonii* ATCC25495^T (64.0%); (ii) *S. coelicolor* A3(2) (22.6 %) or *S. coelicolor* DSM 40233^T (20.8%), it can be assumed that they have a common branching point which is reflected in the phylogenetic tree constructed with 16S rRNA gene sequences**.**

As average nucleotide identity (ANI) is also used as genome relatedness index and substitute for DDH, the process of recognizing novel species, based on the proposed threshold of 98.65%, has been gaining momentum. In the present study, ANI matrix analyses of whole genome of EAG2^T and nine other available whole genomes of genus *Streptomyce*s, *S. albus* SM254, *S. ambofaciens* ATCC 23877, *S. griseus* S4-7, *S. coelicolor* A3(2), *S. afghaniensis* 772, *S. venezuelae* ATCC 15439, *S. nodosus* ATCC 14899, *S. pluripotens* MUSC 135, and *S. vitaminophilus* ATCC 31673 has been presented. The data indicated that none of the strains were synonym species (as per threshold) and the closest species (among the nine distinct species of *Streptomyces*) to EAG2^T was *S. albus* SM254. ANI results presented in this study corroborated with the inference of the earlier MLSA study with the confirmation of status of the 9 species (Labeda *et al*., 2017).

The draft genome of strain EAG3^T is 3.6Mb, distributed in 57 scaffolds. A total of 3811 genes were predicted (using NCBI pipeline) including 3703 protein coding genes, 108 RNA coding genes and 31 pseudogenes. 135 coding DNA sequence (CDS) coded for

proteins with no blast hit. The translated protein sequences of two CDS, EAG3_00904 and EAG3_02364, produced maximum identities of 91% (284/313) and 90% (286/319) with nitronate monooxygenase sequence IDs WP_04670761.1 (length: 327) and WP_049671553.1 (length: 319) respectively of *Bacillus* sp. FJAT-27916. The translated protein sequence of EAG3_00418 produced maximum identities of 92% (451/488) with methylmalonate-semialdehyde dehydrogenase (CoA acylating) sequence ID WP_049670323.1(length: 488) of *Bacillus* sp. FJAT-27916. 3-Nitropropionate (3-NPA), a nitroaliphatic compound, exists in equilibrium with its conjugate base, propionate 3-nitronate (P3N), is found in numerous plants and fungi. Gut bacteria of the herbivores releases 3-NPA (which consequently forms P3N at physiological pH) from the ingested plant parts containing glycosides of 3-NPA (Anderson *et al*.,2005). The mitochondrial electron transport chain and the key enzyme of the Kreb's cycle, succinate dehyrogenase is inhibited by P3N causing acute toxicity (Claus, 1992). Since the earthworm, *E. fetida*, is a voracious dung-eater, similar exposure to P3N is also obvious. Hence, tolerance of *E. fetida* to P3N ought to have been dependent on the gut bacterial capacity to metabolise nitrotoxins. The strain, EAG3^{T}, isolated from gut-content of *E. fetida*, is genetically predisposed to metabolise P3N. Nitronate monooxygenase (NMO) oxidizes P3N to malonic semialdehyde, which can be converted to acetyl-CoA (Francesca *et al*., 2014). NMO is an FMN-dependent enzyme that uses molecular oxygen to oxidize (anionic) alkyl nitronates and nitroalkanes to the corresponding carbonyl compounds and nitrite. It has been reported that bacterial methymalonate-semialdehyde dehydrogenase (CoA acylating) in a two step reaction mechanism catalyses the NAD-dependent oxidation of methylmalonate semialdehyde and malonate semialdehyde into propionyl-CoA and acetyl-CoA respectively (Stines-Chaumeil *et al*., 2006).

Table 2.13: Average nucleotide identity (ANI) matrix analyses of genomic DNA of EAG3 and ten other available whole genomes of Family *Bacillaceae*. Values in parentheses indicate the percentage (%) of alignment.

	EAG3	***Bacillus alcalophilus* ATCC 27647**	***Bacillus amyloliquefaciens* DSM 7**	***Planococcus plakortidis* DSM 23997**	***Planococcus rifietoensis* M8**	***Planomicrobium* sp. ES2**	***Bhargavaea cecembensis* DSE 10**	***Bacillus coagulans* ATCC 7050**	***Bacillus cytotoxicus* NVH 391-98**	***Bacillus methanolicus* PB1**	***Bacillus simplex* SH-B26**
EAG3	*	65.74 *(21.21)*	66.32 *(24.31)*	65.88 *(19.86)*	65.91 *(20.20)*	66.10 *(21.40)*	65.58 *(17.89)*	66.34 *(23.08)*	66.65 *(23.51)*	67.45 *(26.07)*	67.27 *(31.73)*
***Bacillus alcalophilus* ATCC 27647**	65.75 *(18.28)*	*	65.82 *(17.45)*	64.79 *(13.04)*	64.96 *(13.86)*	65.46 *(15.18)*	63.95 *(10.45)*	65.35 *(15.06)*	67.62 *(19.88)*	67.10 *(19.71)*	66.31 *(20.52)*
***Bacillus amyloliquefaciens* DSM 7**	67.37 *(23.74)*	66.55 *(20.45)*	*	67.07 *(19.25)*	66.95 *(19.51)*	67.18 *(19.32)*	66.73 *(18.29)*	67.99 *(23.92)*	67.16 *(24.67)*	68.42 *(24.37)*	67.34 *(29.03)*
***Planococcus plakortidis* DSM 23997**	66.66 *(24.56)*	65.64 *(18.76)*	66.79 *(23.84)*	*	87.63 *(85.45)*	72.83 *(58.01)*	68.44 *(35.36)*	67.51 *(22.66)*	66.28 *(18.78)*	67.03 *(22.28)*	66.42 *(26.25)*
***Planococcus rifietoensis* M8**	66.50 *(23.02)*	65.63 *(18.33)*	66.78 *(22.33)*	87.36 *(80.25)*	*	72.66 *(55.73)*	68.07 *(32.36)*	67.29 *(20.84)*	66.29 *(18.65)*	67.09 *(20.71)*	66.63 *(25.10)*
***Planomicrobium* sp. ES2**	65.86 *(24.97)*	65.16 *(20.19)*	66.17 *(22.69)*	72.45 *(54.86)*	72.57 *(56.63)*	*	67.07 *(32.83)*	66.59 *(21.44)*	66.05 *(19.39)*	66.85 *(21.73)*	66.02 *(25.77)*
***Bhargavaea cecembensis* DSE10**	65.36 *(22.50)*	63.79 *(14.56)*	65.64 *(22.55)*	67.97 *(35.49)*	67.54 *(35.38)*	66.97 *(35.64)*	*	66.36 *(21.80)*	64.35 *(14.20)*	65.60 *(18.95)*	64.97 *(23.12)*
***Bacillus coagulans* ATCC 7050**	67.46 *(26.79)*	66.55 *(20.97)*	68.18 *(27.54)*	67.81 *(21.14)*	67.59 *(21.42)*	67.81 *(21.68)*	67.77 *(21.03)*	*	67.08 *(23.96)*	69.15 *(30.30)*	67.58 *(29.25)*
***Bacillus cytotoxicus* NVH 391-98**	68.06 *(22.10)*	68.57 *(21.67)*	67.47 *(24.03)*	66.99 *(15.18)*	66.99 *(16.01)*	67.45 *(16.91)*	66.46 *(11.99)*	67.22 *(19.47)*	*	69.09 *(24.44)*	68.34 *(26.62)*
***Bacillus methanolicus* PB1**	68.24 *(28.54)*	67.50 *(25.21)*	68.46 *(27.40)*	67.36 *(20.01)*	67.21 *(20.71)*	67.76 *(21.10)*	66.60 *(17.70)*	68.80 *(28.00)*	68.51 *(28.27)*	*	69.07 *(35.71)*
***Bacillus simplex* SH-B26**	68.20 *(24.24)*	67.14 *(18.30)*	67.48 *(22.02)*	66.97 *(16.12)*	66.98 *(16.59)*	67.17 *(17.15)*	66.18 *(15.24)*	67.43 *(19.35)*	68.12 *(22.78)*	69.19 *(24.83)*	*

Average nucleotide identity (ANI) scores generated during global comparisons of the genome sequence (Lee, 2006) of EAG3^T with previously deposited whole genome sequences of *Bacillus methanolicus* PB1, *Bacillus simplex* SH-B26, *Bacillus cytotoxicus* NVH 391-98, *Bacillus coagulans* ATCC 7050, *Bhargavaea cecembensis* DSE10, *Planomicrobium* sp. ES2, *Planococcus rifietoensis* M8, *Planococcus plakortidis* DSM 23997, *Bacillus amyloliquefaciens* DSM 7 and *Bacillus alcalophilus* ATCC 27647 indicates sufficient distance from and supports the findings of 16S rRNA gene phylogeny.

Table 2.14: Genome-to-Genome Distance between strain EAG3^T and other taxonomic neighbours (Family *Bacillaceae*) calculated using GGDC 2.1

*The reference genomes (from Sl. No. 1 to 7) used in this study were of the bacteria *Bacillus amyloliquefaciens* DSM7, *Bacillus simplex* DSM 1321, *Bacillus subtilis* NCIB 3610, *Planococcus rifietoensis* M8, *Bacillus cohnii* DSM 6307, *Planococcus plakortidis* DSM 23997 and *Bacillus coagulans* ATCC 7050 respectively.

Sl. No.	Reference genome*	Formula 1				Formula 2				Formula 3				G+C difference
		DDH	Model C.I.	Distance	Prob. DDH >= 70 %	DDH	Model C.I.	Distance	Prob. DDH >= 70 %	DDH	Model C.I.	Distance	Prob. DDH >= 70 %	
1	**FN597644.1**	12.8	[10.1 - 16.1 %]	0.9811	0	26.6	[24.2 - 29%]	0.1631	0.02	13.2	[10.9 - 16%]	0.9842	0	3.54
2	**CP017704.1**	12.9	[10.2 - 16.2 %]	0.9783	0	24	[21.7 - 26.5 %]	0.1818	0	13.3	[10.9 - 16%]	0.9823	0	2.65
3	**CP020102.1**	12.9	[10.2 - 16.2 %]	0.9792	0	24.9	[22.6 - 27.4 %]	0.175	0.01	13.3	[10.9 - 16%]	0.9828	0	0.97
4	**CP013659.2**	12.7	[10 - 15.9 %]	0.9899	0	35.4	[33 - 37.9 %]	0.1165	0.75	13.1	[10.7 - 15.8 %]	0.9911	0	5.91
5	**CP018866.1**	12.8	[10.1 - 16.1 %]	0.9836	0	31.2	[28.8 - 33.7 %]	0.1357	0.17	13.2	[10.9 - 16%]	0.9858	0	6.41
6	**CP016539.2**	12.7	[10 - 15.9 %]	0.9905	0	38.7	[36.3 - 41.3 %]	0.1037	1.97	13.1	[10.7 - 15.8 %]	0.9915	0	7.43
7	**CP009709.1**	12.7	[10 - 16%]	0.9875	0	32.6	[30.2 - 35.1 %]	0.1289	0.29	13.1	[10.8 - 15.9 %]	0.9891	0	4.35

The ANI score for strain comparisons between EAG3^T and *Bacillus methanolicus* PB1 was the highest (67.45%) which is far below 95- 96 % cut-off value for novel species determination by this approach. The GGDC calculations with BLAST+ for strain EAG3^T and other taxonomic neighbours Family *Bacillaceae* gave DNA–DNA homology values

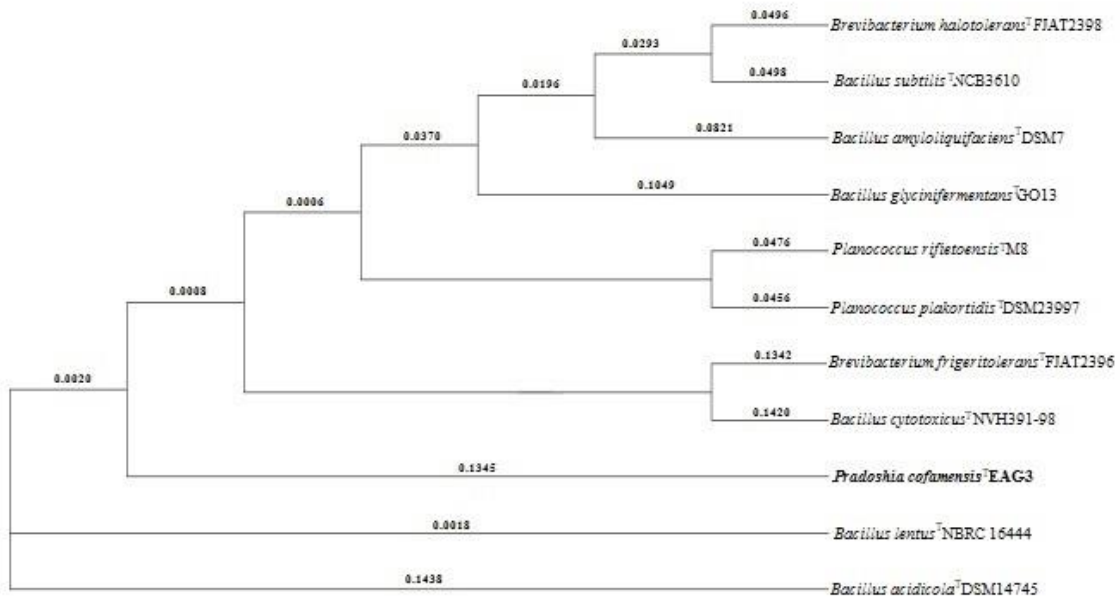


Fig. 2.30: Phylogenetic relationship of strain EAG3^T (*Pradoshia cofamensis*, gen. nov., sp. nov.) and closely related genomes. An assembly and alignment-free (AAF) method was used to construct phylogeny from next-generation sequencing data.

of <13% by model 1, and just above 13% for the third equations in the program. The probable DDH values were "0" for model 1 and 2 and far less than 1 for most of the compared genomes except Planococcus rifietoensis M8 (0.75), Planococcus plakortidis DSM 23997 (1.97). The G+C content difference for these two species of Planococcus were sufficient enough (5.91 and 7.43) for delineating Genus nov.

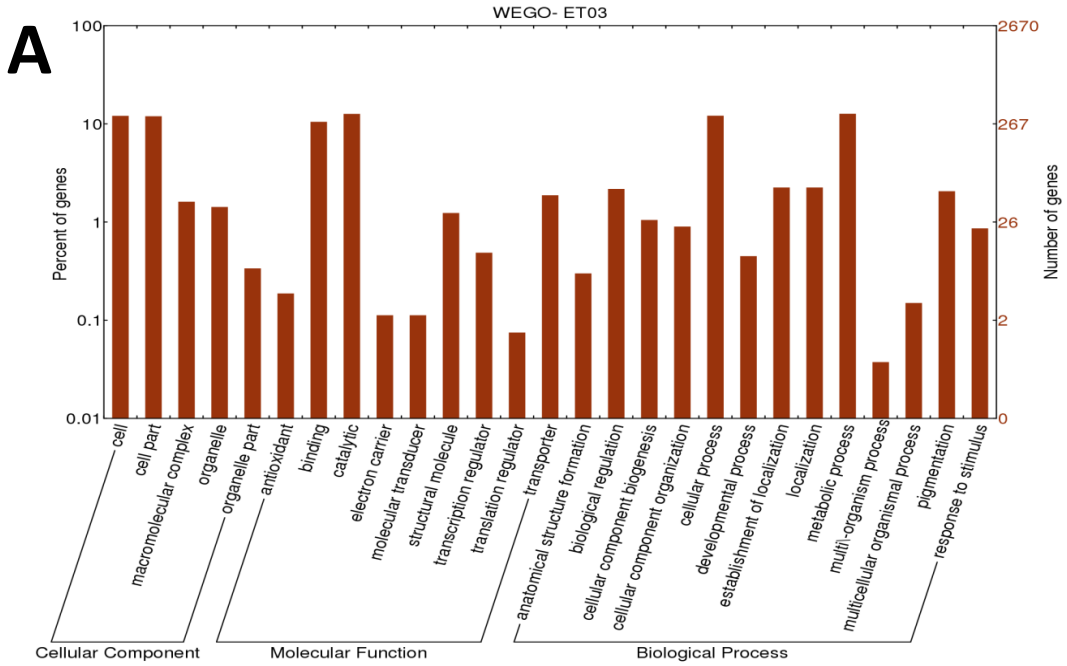
2.3.7.4. Functional analysis by gene annotation (BLASTx) and Gene ontology (Blast2Go)

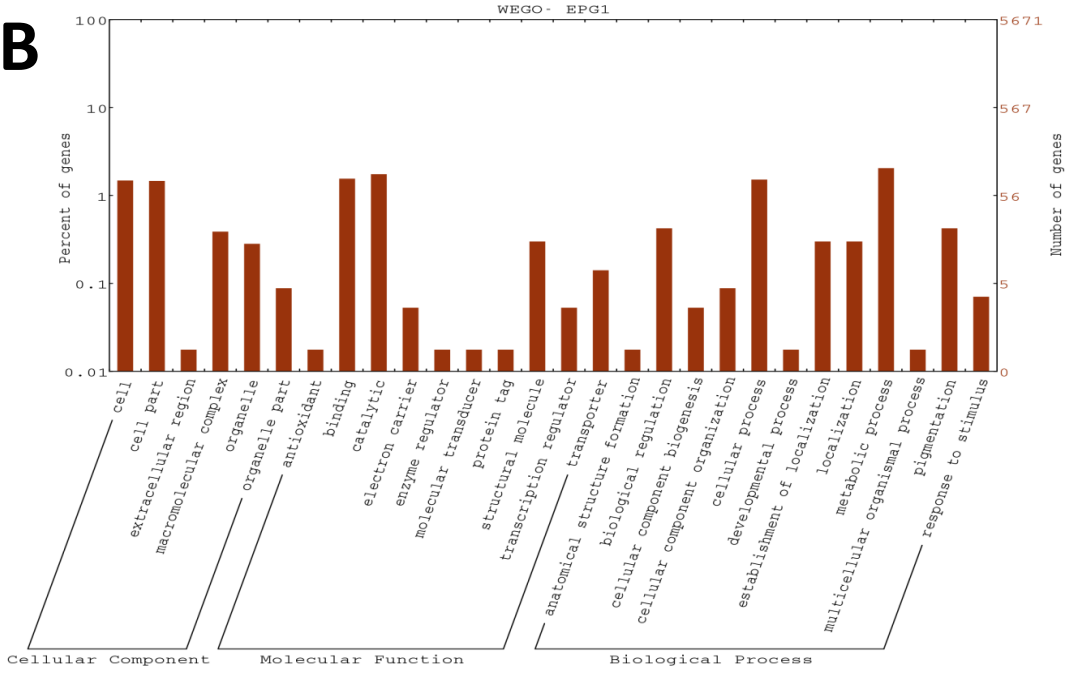
GO mapping provided ontology of defined terms representing gene product properties grouped into three main domains: Biological process, Molecular function and Cellular component.

Table.2.15: GO mapping provided gene ontology

Sl No.	Sample Name	Biological Processes	Cellular Component	Molecular Function
1	EAG2	1,218	735	1,412
2	EAG3	1,670	1,154	1,815
3	EPG1	135	85	154
4	ET03	438	323	476

The WEGO plots generated are depicted in Fig: 2.31-


A
WEGO- ET03
Percent of genes
Number of genes
100
10
1
0.1
0.01
2670
267
26
2
0
cell
cell part
macromolecular complex
organelle
organelle part
antioxidant
binding
catalytic
electron carrier
molecular transducer
structural molecule
transcription regulator
translation regulator
transporter
anatomical structure formation
biological regulation
cellular component biogenesis
cellular component organization
cellular process
developmental process
establishment of localization
localization
metabolic process
multi-organism process
multicellular organismal process
pigmentation
response to stimulus
Cellular Component
Molecular Function
Biological Process


B
WEGO- EPG1
Percent of genes
Number of genes
100
10
1
0.1
0.01
5671
567
56
5
0
cell
cell part
extracellular region
macromolecular complex
organelle
organelle part
antioxidant
binding
catalytic
electron carrier
enzyme regulator
molecular transducer
protein tag
structural molecule
transcription regulator
transporter
anatomical structure formation
biological regulation
cellular component biogenesis
cellular component organization
cellular process
developmental process
establishment of localization
localization
metabolic process
multicellular organismal process
pigmentation
response to stimulus
Cellular Component
Molecular Function
Biological Process

Cont….

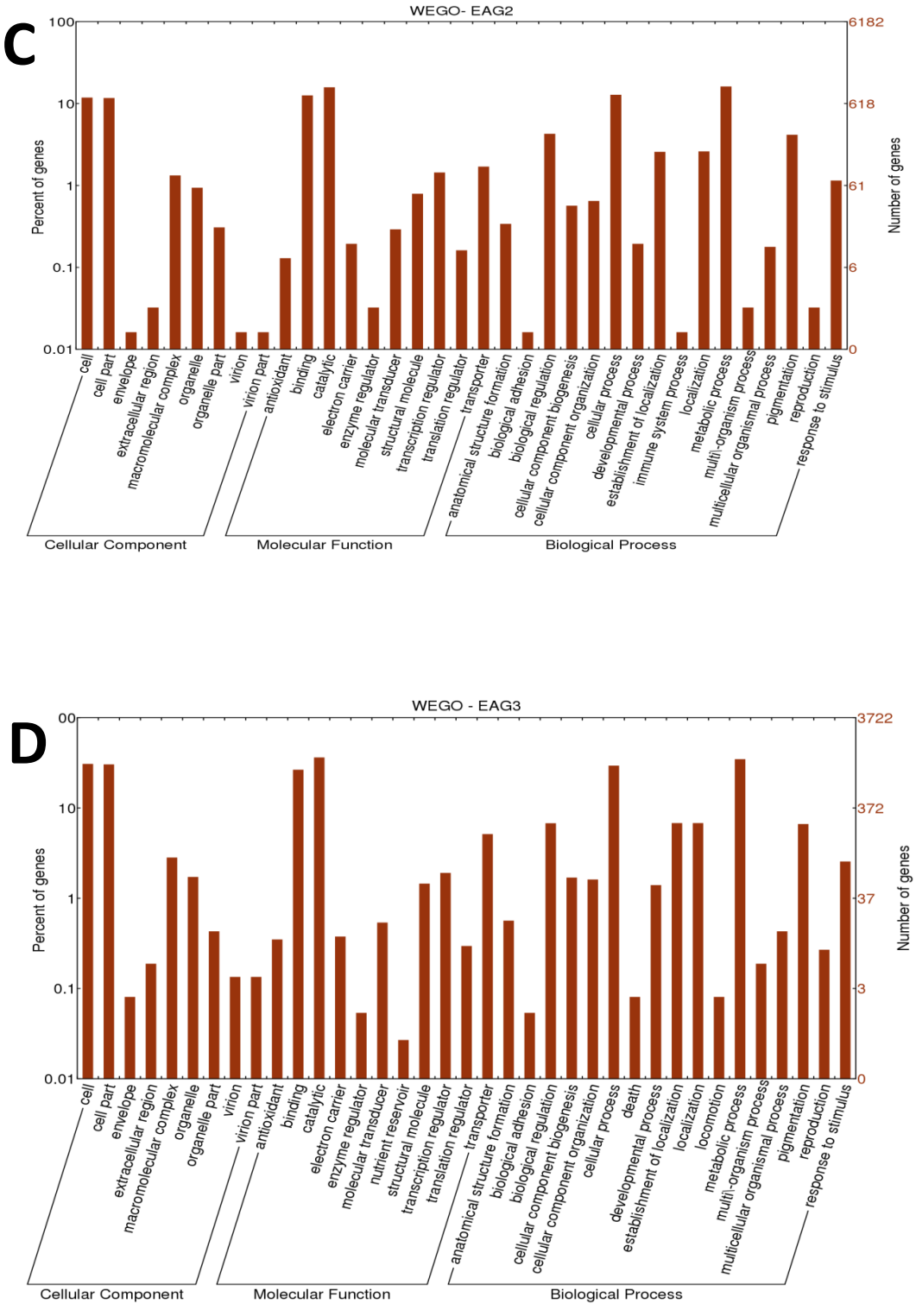


Fig: 2.31: WEGO plots of the four strains. A- ET03^{T}, B-EPG1^{T}, C-EAG2^{T} and D-EAG3^{T}

CHAPTER - 3

MODULATION OF COELOMIC FLUID CONSTITUENTS UNDER BACTERIAL CHALLENGE

3.1 Introduction

Eisenia fetida (Savigny, 1826), the clitellate annelid is a universally accepted vermicomposting earthworm living in and on the decomposing organic matter. Soil and organic debris contain different groups of bacteria along with other microorganisms. *Eisenia,* in context to its living zone certainly has to face the bacterial challenge (Aira *et al.,* 2007). *E. fetida* hosts $6X10^5$ /ml ($0.9X10^5$ /adult individual) bacteria in its coelom (Edwards and Lofty, 1977). The Superior numbers of phagocytes along with various humoral factors prevent the microorganisms from outgrowth (Bilej *et al.,* 2000; Ko¨hlerova´, 2004). If the bacterial burden surpasses the optimum limit it may cause disease in the earthworm. Smirnoff and Heimpel (1961) reported that when large doses of *Bacillus thuringiensis* invaded the body cavity of the earthworm, it caused an extensive septicemia and eventual death. Hiempel (1966) reported that blister disease of *E. fetida* was found to contain crystalliferous bacteria in all the lesions. So, logically *E. fetida* seems to have certain defence mechanisms that would be able to protect it against invading pathogens. Coelomocytes in the coelomic cavity are reported to be the major players of the immune system (Dhainaut and Scaps, 2001; Dvořák *et al.,* 2016). Semi-quantitative Expressed sequence tag (EST) based genomic study on a closely related species *E. andrei* following bacterial challenge reported 24 immune-related and cell defense genes (Bilej 2014). However, there is limited genetic information available for immunological pathways in *E. fetida* under bacterial challenge. This study compared transcriptomic profiles of coelomocytes from normal and bacteria infected *E. fetida* by Illumina-based paired-end sequencing to explore the molecular immune mechanism of *E. fetida* against bacterial infection. Adult *E. fetida* earthworms maintained at 22 ℃ in compost were transferred 2 days before experiments on filter paper soaked with an isotonic balanced salt solution (BSS) (Stein and Cooper, 1981). Bacterial inoculation in the coelomic fluid was done following the protocol described by Kauschke and Mohrig in 1987 with minor modifications.

Percoll (Pharmacia) was used as a cell separation media by Hamed *et al.,* 2005. Coelomocytes thus isolated was evaluated for living cell numbers by trypan-blue exclusion method; and Wright-Giemsa staining was performed for cytological study. Cell

size as well as nuclear diameter are measured with a calibrated ocular scale (10x), using a Zeiss microscope (40x) objective and a microcytometer.

Modern medical research has indicated that the coelomic fluid (CF) of earthworms contains an abundance of bioactive substances including lectin (Suzuki *et al.,* 2009), polysaccharide (Wang *et al.,* 2007), protease (Sugimoto *et al*., 2003), antibacterial peptide (Wang *et al.,* 2003), metalloenzyme (Sturzenbaum *et al.,* 2001), fibrinolytic enzyme (Wang *et al.,* 2005), and so on. Earthworm proteins and peptides have exhibited various biological activities (Liu *et al.,* 2004; Wang *et al.,* 2007). Fetidin, with apparent molecular weight of 40 kDa, was purified from earthworm coelomic fluid and its bioactivities of antibacterial action, hemolysis and hemocoagulation were estimated (Milochau *et al.,* 1997). A 42 kDa protein, named coelomic cytolytic factor 1 (CCF-1), was reported to have cytolytic, opsonizing and hemolytic properties (Beschin *et al.,* 1998). Lumbricin I which was isolated and characterized from the earthworm showed antimicrobial activity *in vitro* against a broad spectrum of microorganisms without hemolytic activity (Cho *et al.,* 1998).

ECFP, 38.6 kDa protein, was isolated from *E. fetida* and shown to possess significant hemolytic activity to chicken red blood cells (CRBC) (minimal hemolytic concentration 0.39 µg/mL) and antibacterial effect against *Escherichia coli* (minimal bactericidal concentration, MBC 180 µg/mL) and *Staphylococcus aureus* (MBC 90 µg/mL) (Hua *et al.,* 2011).

Only fragmented data regarding the immune signaling and function of immune system in lumbricids are available, so, there is ample scope to undertake through investigation to unveil the immune pathways in these group.

Whole transcriptome analysis is a powerful tool for understanding different cellular, metabolic and developmental pathways that continue in an organism (or in any particular tissue or cell type that is sampled) under a specific condition like bacterial challenge (Khoo *et al*., 2012). Recently, NGS transcriptome sequencing technology has gradually replaced the gene chip technology and becomes a major tool for studying gene expression (Mu *et al*., 2014; Zhou *et al*., 2015; Dheilly *et al*., 2014). The high-throughput transcriptome sequencing is particularly important to provided insights into the immunogenetics of nonmodel

organisms, lacking reference genomes (Hanelt *et al*., 2008).

In earthworms, coelomocytes are involved in immune responses involving phagocytosis, encapsulation, coagulation etc. (Li *et al*., 2013; Gliński *et al*., 2000; Eliseikina *et al*., 2002; Dolmatova *et al*., 2004; Mutz *et al*., 2013). In this study, we focused on coelomocytes for the transcriptome analysis to have a deeper look into the immune mechanism of *E. fetida* against bacterial infection.

Zhang *et al.,* (2018) studied transcriptome of coelomocytes of Onchidium struma after bacterial challenge. Hanelt *et al.,* (Adema *et al*., 2010) conducted a similar transcriptome study of *Biomphalaria glabrata* snails challenged with *E. coli*. No studies have been reported on the coelomocytes transcriptome of *E. fetida* in response to bacterial infection. There is limited genetic information on *E. fetida* and further study is needed for understanding its disease resistance mechanism.

To gain deeper insight into the molecular immune mechanism of *E. fetida* against bacterial infection, *de novo* transcriptome sequencing of its coelomocytes after infection with *Bacillus thuringiensis* was performed on Illumina NextSeq 500 using 2 x 75 bp chemistry, and performed a global survey of immune-related genes, annotation of immune signaling pathways and determination of gene expression.

3.2. Materials and methods

3.2.1 Microbial challenge to *E. fetida*

Adult *E. fetida* worms were acclimatized to the laboratory conditions and were then distributed in experimental and control groups. The experimental animals were challenged by *Bacillus thuringiensis*

3.2.1.1 Dose determination for microbial challenge experiment

The LD_{50} value for the number of live *B. thuringinsis* cell injected in adult *E. fetida* worm (Body wt. ~ 0.3g) was calculated using arithmetic method of Karber as adapted by Alui and Nwude (1982). Log phase culture (~1 OD) of *B. thuringiensis*, supposed to carry 5.0E+8 cfu/ml (Fisheret *et al., 2001)* was taken in microfuge tubes, centrifuged at 2000g for 5 minute at 4°C, washed twice with sterile MiliQ water and finally re-suspended to the concentrations 10^7 cfu/ml, 10^8 cfu/ml, 10^9 cfu/ml and 10^{10} cfu/ml. Approximately 10 µl of this bacterial suspension was injected in each worm of the four

dose-groups (each group contained 10 animals) using insulin syringe (Dispo-van, India) with 31 gauze fine needle. The lethal effects of various doses of bacterial injection are shown in Table 3.1.

$1/10^{th}$ of LD_{50} was taken as the dose for injecting each individual of the experimental groups of the earthworm.

3.2.1.2 Injecting microbes into coelomic space

$1/10^{th}$ of LD_{50} dose of live *B. thuringiensis* is administered as the sub-lethal bacterial challenge per adult *E. fetida.* For this, 400 µl of Log phase culture (~1 OD) of *B. thuringiensis* was taken in the microfuge tube, centrifuged at 2000g for 5 minutes at 4°C, washed twice with sterile MiliQ water and finally resuspended in 1 ml. Approximately 10 µl of this bacterial suspension (or ~2 x 10^6 live *B. thuringiensis*) was injected in each worm of the three treated groups using insulin syringe (Dispo-van, India) with 31 gauze fine needle. All *E. fetida* of other three Petri-plates were injected with 10 µl of sterile water.

The experiments were carried out in three (03) different sets. In each set the worms were kept in Eight (08) Petri-plates distributed equally in experimental and control groups (each plate with twenty-five adult earthworms) and then incubated at 25°C and fed on rehydrated (moisture, 80%) cow dung chips.

3.2.2. Comparative study of coelomocytes in challenged and control groups of *E. fetida*

Coelomic fluid was periodically withdrawn from the body of *E. fetida* of the experimental and control groups. Coelomocytes were isolated, identified, counted and studied for their phagocytic nature.

3.2.2.1 Periodic withdrawal of coelomic fluid of E. fetida

Coelomic fluid from the treated and control groups of earthworm was periodically (at 12h, 24h, 36h and 48h) pooled out in separate siliconized centrifuge tubes. Sterile, sharpened capillary tubes were used to withdraw the coelomic fluid from surface cleaned earthworms (Bilej *et al.,* 1990). Sterile 1x PBS, pH7.4, prepared by using PBS powder (Himedia M1866) and MiliQ water was used to wash the coelomocytes. Total process

was done on ice.

Coelomocytes were separated by centrifugation (100xg for 10 min) and cell-free coelomic fluid was collected and kept at -20 °C. Isolated cells were resuspended in PBS.

3.2.2.2 Separation of coelomocytes

A portion from each group of the Isolated cells was separated on a Percoll gradient (Pharmacia, Sweden) created by mixing with 0.15 M NaCl to prepare six step concentration gradients, 5-15%, 25-35% and 45-55% from top to bottom of the glass centrifuge tube (Hamed *et al.,* 2005). Coelomocyte suspension was put on top of the concentration-gradient and centrifuged at 1000g for 10 minutes. Coelomocyte bands from different layers were collected carefully by Pasteur pipette. Separated coelomocytes were washed twice in PBS.

3.2.2.3 Total count (TC) and Differential count (DC) of coelomocytes

Coelomocytes collected from each group of earthworm were counted using an improved Neubauer haemocytometer using standard methods (Kirk *et al.,* 1975) and reported as the mean ± standard deviation (SD) per ml. Cell viability was determined by staining with 0'4% trypan blue (Sigma T0887). Viability was expressed as the percentage of live-cells at counting.

Differential counts were made for each group of worms by placing a drop of coelomocyte suspension in PBS onto a clear glass slide and making a smear. After air-drying, coelomocytes were fixed in methanol for 20 s and stained with modified Wright's stain (Sigma WS16) for 20 s. One hundred leukocytes were classified under oil immersion using 100 x objective lens.

3.2.2.4 Micrometry of coelomocytes

The diameter of the stained coelomocytes and their nuclei were measured using an ocular micrometer which was calibrated using a stage micrometer. Three readings were taken to determine the average diameter in each case.

3.2.2.5 Study of phagocytic behaviour of coelomocytes

A carbon particle suspension in PBS (pH, 7.4) solution was prepared by centrifuging Higgins India ink at 1,300g for ten minutes, removing the supernatant, and washing the pellet four times in PBS (Stein *et al.,* 1975). After the final wash, PBS amount was adjusted such that the carbon pellet was 25% (V/V). 200µl of live coelomocyte suspension in PBS (adjusted to 10^6 cells/ml) was taken on a clear glass slide with 20µl of the carbon suspension and incubated at 28 °C in clean moist Petri plate chambers. Three such slides were prepared from each experimental and control groups of *E. fetida* for time-lapse study of phagocytosis percentage at 30min, 1h and 1h 30min. After incubation at different time intervals the slides were taken out, gently smeared, fixed and stained following procedure mentioned in section 3.2.2.3 and observed under the microscope.

3.2.2.6 Scanning Electron Microscopy (SEM) of the coelomocytes.

Coelomocytes from different experimental and control groups of *E. fetida* were spread on coverslips, semi-dried and were then fixed in the mixture of 1% paraformaldehyde and 2,5% glutaraldehyde, then postfixed in OsO4. The coverslips containing coelomocytes were completely dehydrated by passing through increasing grade of alcohol and drying in a critical point dryer. The cells were gold metal coated and observed in an scanning electron microscope (JS MIT 100, JEOL Ltd., Tokyo, Japan) at USIC, NBU.

3.2.3. Comparative study of coelomic fluid in challenged and control groups of *E. fetida*

Coelomic fluid and coelomocyte lysate (obtained by five cycles of freezing and thawing and treated with a proteinase inhibitor cocktail; Sigma P2714) from the treated and control groups of earthworms at different time intervals were analyzed for total protein estimation and SDS-PAGE.

3.2.3.1 Estimation of total protein content

Protein concentration was determined by the Bradford method (1976), using bovine serum albumin (BSA) to construct the standard curve. The standard curve was established as follows: y = 0.018x + 0.175 (R2=0.967) [x-axis represented the protein concentration and y-axis represented optical density (OD)].

3.2.3.2 SDS-PAGE analysis

The crude coelomic fluid and coelomocyte lysate after making bacteria free by filtering through 0.2 um membrane were separated in SDS-PAGE (chamber 0.1×10×10cm) using 10% polyacrylamide resolving gel and a 4% stacking gel in the presence of 1% SDS. The migration buffer consisted of 25mM Tris and 192mM glycine (pH 8.6). After migration, the gel was stained with Coomassie brilliant blue R-250, and decolored with destaining solution (10% methanol:20% ice-cool acetic acid:70% dH_2O).

3.2.4. Dynamics of indigenous *Bacillus* spp. in coelomic fluid of *E. fetida* under challenge by *B. thuringiensis*

Five earthworms (*E. fetida*), at different intervals (12h, 24h, 36h and 48h), were taken from each of the three experimental plates and washed several times with sterile distilled water. Thoroughly washed individual earthworm was held firmly in a sterile tissue paper and coelomic fluid (~10 µl) was collected in a fine- sharp sterile capillary tube (diameter ≤ 1.0 mm) by puncturing the coelomic cavity. The coelomic fluid contents of the capillary tubes were pooled in a PCR microfuge tube, centrifuged at 150g for 5min to pellet out the coelomocytes. The supernatent was serially diluted in sterile PBS. Multiple dilutions and plating on HBA plates were done to enumerate the number of *B. thuringiensis* content (CFUs) in the coelomic fluid and differentiation between various species of *Bacillus* along with viable count respectively.

3.2.5. Comparative gene expression studies via whole transcriptomics in coelomocytes of *'B. thuringiensis* challenged' and 'control' *E. fetida*

3.2.5.1 Total RNA Isolation and cDNA library construction

Total RNA was isolated from both the treated and control coelomocyte samples of *Eisenia fetida* using commercially available Quick-RNA Miniprep Plus kit (ZYMO Research) as per the manufacturer's instruction. The quality and quantity of the isolated RNA was checked on 1% denaturing RNA agarose gel and NanoDrop, respectively.

The RNA-Seq paired end sequencing libraries were prepared from the QC passed RNA samples using illumina TruSeq Stranded mRNA sample Prep kit. Briefly, mRNA was enriched from the total RNA using poly-T attached magnetic beads, followed by

enzymatic fragmentation, 1st strand cDNA conversion using SuperScript II and Act-D mix to facilitate RNA dependent synthesis. The 1st strand cDNA was then synthesized to second strand using second strand mix. The ds-cDNA was then purified using AMPure XP beads followed by A-tailing, adapter ligation and then enriched by limited no of PCR cycles. The PCR enriched libraries were analyzed in 4200 Tape Station system (Agilent Technologies) using High Sensitivity D1000 Screen Tape as per manufacturer instructions.

3.2.5.2 Sequencing and assembly

After obtaining the Qubit concentration for the libraries and the mean peak sizes from Agilent Tape Station profile, the PE illumina libraries were loaded onto NextSeq 500 for cluster generation and sequencing. Paired-End sequencing allows the template fragments to be sequenced in both the forward and reverse directions on NextSeq 500 using 2 x 75 bp chemistry. The kit reagents were used in binding of the samples to complementary adapter oligos on the paired-end flow cell. The adapters were designed to allow selective cleavage of the forward strands after re-synthesis of the reverse strand during sequencing. The copied reverse strand was then used to sequence from the opposite end of the fragment.

The filtered high-quality reads of the 2 samples were pooled together and assembled into transcripts using Trinity de novo assembler (V2.5) with a kmer of 25 and minimum contig length of 200. The assembled transcripts were then further clustered together using CD-HIT-EST-4.5.4 to remove the isoforms produced during assembly. This resulted in sequences that can no longer be extended. Such sequences are defined as Unigenes. Only those unigenes which were found to have >90% coverage at 3X read depth were considered for downstream analysis.

3.2.5.3 Coding sequence (CDS) Prediction

TransDecoder-v2.0 was used to predict coding sequences from the above-mentioned unigenes. TransDecoder identifies candidate coding regions within unigene sequences. TransDecoder identifies likely coding sequences based on the following criteria:

A minimum length open reading frame (ORF) is found in a unigene sequence

A log-likelihood score similar to what is computed by the GeneID software is > 0.

The above coding score is greatest when the ORF is scored in the 1st reading frame as compared to scores in the other 5 reading frames. If a candidate ORF is found fully encapsulated by the coordinates of another candidate ORF, the longer one is reported. However, a single unigene can report multiple ORFs (allowing for operons, chimeras, etc.).

3.2.5.4 Functional unigene annotation and classification

Functional annotation of the genes was performed using DIAMOND program, which is a BLAST-compatible local aligner for mapping translated DNA query sequences against a protein reference database. DIAMOND (BLASTX alignment mode) finds the homologous sequences for the genes against NR (non-redundant protein database) from NCBI. Majority of the blast hits were found to be against *Capitella teleta* which also belongs to the phylum Annelida.

To identify sample wise CDS from above mentioned pooled set of [control(Cn) and treated (Ex)] CDS, reads from each of the samples were mapped on the final set of pooled CDS using bwa (-mem) toolkit. The read count (RC) values were calculated from the resulting mapping and those CDS having 90% coverage and 3X read depth were considered for downstream analysis for each of the samples. Sample-wise CDS statistics have been summarized in the following table.

3.2.5.5 Gene Ontology Analysis

Gene ontology (GO) analyses of the CDS identified for each of the 2 samples were carried out using Blast2GO program. GO assignments were used to classify the functions of the predicted CDS. The GO mapping also provides the ontology of defined terms representing gene product properties which are grouped into three main domains: Biological Process (BP), Molecular Function (MF) and Cellular Component (CC).

GO mapping was carried out in order to retrieve GO terms for all the functionally annotated CDS. The GO mapping uses following criteria to retrieve GO terms for the functionally annotated CDS:

BlastX result accession IDs are used to retrieve gene names or symbols, identified gene names or symbols are then searched in the species-specific entries of the gene- product tables of GO database.

BlastX result accession IDs are used to retrieve UniProt IDs making use of PIR which includes PSD, UniProt, SwissProt, TrEMBL, RefSeq, GenPept and PDB databases.

Accession IDs are searched directly in the dbxref table of GO database.

BlastX result accession IDs are searched directly in the gene product table of GO database.

3.2.5.6 Functional Annotation of KEGG Pathway

To identify the potential involvement of the predicted CDS in biological pathways, 21,722 and 22,143 CDS of Cn and Ex samples respectively were mapped to reference canonical pathways in KEGG.

3.2.5.7 Differential Gene Expression Analysis

To perform differential expression analysis for Cn andEx samples, reads from the individual samples were mapped on the final set of CDS using bwa (-mem) toolkit. The read count (RC) values were calculated from the resulting mapping and only those CDS were considered for differential expression analysis which passed the 90% coverage at 3X depth criteria. Then by employing a negative binomial distribution model in DESeq package (version1.22.1 http://www.huber.embl.de/users/anders/DESeq/) differential gene expressions were calculated. Dispersion values were estimated with the following parameters: method = blind, sharingMode = fit-only and fittype = local. Finally Log2 fold change (FC) value was calculated from the aforementioned RC values.

$$FC = Log2\,(treated/control)$$

The CDS with FC value greater than zero were considered as up-regulated whereas less than zero as down-regulated. P-value threshold of 0.05 was used to filter statistically significant results. The CDS with FC value equal to zero are considered as "not regulated".

The combination used to identify the differentially expressed genes was 'Cn' for control and 'Ex' for experiment or treated groups respectively.

3.3. Results and discussion

3.3.1 Microbial challenge to *E. fetida*

3.3.1.1 Dose determination for microbial challenge experiment

Group	**Dose (cfu/ml)**	**Dose Difference (cfu/ml)**	**No. o f Animal per group.**	**No. Dead**	**Mean Dead**	**Dose difference X mean Dead**
1	control	-	10	0	0	0
2	10,000,000	10,000,000	10	0	0	0
3	100,000,000	90,000,000	10	4	2	180,000,000
4	1000,000,000	900,000,000	10	7	5.5	4950,000,000
5	10000,000,000	9000,000,000	10	10	8.5	76500,000,000
LD_{50} (*B.thuringiensis)* =10000,000,000-(81630,000,000/10)= 1837,000,000 cfu/ml						Sum=81630,000,000

The LD_{50} of live *B. thuringiensis,* injection was found to be 1.8 x 10^9 cfu/ml or ~2 x 10^7 live *B. thuringiensis* in 10 µl dose per adult *E. fetida*. 1/10^{th} of LD_{50} was taken as the dose for injecting each individual of the experimental groups of the earthworm.

3.3.2. Coelomocytes in challenged and control groups of *E. fetida*

3.3.2. 1 W*ithdrawal volume of coelomic fluid of E. fetida*

Coelomic fluid was periodically withdrawn from the body of *E. fetida* of the experimental and control groups. 200 µl of coelomic fluid (from~20 earthworms) was extracted from each of the experimental and control groups of the different sets.

3.3.2.2 Separation of coelomocytes

Four semi-separated cell bands were observed at percoll concentrations of 10, 25, 40 and 50% contining amoebocytes, large granulocytes, eleocytes and small granulocytes as the

main components correspondingly. This observation is comparable to the similar report by Hamed *et al.*, (2005), where the authors found bands of acidophils, basophils, chloragocytes and nutrophils at more or less similar percoll gradients.

3.3.2. 3 Total count (TC) and Differential count (DC) of coelomocytes

Light microscopic studies of coelomocytes in the untreated adult (Bd wt. 0.3±0.05g) *Eisenia fetida* coelom contains 1.2 ± 0.7 x 10^7 coelomocytes (TC) /ml, with the following morphologically distinct groups - amoebocytes (or hyaline amoebocytes)(23±9%), large granulocytes (or basophils) (18±7%), eleocytes (or chloragocytes) (6±3%) and small granulocytes (or granular amoebocytes) (51±8%).

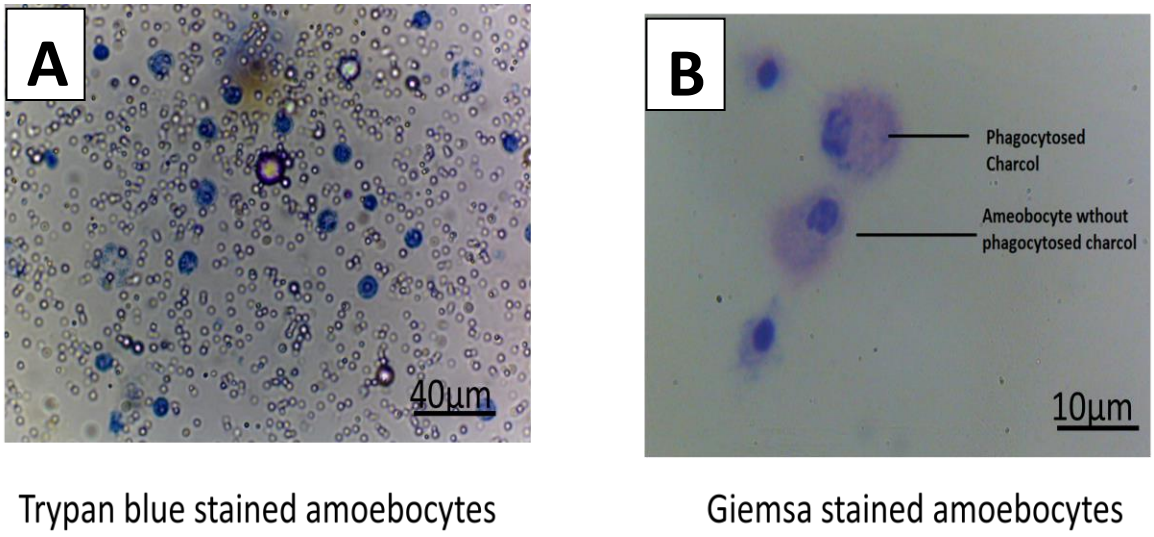


Fig. 3.2: Microscopic observations during Total count (A) and Differential count (B) of coelomocytes of *E. fetida*

Amoebocytes with phagocytotic activity can be identified from the size and shape of the nucleus. The cells tend to be smaller, as little as 8 µm, but occasionally may be as large as 15 µm with fewer granules, eccentric nucleus and large pseudopodia. Granulocytes have cytoplasm completely filled with small granules (basophilic) and have the diameter of 15-22 µm. The eleocytes are large cells with 30-60 µm diameter (with fluorescent yellow pigmentation). The fourth group of cells, small granulocytes are generally, <6 µm in diameter and are densely packed with granules (acidophilic).

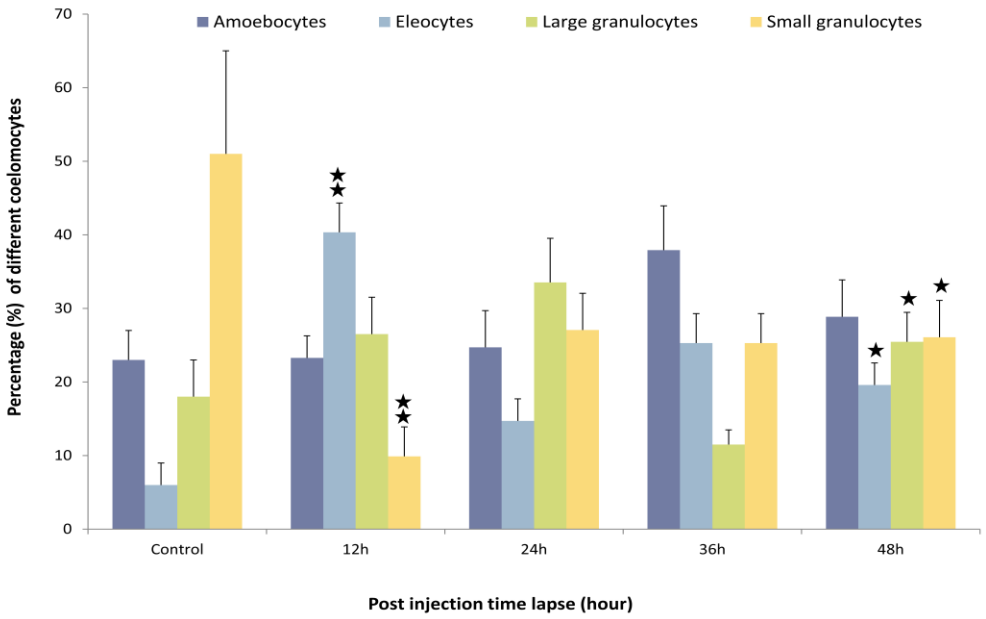


Fig. 3.2: Differential count of coelomocytes in *B. thuringiensis* challenged *E. fetida* at different time intervals

3.3.2. 4 Microscopic attributes of the coelomocytes

Micrometric reports of the different coelomocytes are enlisted below-

Table3-2: Cell and nuclear size of different coelomocytes

Cell Type	Size of the Cell	Size of Nucleus
Amoebocytes	8-15 µm	3-5 µm
Large granulocytes	15-22 µm	4-6 µm
Eleocytes	30-60 µm	5-8 µm
Small granulocytes	<6 µm	<4 µm

3.3.2.5 Phagocytic behavior of coelomocytes

Both amoebocytes and granulocytes were frequently found with inclusion bodies including nano-carbon particles. In the treated groups, these cells were often found to interact with bacilli especially at 24 and 36 h post inoculation groups (Fig. 3.3). These

coelomocytes were frequently observed to accumulate towards the activated carbon particles or formalin fixed *B. thuringiensis* cells (Fig. 3.4), which indicate their increased efficiency of chemotaxis and phagocytosis.

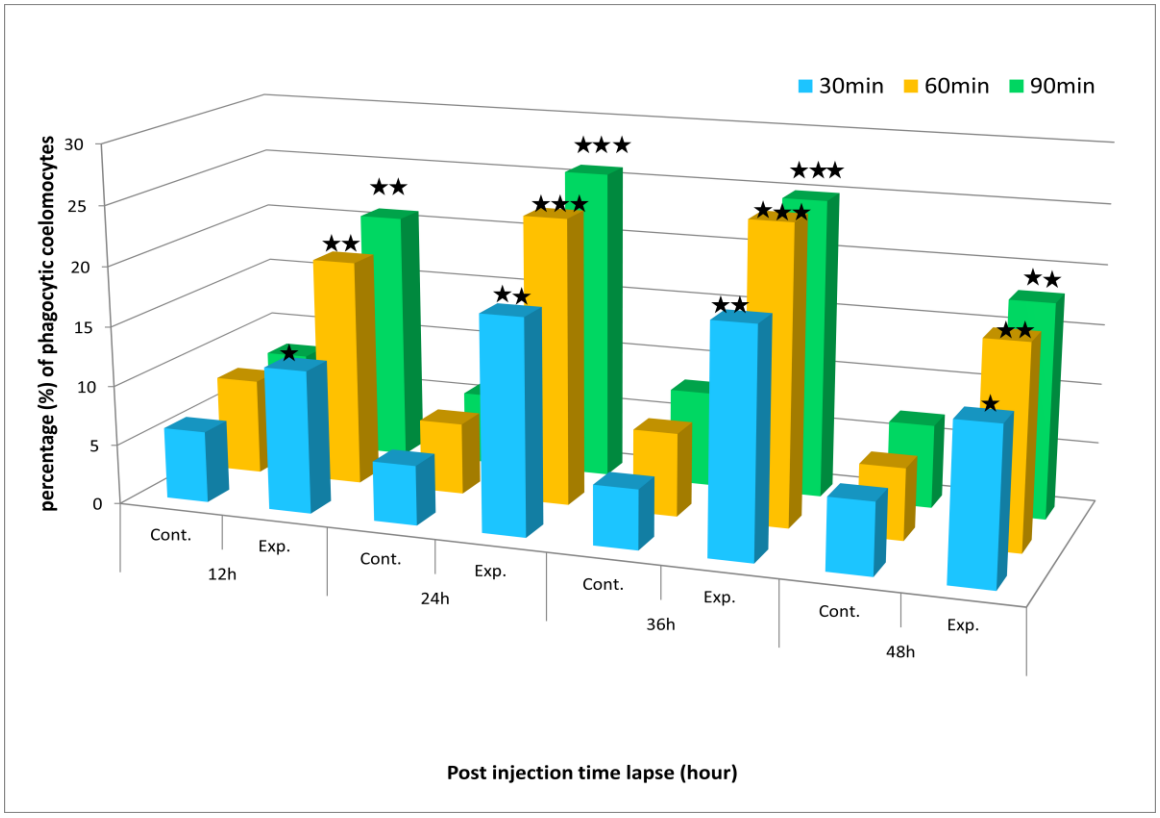


Fig. 3.3: Bar graph showing percentage of coelomocytes with phagocytic inclusion bodies in the experimental and control *E. fetida* at different time intervals after challenge.

3.3.2.6 Scanning Electron Microscopy (SEM) of the coelomocytes

Outer morphology of the four types of coelomocytes were revealed by SEM. Fig 3.5 depicts the SEM plates of glutaraldehyde fixed coelomocytes from *E. fetida*. Amoebocytes (Fig. 3.5-A) are seen to throw extensive pseudopods in all directions. Large granulocte (Fig. 3.5-B) can be easily identified with its numerous small pseudopodia. Eleocyte (Fig.3.5-C) are large spherical cells. Small granulocyte (Fig. 3.5-D) are the smallest spherical cells. Different types of coelomocytes present in coelomic fluid of *E. fetida* interact with bacteria and possibly keep the bacterial population under control by regularly engulfing them.

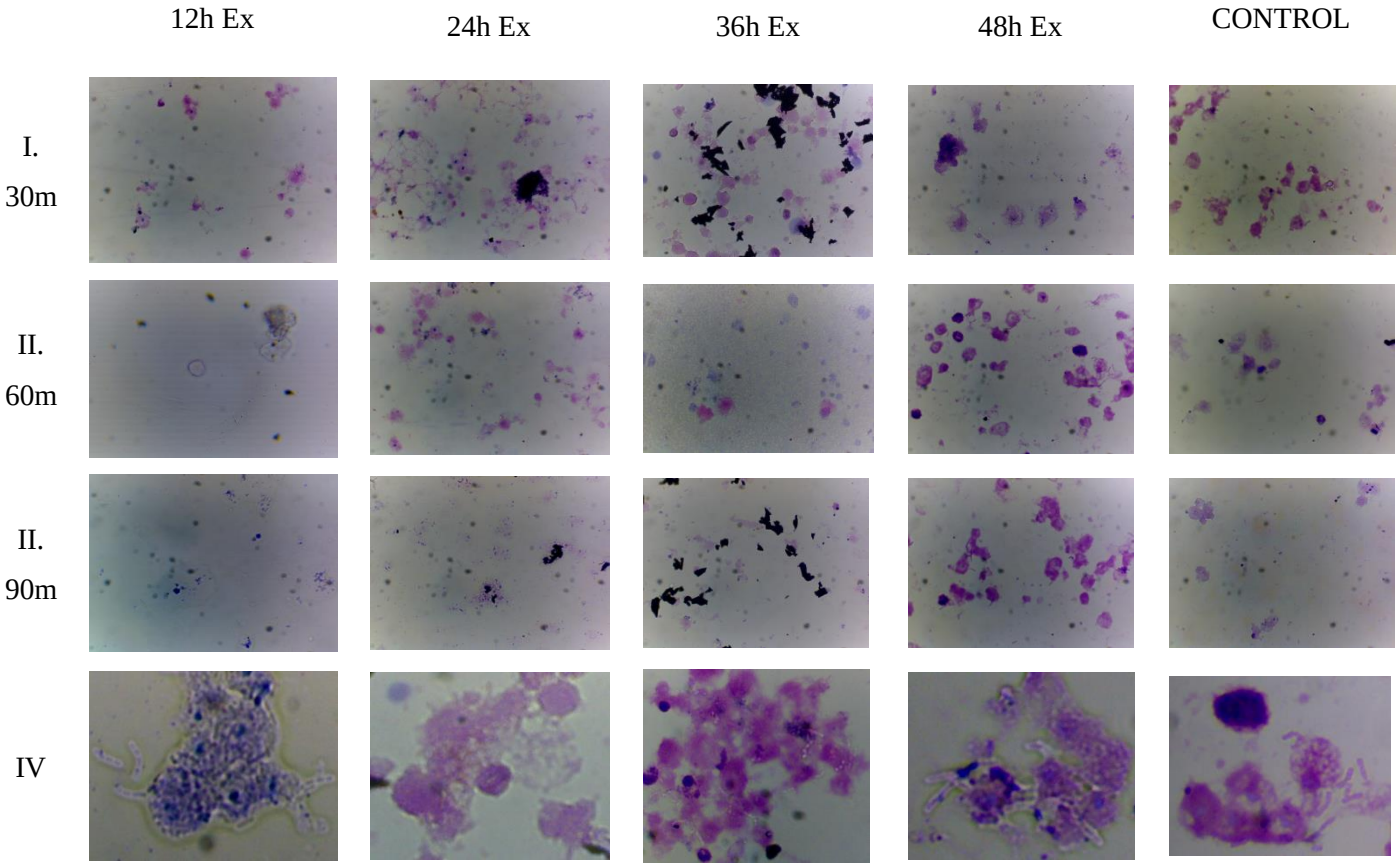


Fig.3.4: Photographic plates of giemsa stained coelomocytes from *B. thuringiensis* challenged *E. fetida* at different time intervals with activated carbon (Row I, II and III; 40x objective) or formalin fixed *B. thuringiensis* cells (Row IV; 100x objective).

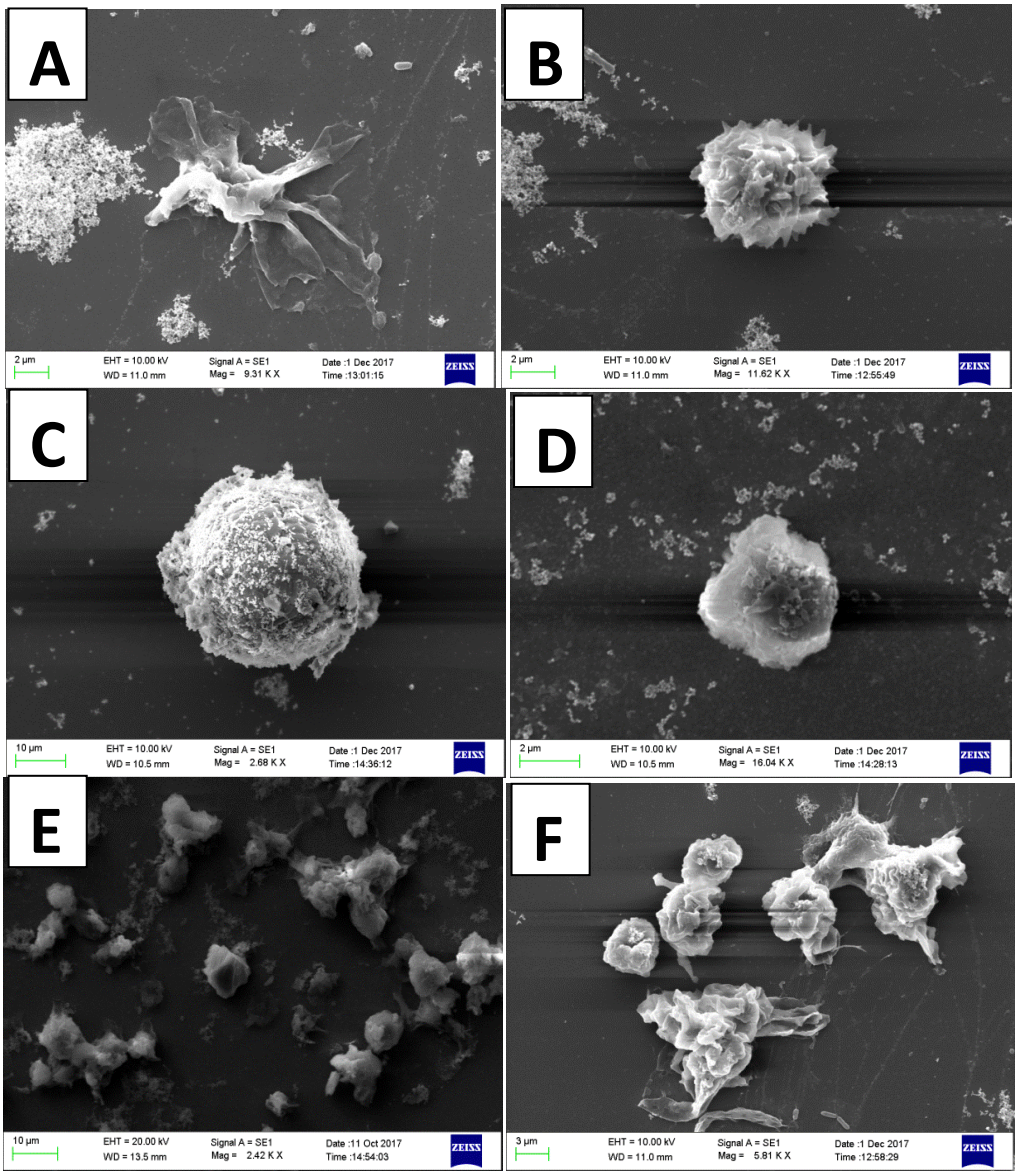


Fig.3.5**:** SEM plates of glutaraldehyde fixed coelomocytes from *E. fetida*. A- Amebocyte, B- Large granulocte, C- Eleocyte, D- Small granulocyte, E- Different types of coelomocytes present in coelomic fluid of *E. fetida* and F- Coelomocytes interacting with bacteria.

3.3.3. Protein analysis of coelomic fluid in challenged and control groups of *E. fetida*

Coelomic fluid protein content was analysed in the following two steps-

3.3.3.1 Total protein content

Total protein content estimated in the untreated animals was 1.2±0.2 mg/ml in coelomic fluid and 6.9±0.6 mg/ml in the coelomocytes. Time lapse analysis of the total protein content is observed to increase in both coelomic fluid and coelomocytes after inoculation

of live *Bacillus thuringiensis* after 24 and 36h lapse to upto 10.2±0.4 mg/ml in coelomic fluid and 12.5±0.8 mg/ml in the coelomocytes (Fig. 3.6). This trend is observed to decrease at 48h onwards.

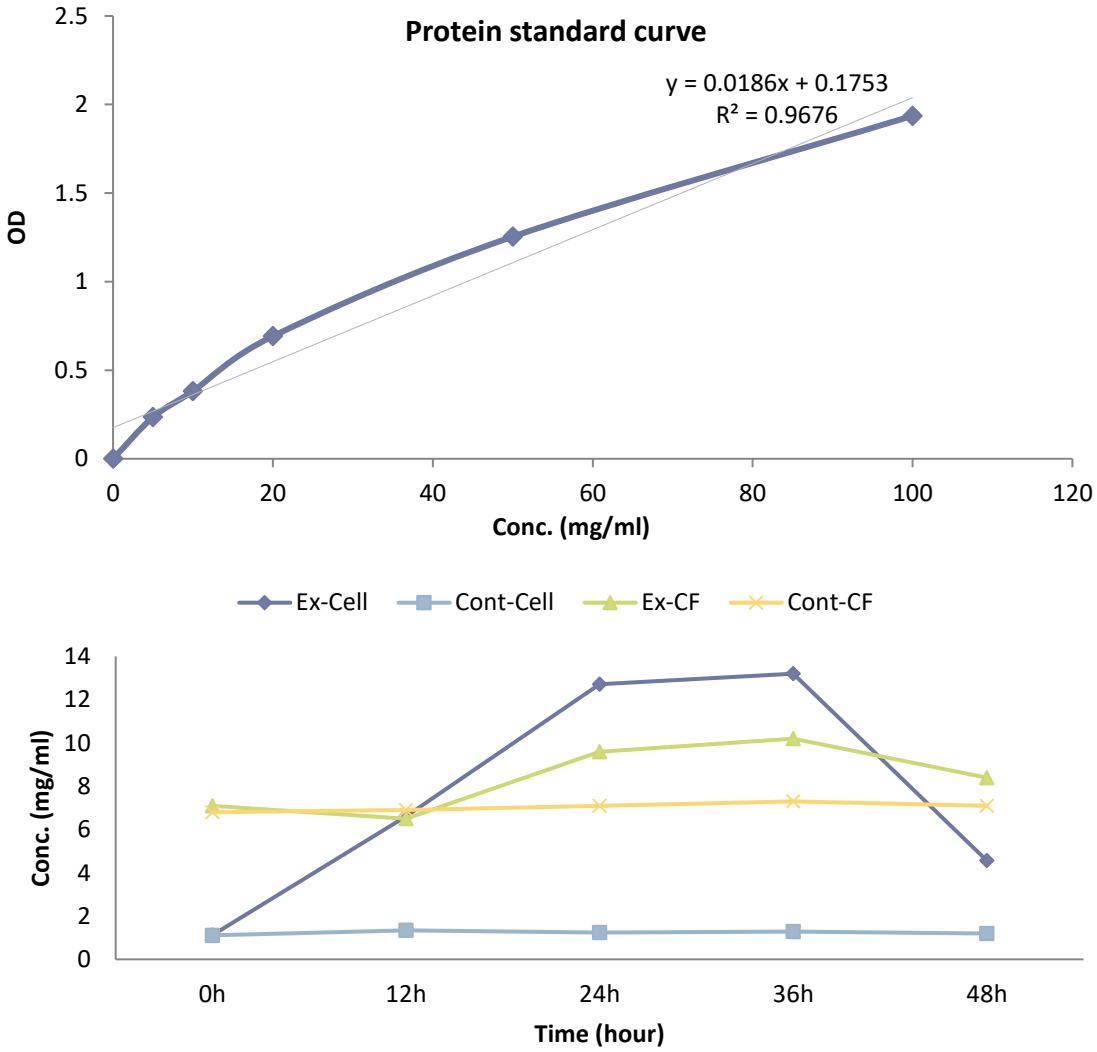


Fig.3.6: Time lapse analysis of total protein contents in coelomic fluid and in the coelomocytes of *B. thuringiensis* challenged and control *E. fetida*

3.3.3.2 SDS-PAGE analysis

Total protein content estimated in the coelomocyte lysate was approximately 5-6 times higher.

Thin bands of ~ 38 kDA and ~ 42kDA were found in the coelomic fluid lanes, which may be corresponding to the ECFP and CCF1 proteins described earlier. Some other smaller protein bands of unknown function were also found to be expressed. The cell lysate lanes could not be separated well may be because of different types of proteins expressed in the cells.

3.3.4. Dynamics of indigenous *Bacillus* species in coelomic fluid of *E. fetida* under *B. thuringiensis* challenge

Study of coelomic fluid at different intervals (12h, 24h, 36h and 48h), from each of the experimental and control plates, clearly reveals that the population of introduced *B. thuringiensis* started declining from the day 2 onwards and at day6 there was significant reduction in pathogen count (Fig. 3.7).

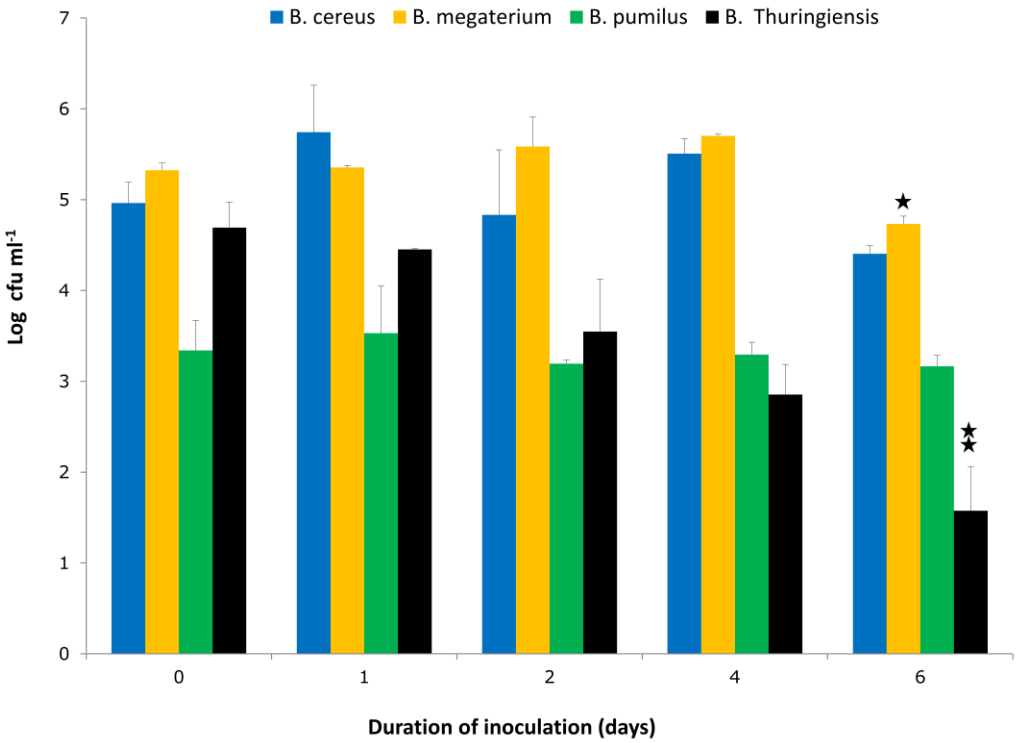


Fig.3.7: Dynamics of indigenous *Bacillus* species in coelomic fluid of *E. fetida* under challenge by *B. thuringiensis*

3.3.5. Comparative gene expression studies via whole transcriptomics in coelomocytes of *'B. thuringiensis* challenged' and 'control' *E. fetida*

3.3.5.1 Total RNA Isolation and cDNA library construction

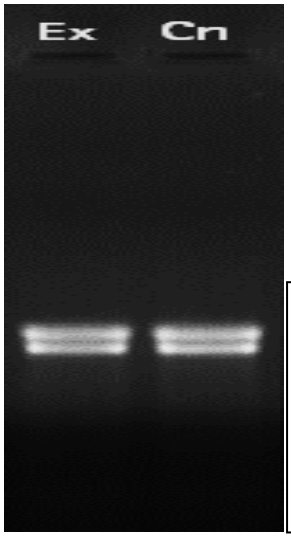


Fig.3.9: Isolated RNA from the Experimental (Ex) and Control (Cn) samples, run on on 1% denatured Agarose gel shows mRNA smears below the 23S and 16S bands.

The total RNA isolated from the two samples (Experimental and control) were Qualified using NanoDrop and Qubit. The results (Table- 3.1) indicated good quality RNA isolation sufficient for the transcriptomic pipeline.

Table 3.1: Quantification of receivedRNAsamples using NanoDrop/Qubit

Sr. No.	**Sample ID**	**NanoDrop Readings** (ng/µl)	**Qubit Readings** (ng/µl)	**NanoDrop** OD $A_{260/280}$	**NanoDrop** OD $A_{260/230}$
1	*Ex*	*247.8*	*124.0*	*2.13*	*1.64*
2	*Cn*	*266.5*	*144.0*	*2.13*	*2.27*

3.3.5.2 Sequencing and assembly

The paired-end (PE) libraries that were prepared from total RNA had mean fragment size distributions of 442bp and 457bp for the samples Ex and Cn, respectively. After the transcriptome was conducted, 34.67 Mb and 36.07 Mb high quality reads (QV>20) were obtained from coelomocyte samples of unchallenged and *B. thuringiensis* challenged *E. fetida*, respectively (Table 1). Total clean data generated were 5.5 Gb and 5.7 Gb from two groups, respectively (Table 1). The clean reads from the two libraries were assembled into 83,192 unigenes. The total length of all unigenes was 88,326,196 bp and average length was 1062 bp.

N_{50} was 1,799 bp.

The length distribution of assembled unigenes obtained from coelomocytes in unchallenged and *B. thuringiensis* challenged *E. fetida* revealed that most of unigenes ranged from 200 bp to 2000 bp, and approximately 6.5% of unigenes were over 3000 bp in length (Figs. 1–3). The sequence data of unigenes were deposited in the Sequence Read Archives (SRA) at NCBI under accession number SRP6138677.

3.3.5.3 Coding sequence (CDS) Prediction

TransDecoder predicted 28,770 CDS. Functional annotation of the CDS using DIAMOND (BlastX mode) against the NCBI 'Nr" database cloud classify 18,602 CDS most of which annotated against Capitella telata, a polychaete worm. 10,168 CDS were without Blast hit.

3.3.5.4 Functional unigenes annotation and classification

Majority of the blast hits were found to be against *Capitella teleta* which also belongs to the phylum Annelida.

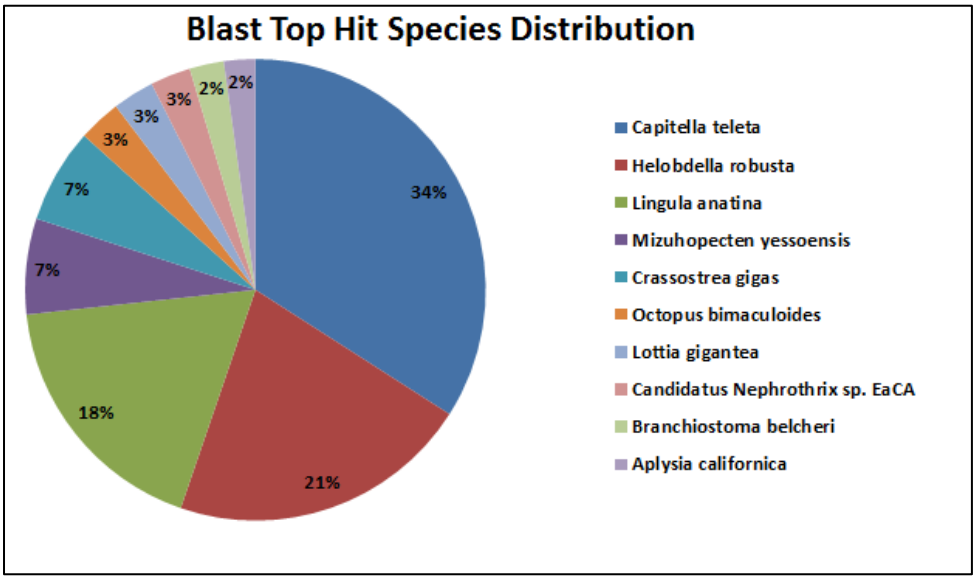


Figure 3.9: Top blast hit species distribution of pooled CDS

3.3.5.5 Gene Ontology Analysis

Gene Ontology (GO) analysis using B2G framework, categorized 4,138 and 4,206 CDS from the control and treated samples respectively to three main domains: Biological process (BP) Molecular function (MF) and Cellular

component(CC) (Table 2).

3.3.5.6 Functional Annotation of KEGG Pathway

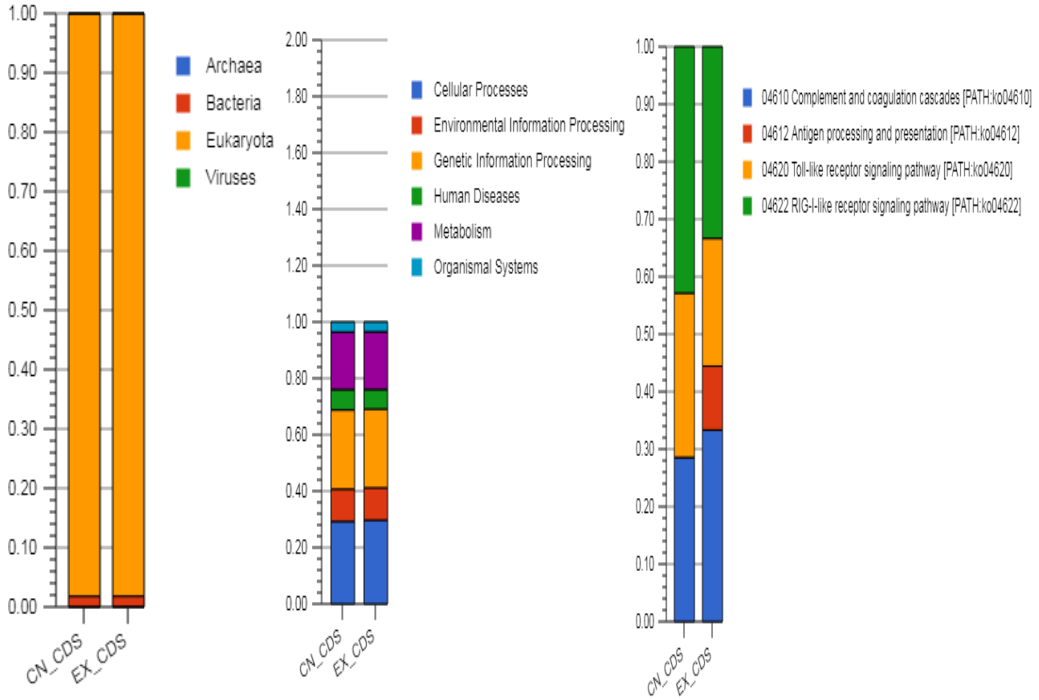


Fig.3.10: KO analysis of the Control (CN_CDS) and experimental (EX_CDS) CDS in MG-RAST

KEGG pathway analysis was carried out for all the individual set of CDS. A total of 5,343 and 5,425 CDS of the control and treated samples respectively could be categorized into 29 different pathways, the majority of which were signal transduction, transport and catabolism, translation, endocrine and immune system etc. The output of KEGG analysis includes KEGG Orthology (KO) assignments and corresponding Enzyme Commission (EC) numbers and metabolic pathways of predicted CDS using KEGG automated annotation server, KAAS. The results are briefed in the table S2. KO analysis of the Control (CN_CDS) and experimental (EX_CDS) CDS in the MG-RAST website also produces similar results (Fig: 3.10)

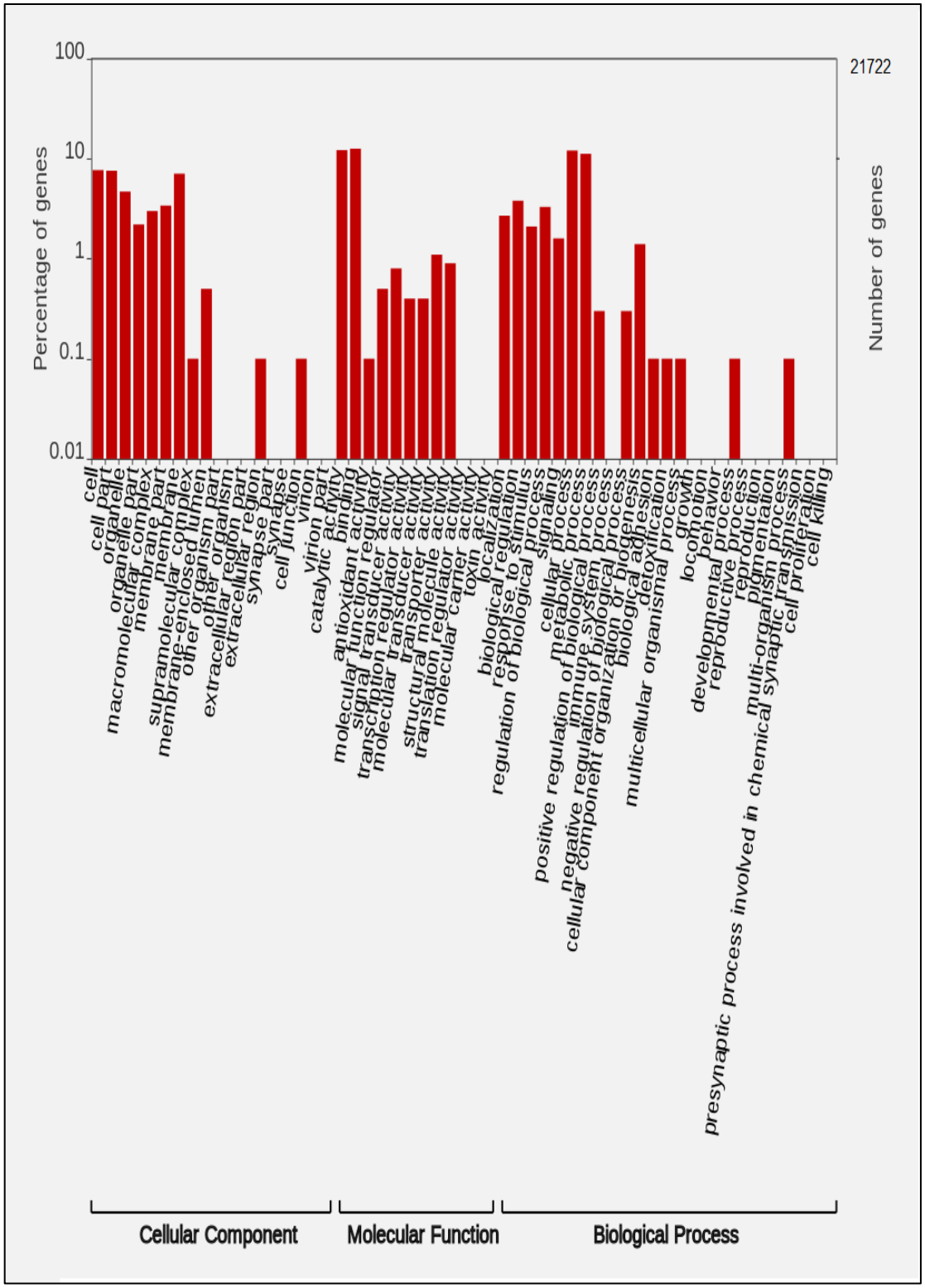


Fig. 3.11: WEGO Plot for Cn Sample

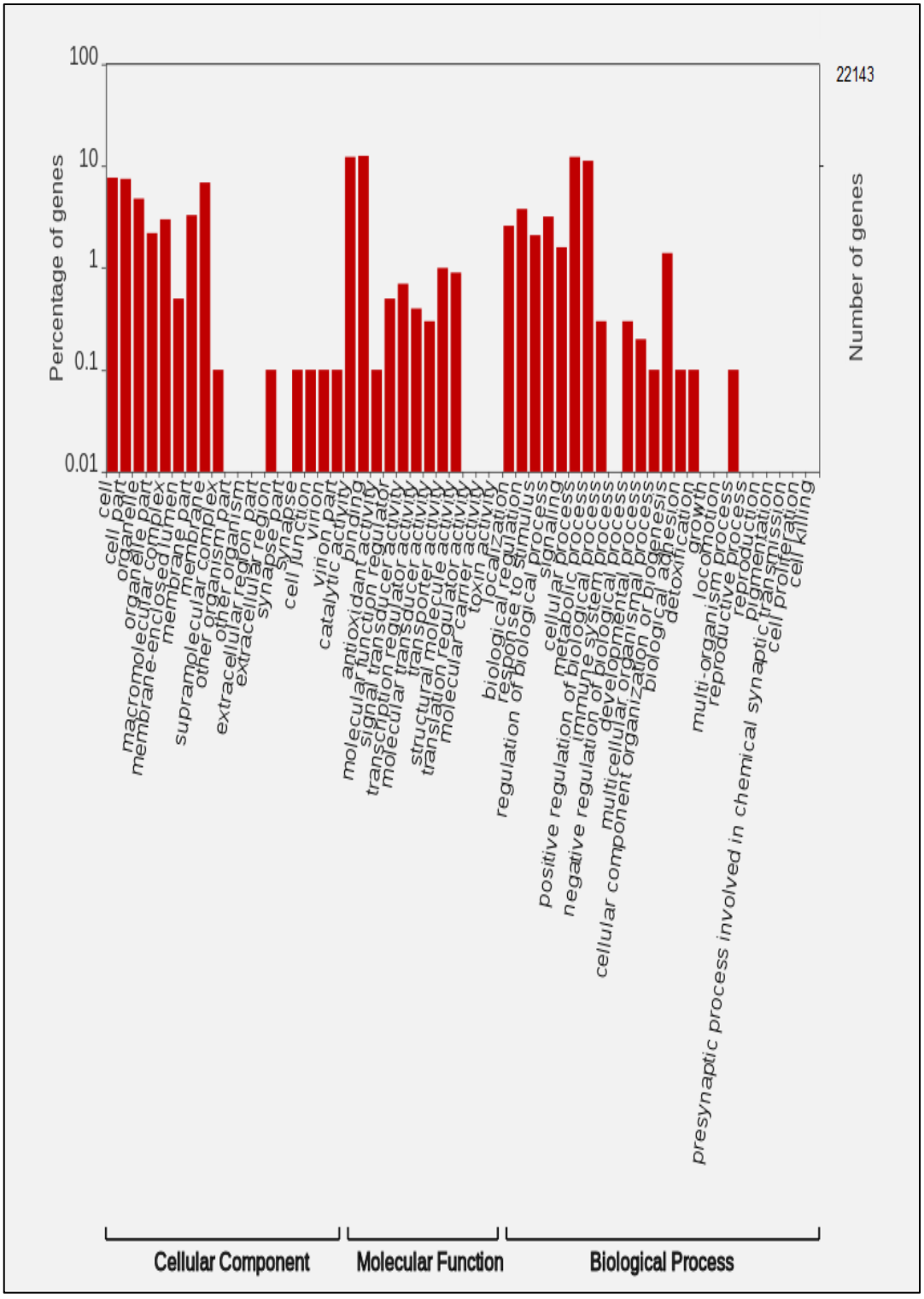


Fig. 3.12: WEGO Plot for Ex Sample

3.3.5.7 Differential Gene Expression Analysis

Differential gene expression analysis revealed that the expression of 646 genes was significantly up-regulated and 512 genes were significantly down-regulated in the treated group. 19,616 genes commonly expressed in both the samples. 98 unigenes were up-regulated more than two folds, and 15 unigenes were down regulated more than 2 folds, respectively (Table S1). Volcano plot (Fig 3.14) and Scatterplot (Fig 3.15) of differentially expressed genes have been created. Green dots represent the down-regulated (significant) and red dots represent the up-regulated (significant) genes. Fig 3.13 depicts the heat map of the top 50 differentially expressed genes (significant). COG analysis of DEGs showed that 621 DEGs (11.4%) were related with the immune system and diseases.

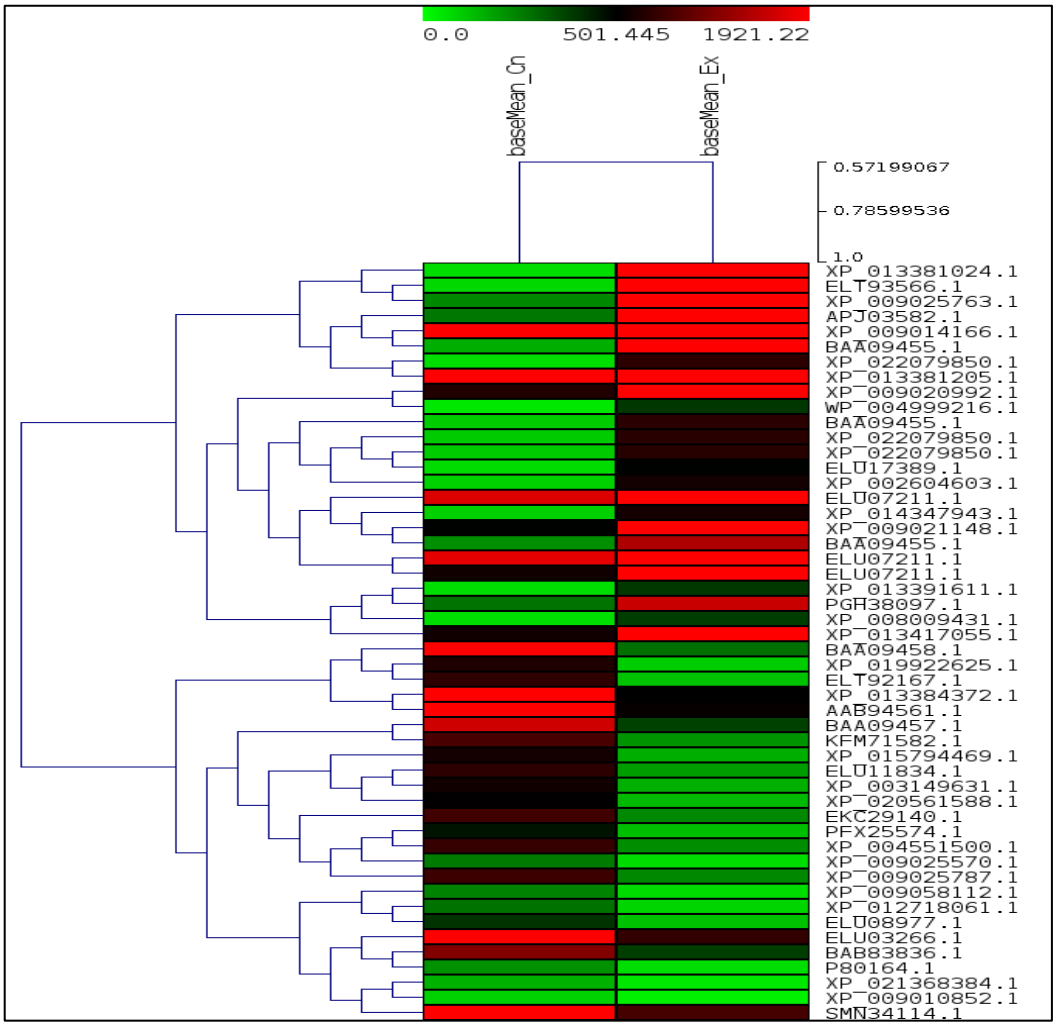


Fig 3.13: Heat map **of** the top 50 differentially expressed genes (significant); baseMean_Cn represents the normalized expression values for Cn sample **&** baseMean_Ex represents the normalized expression values for Ex sample.

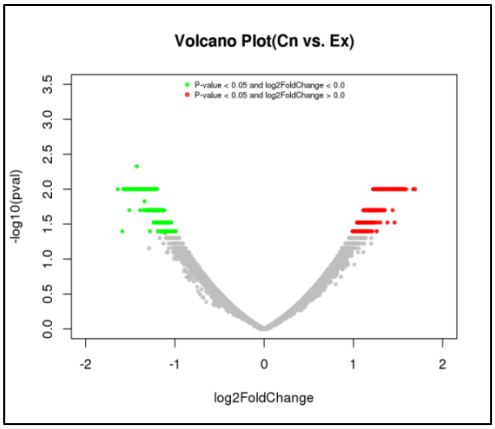


Fig. 3.14: Scatterplot of differentially expressed genes; green dots represent the downregulated (significant) and red dots represent the upregulated (significant) genes; baseMean_Cn represents the normalized expression values for Cn sample and baseMean_Ex represents the normalized expression values for Ex sample.

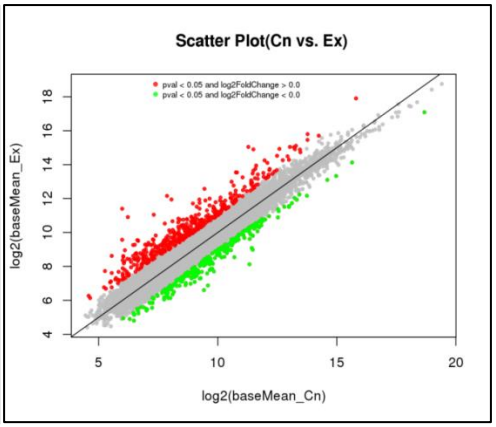


Fig.3.15: Volcano plot of differentially expressed genes; green dots represent the downregulated (significant) and red dots represent the upregulated (significant) genes.

Identification of immune-related genes and pathways

The immune related pathways significantly up-regulated were Toll-like receptor signaling pathway (PATH:ko04620), NOD-like receptor signaling pathway (PATH:ko04621), RIG-I-like receptor signaling pathway (PATH:ko04622), FcyR-mediated phagocytosis (PATH:ko04666), leukocyte trans-endothelial migration (PATH:ko04670).

Toll-like receptor (TLR) signaling pathway

Fifteen genes involved in the TLR pathway were found to be regulated in *B. thuringiensis* infected *E. fetida* compared to non-infected control. The DEGs, Lipopolysaccharide binding protein (LBP), TLR2, TLR13, FADD and TRAF3 were up-regulated (Fig. 3.16). The LBP expression was found to be proliferated by 2 fold suggesting an accurate and rapid response of *E. fetida* to bacterial invasion (Table 3). The TLR4 located adjacently down-stream of LBP was not found in *E. fetida,* compared to other animals, suggesting an alternative LBP sub-pathway.

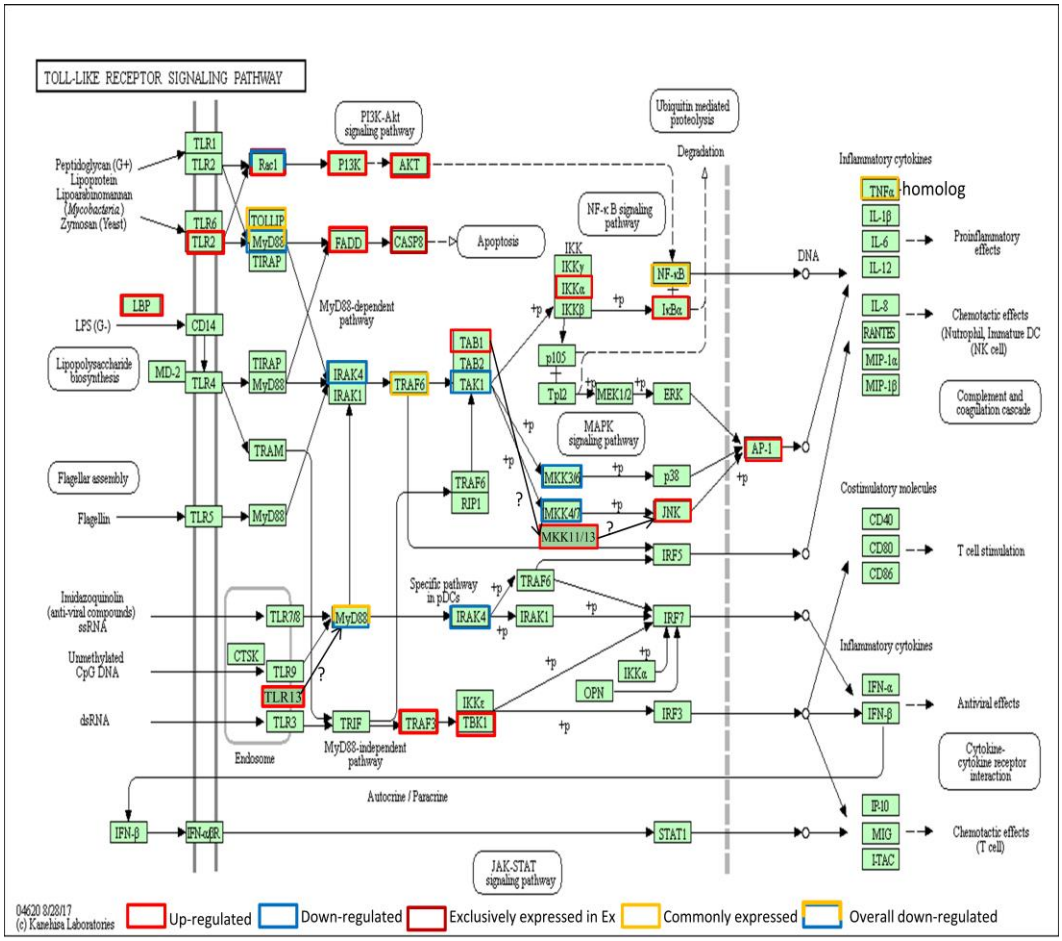


Fig. 3.16: Mediators of Toll-like receptor (TLR) signaling pathway that are differentially modulated in the experimental groups compared to the control (The KEGG pathway was adopted from Kanehisa Laboratories, Japan).

A specific study on TLR immunity pathway in a marine mollusk (*Crassostrea gigas*) showed that time-dependent patterned expression regulation of MyD88 along with TLR4 was observed while incubation with LPS. The expression of TLR4 in *C. gigas* was initially depressed and followed by a gradual proliferation. MyD88 is the first molecule acti- vated in the TLR signaling pathway (Muzio *et al.*, 1997; van dar Sar *et al.*, 2006), and has been found in vertebrates (Bonnert *et al.*, 1997; Wheaton *et al.*, 2007; Zhang *et al.*, 2012[a]) and invertebrates (Zhang *et al.*, 2012[b]; Ren *et al.*, 2014; Chu *et al.*, 2013) Adema *et al.* (Morozova *et al.*, 2009) found that snails challenged with *B. thuringiensis* showed more than 2 fold upregulation on JNK-interacting protein 3 which involved in signal transduction after TLR 4 activation. In our study, the MyD88 gene's expression was common

in both *B. thuringiensis* challenged and control animal with an overall down regulation, implying a similar regulation pattern (Fig. 3.16). The TLR2 and TLR13 expressions are up-regulated. Although the intermediate proteins of the TLR2 pathway are observed to be up-regulated leading to caspase 8 mediated apoptosis which is only seen in experimental worms along with activation of AP-1 complex leading to the expression of pro-inflammatory cytokines.

RIG-I-like receptor signaling pathway

RIG-I is one of the three RLR signaling pathway reported to be mainly involved in intracellular virus sensing. Non-self RNA appearing in a cell as a result of intracellular

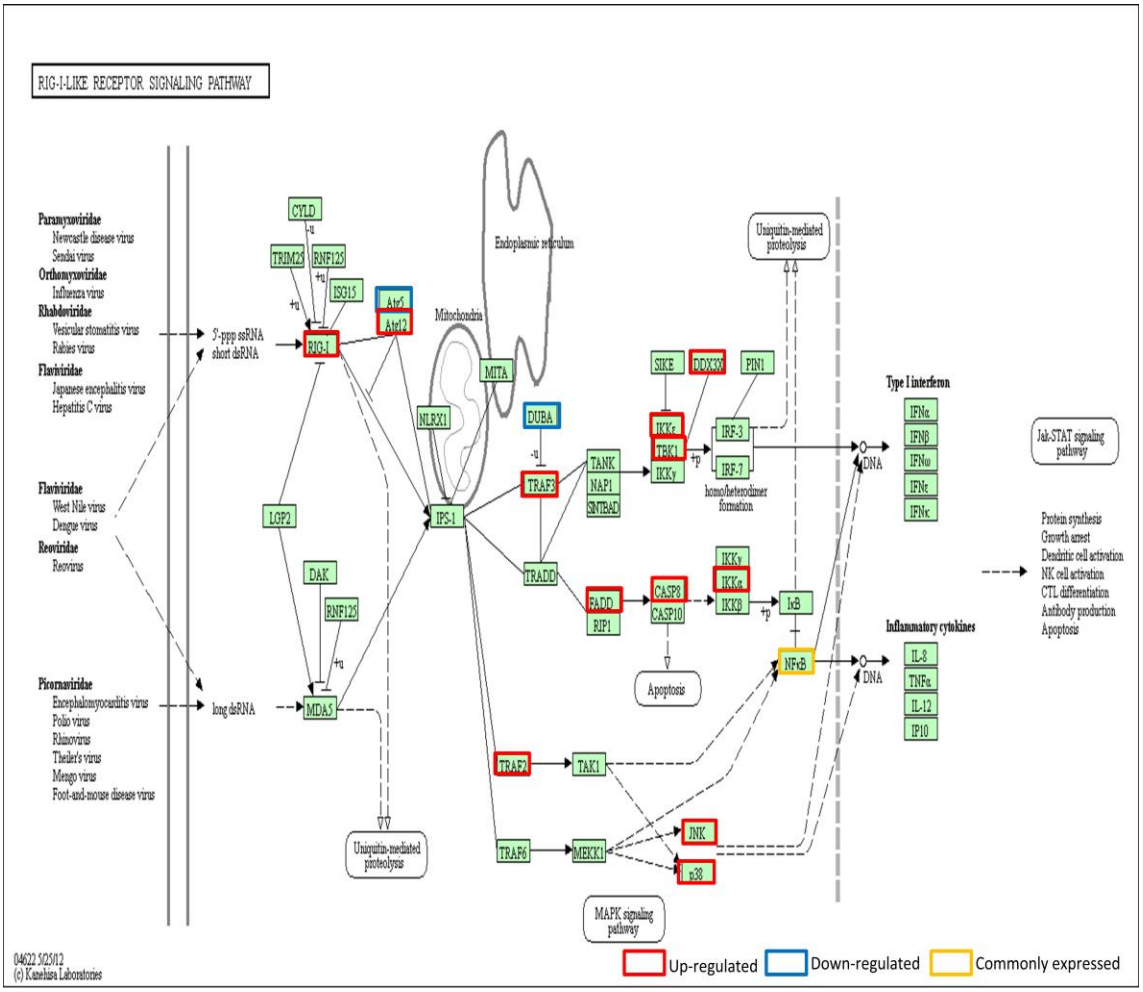


Fig. 3.17: Mediators of RIG-I-like receptor signaling pathway that are differentially modulated in the experimental groups compared to the control (The KEGG pathway was adopted from Kanehisa Laboratories, Japan).

viral replication is recognized by a family of cytosolic RNA helicases termed RIG-I-like receptors (RLRs). Fourteen genes involved in the RIG-I pathway were found to be regulated in *B. thuringiensis* infected *E. fetida* compared to non-infected control. The DEGs, RIG-I, Ate12, TRAF2, TRAF3, FADD, CASP8, IKK, p38, JNK, TBK1 and DDX3X were up-regulated (Fig. 3.17). Inhibitors in the pathway such as Ate5 and DUBA were down regulated in the *B. thuringiensis* injected worms compared to the control ones. The IPS-1 located adjacently down-stream of RIG-I was not found in *E. fetida*, compared to other animals, suggesting an alternative RIG-I sub-pathway. Upon recognition of viral nucleic acids, RLRs recruit specific intracellular adaptor proteins to initiate signaling pathways that lead to the synthesis of type I interferon and other inflammatory cytokines, which are important for eliminating viruses.

C-type lectin pathway

C-type lectin receptors (CLRs) are a large super-family of proteins characterized by the presence of one or more C-type lectin-like domains (CTLDs). CLRs function as pattern-recognition receptors (PRRs) for pathogen-derived ligands in dendritic cells, macrophages, neutrophils, etc., Ten genes involved in the C-type lectin pathway were found to be regulated in *B. thuringiensis* infected *E. fetida* compared to non-infected control. The DEGs, p38, JNK, AP-1, CYLD and CALM were up-regulated (Fig. 3.18). Two novel lectin receptors Galectin and Rhamnose binding lectins were found to highly up-regulated in the experimental worms after bacterial challenge. Many other factors involved in this pathway viz. Syk, Pak-1, Bcl-3, Bcl-10, PLC$_{\gamma}$2 and NEMO were highly expressed in both experimental and control group of animals. Upon ligand binding, CLRs stimulate intracellular signaling cascades that induce the production of inflammatory cytokines and chemokines, consequently triggering innate and adaptive immunity to pathogens. *B. thuringiensis* is a potential pathogen for many invertebrate groups (WHO, 2003). In this study, the differently expressed unigenes (DEGs) were analyzed to enrich our knowledge and the DEGs information obtained will help to understand molecular immune mechanism of

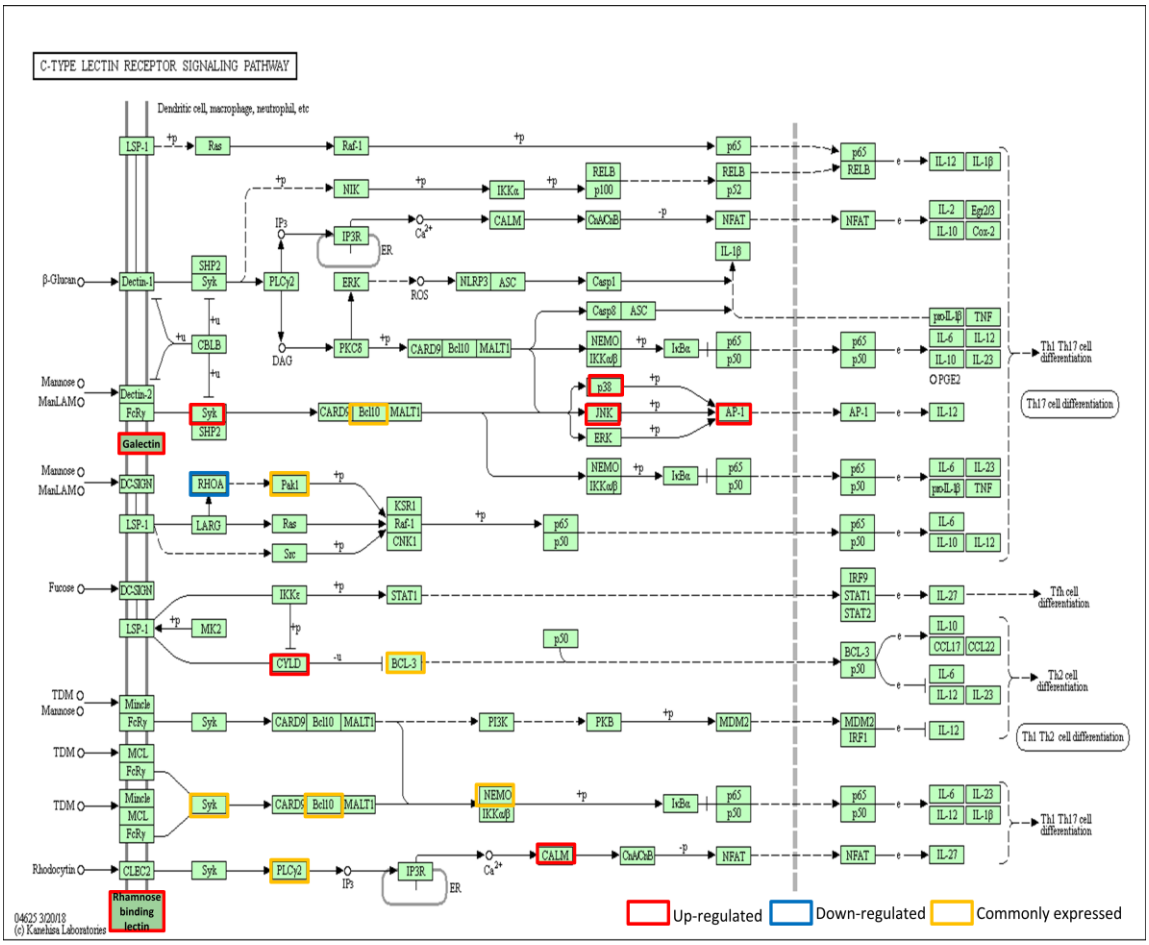


Fig. 3.18: Mediators of RIG-I-like receptor signaling pathway that are differentially modulated in the experimental groups compared to the control (The KEGG pathway was adopted from Kanehisa Laboratories, Japan).

E. fetida in response to bacterial infection. *E. fetida* is a cosmopolitan representative of the Phylum Annelida which is the first triploblastic eucoelomate group, thus to be a potentially valuable representative species for providing information on evolution of immune system in the eucoelomates (Qlu, 1991; Zhang *et al.*, 2017; Sun *et al.*, 2016).

However, studies on the generation, development, differentiation and immune system of *E. fetida* are limited due to the lack of transcriptomic and genomic information. The transcriptomic and genomic information obtained in the present study will be helpful for future studies of *E. fetida*.

CHAPTER 4

WOUND HEALING & REGENERATION IN *EISENIA FETIDA*

4.1.Introduction

The process of wound healing and regeneration in the amputated body parts is studied in few isolated groups of animals. This bio-medically relevant field of scientific research still have the potential in revealing knowledge about the differential regeneration capability of different animals and their cellular and molecular basis. Some diploblastic groups, acoelomate triploblasts, annelids, insects, asteroid echinoderms, fishes, amphibians and few lizards has the ability to regenerate amputated body parts (Sa´nchez Alvarado, 2000; Myohara, 2004; Bely & Nyberg, 2010). Some lower group of animals like hydra, planarians can regenerate complete individuals from small pieces of the body while in the complex forms like fish, amphibians and lizards replacement of lost fins appendages or tail respectively is observed. In between these lower and higher groups of animals, annelids occupy an intermediate position being the simplest eucoelomate triploblast. Hence, reasonably annelids particularly earthworms have long been subjects of regeneration research. Regeneration in annelid groups like nereids (errant polychaetes), enchytraeids (Sedentaria), naidids (Sedentaria), capitellids (Sendentaria), and Lumbricids (oligochaetes) have been studied (Berrill, 1952). Among these different groups of annelids successful regeneration of posterior segments have been reported, but, the ability to regenerate anterior segments were found to be futile in most cases (Bely, 2014). Because of the easy availability, convenience in culture and handling in the laboratory and rapid regenerative power, the earthworm *E. fetida* has been preferred in regeneration research (Edwards & Bohlen, 1996). Earlier reports regarding the correlation of the survival rate and regeneration capacity of immature and adult *E. fetida* are available which indicate that survival and regeneration proportionally depend on the number of annules retained in the amputated earthworm (Xiao *et al.*, 2011). The processes of wound healing and cell proliferation at the wound site after amputation were described by Hill (1970). The nature of tissue interactions between the epidermis, nerve cord, and gut endoderm in regenerating earthworms was reported by Fitzharris & Lesh (1972). Few histological studies have postulated that all of ectoderm, mesoderm and endoderm proliferate near the wound site (Clark, 1972; Cornec *et al.*, 1987). Neoblast cells, observed near the severed region in many clitellate annelids, has been in focus for their origin and function (Cornec *et al.*, 1987; Tadokoro *et al.*, 2006; Myohara, 2012;

Sugio *et al.*, 2012). These neoblast cells have been postulated to form 'regeneration blastema' in the regenerating region. The blastema is said to form most of the missing structures (Müller *et al.*, 2003; Müller, 2004; Müller & Henning, 2004; Zattara & Bely, 2011). Tweeten and Reiner (2012) pointed out that gut endoderm has an important role in the regeneration of gut in the amputated animals. Park *et al.*, (2013) described circular muscle layer to be the major origin of blastemal cells during regeneration in *Eisenia andrei* with histological evidences of de-differentiation restricted to 1 and 3 day(s) after amputation. There was no available report on the origin of vital tissues like epithelium, circular muscle layer or gut wall in the regenerating region of *E. fetida*. Under this scenario, further research attempting revelation of mechanism of tissue reorganization and regeneration of body wall, muscle layers and gut lining is needed.

Some authors have recently speculated about the possible contribution of other tissues in the blastema formation (Park *et al.*, 2013). In this study, a histological investigation of de-differentiation followed by re-differentiation is carried out from day 1 to day 11 following amputation of posterior one-third of the *E. fetida*'s body. In particular we investigated the importance of the vital organs in relation to the survival and regeneration of the amputated fragments in *Eisenia fetida* focussing mainly into the process of regeneration with the classical histology tool.

4.2. Materials and methods

4.2.1. Amputation experiment of *E. fetida*

Earthworms *(E. fetida)* were reared in cow manure at 22 ± 1◦C, in a laboratory climate chamber for several generations. Cocoons were collected from the culture beds and reared in small aerated plastic chambers with adequate cow dung manure, humidity, and temperature maintenance. Newly hatched earthworms took about 8 weeks to develope clitella until which they were regarded as immature. Amputation experiments were done on the adult *E. fetida* by severing the worm on four different amputation sites (Fig. 4.1). Adult *E. fetida* has approximately 97 segments (Halder, 1999) including a well-developed clitellum on annule number 25–31 (Fig. 4.2a) and generally weighs 0.35±0.03g.

4.2.1.1 Amputation, rearing and estimation of survivability of amputated parts

Four types of amputations (in between segments 10-11 [A1 (Anterior); P1, (Posterior)], 24-25[A2 (Anterior); P2, (Posterior)], 31-32 [A3 (Anterior); P3, (Posterior)] and 66-67[A4 (Anterior); P4, (Posterior)] were done. Sets of amputated segments {A1, P1}, {A2, P2}, {A3, P3}, and {A4, P4}, were selected as amputation sites to represent anterior segments without esophagus and crop (A1), without clitellum (A2), with clitellum (A3), and body part (2/3rd) with most known vital organs (A4); on the other hand, the posterior segments with esophagus and crop (P1), with clitellum (P2), without clitellum (P3)and body part (1/3rd) without most known vital organs (P4).

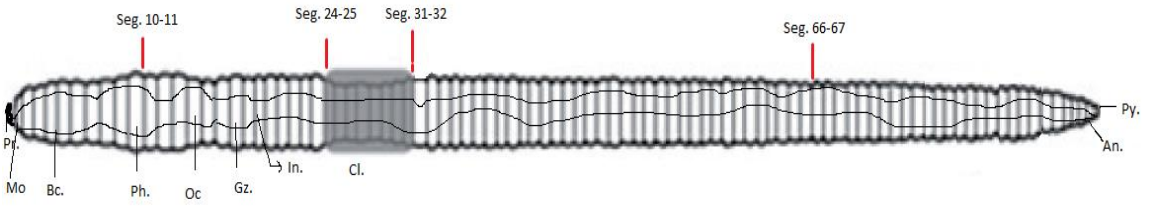


Fig. 4.1: Schematic drawing of an adult *E. fetida* depicting the four selected regions of amputation (An-Anus, Bc- Buccal cavity, Cl- Clitellum, Gz- Gizzard, In- Intestine, Mo-Mouth, Esophagus, Ph- Pharynx, Pr- Prostomium, Py- Pygidium).

The above-mentioned scheme of amputation was designed to observe the effects of position of amputation on survival and recovery rate (weight gain). Controls were *E. fetida* with no amputations. All amputations (as per scheme) were replicated three times with 20 earthworms per treatment; each kept in an aerated plastic container with cow manure and hydrated tissue paper (Fig.4.2c). Amputation was done following standard method (Xiao *et al.*, 2011).

4.2.1.2 Estimation of body weight alteration of amputated parts

The survival rates, growth in terms of body weight (including gut content) were checked every 2 days for the first week and then weekly up to 10 weeks.

4.2.1.3 Periodic histological observations

For histological studies, surviving representatives of A4 (as A4 regenerated most successfully) were selected and sacrificed. The nearby region of amputation of each A4 was severed out with a sterile sharp scalpel and processed for histological observation

(Cui, 2011). Briefly, the whole tissue was fixed in Bouin's fixative for 18hours. After ethanol washings at different increasing grades upto absolute for dehydration, xylol was used for clearing. The tissues were then infiltrated with paraffin (MT56-58◦C) by warming in oven. Microsections (6 µm) were cut (using Lipshaw type rotary microtome, York Scientific Industries Pvt. Ltd.) and mounted on slides. Microsections were stained with haematoxylin and eosin and observed under a light microscope (Dewinter, Model DEW/002, Dewinter Optical Inc., INDIA) and photographed with DIGIEYE powered through USB 2.0 equipped desktop computer.

4.2.2. Dynamics of indigenous *Bacillus* species in coelomic fluid of the amputated body parts of *E. fetida*

4.2.2.1 Periodic withdrawal of coelomic fluid of E. fetida

Coelomic fluid from the survived body part (A4) of the earthworms was periodically (at day1, 3, 7 and 11) were pooled out in separate siliconized centrifuge tubes. Approximately 100µl from ~ 20 A4 animal (part) was collected at each study day. Sterile, sharpened capillary tubes were used to withdraw the coelomic fluid from surface cleaned earthworms (Bilej *et al.,* 1990). Sterile 1x PBS, pH7.4, prepared by using PBS powder (Himedia M1866) and MiliQ water was used to wash the coelomocytes. Total process was done on ice.

Coelomocytes were separated by centrifugation (100xg for 10 min) and cell-free coelomic fluid was collected and kept at -20 °C. Isolated cells were resuspended in PBS

4.2.2.2 Serial dilution, plating and incubation

The supernatant collected in the previous step was serially diluted in sterile PBS. Multiple dilutions and plating on HBA plates were done to enumerate the number of indigenous *Bacillus* spp. content (CFUs) in the coelomic fluid during different stages of regeneration.

4.2.3. Compositional changes of coelomic fluid relating to coelomocytes of selected regenerating *E. fetida*

4.2.3.1 Isolation of coelomocytes

Coelomocytes separated on the previously mentioned step on day 1,3,7 and 11 from the A4 body part of *E. fetida* were isolated, identified, counted (DC) and studied for their phagocytic nature.

4.2.3.2 Differential count (DC) of coelomocytes

Differential counts were made for each group of regenerating worms by placing a drop of coelomocyte suspension in PBS onto a clear glass slide and making a smear. After air-drying, coelomocytes were fixed in methanol for 20 s and stained with modified Wright's stain (Sigma WS16) for 20s. One hundred leukocytes were classified under oil immersion using 100 x objective lens.

4.2.3.3 *Study of phagocytic behaviour of coelomocytes*

Time-lapse study of phagocytosis percentage at 30min, 1h and 1h 30min from the coelomocytes collected from A4 regenerating segment of *E. fetida* was performed following the same process described in section 3.2.2.5.

4.3. **Results**

4.3.1 Amputation of *E. fetida*

4.3.1.1 *Survivability of amputated parts*

The survival rates of the anterior and posterior segments of the amputated *E. fetida* (as per scheme) are calculated as survival percentage for each amputated part by using the formula, (surviving amputated parts / total initial number of that amputated part i.e.20) X 100. A2, A3, and A4, have shown the higher rate of survival (78%, 78% and 89% respectively after 10 weeks) during the process of regeneration, while A1 failed to survive and 89 % died within day 1 and the remaining (11%) in day2 (Fig 4.3a).

All the posterior amputated parts P1, P2, P3, P4, have strikingly low survival rates during the process of regeneration (Fig. 4.3b). All the P1 parts died within 2 weeks of amputation. P2, P3 and P4 have 22%, 44% and 33% survival rate respectively up to 10 weeks of the experiment period.

4.3.1.2 *Body weight alteration of amputated parts*

The mean body weights of the survived amputated parts were taken into account for calculating the alteration of body weight. The reduction in weight with time was noticed among the survived anterior amputated parts, A2, A3, A4, up to 6 weeks (Fig. 4.4a). A3 and A4 gained weight from 7th week which even surpassed their initial (Day 0) weight. The weight of amputated amputated part, A2 remained unchanged. All the posterior amputated parts (P1, P2, P3 and P4) of the amputated *E. fetida* have shown progressive decline in body weight (Fig.4.4b). The decrease in body weight of the posterior amputated parts was found to be much faster than the decrease shown by the anterior parts in the initial days of observation.

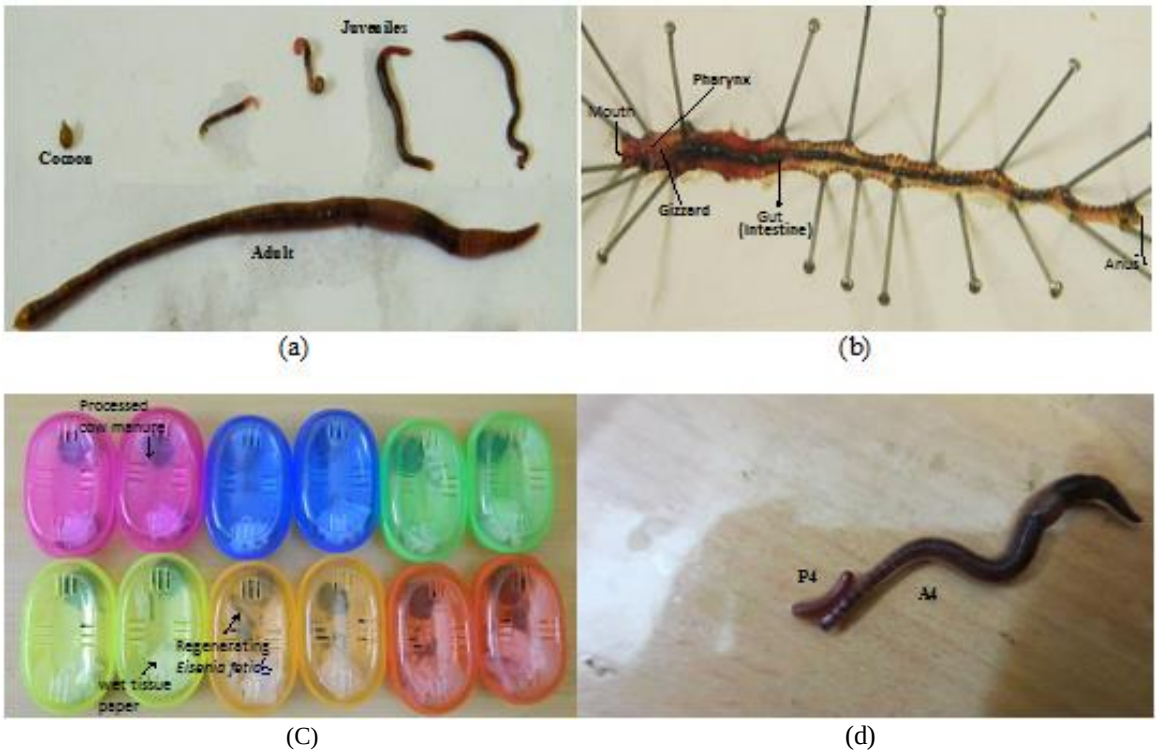
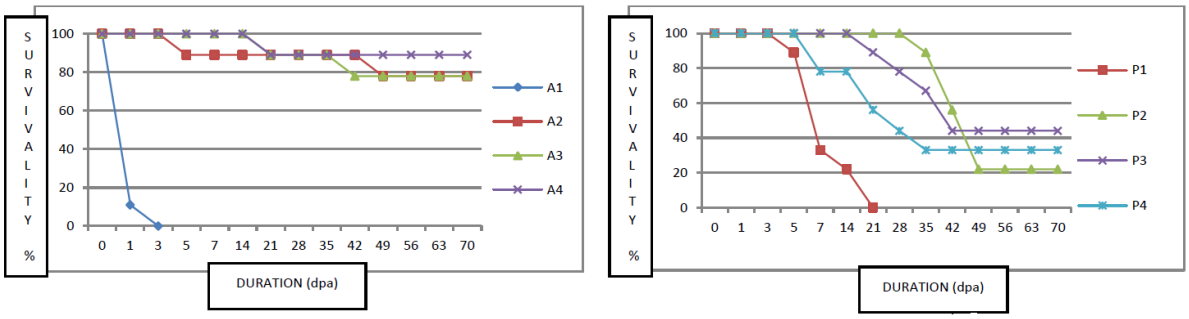
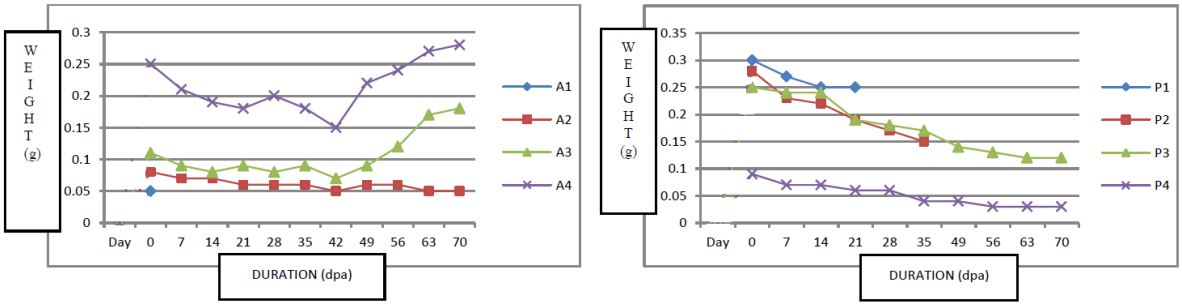


Fig.4.2: Photographs taken during different phases of *E. fetida* regeneration experiment. (a) Adult *E. fetida* compared with juveniles (b) Dissected adult *E. fetida* with intact gut (c) Rearing containers of individual amputated *E. fetida* (d)11 dpa A4 amputee with regenerated posterior part ; survived P4 amputee without regeneration



(3a) (3b)

Fig. 4.3 Survival (%) of the amputated *E. fetida* in respect to days post amputation. (a)Anterior parts (b)Posterior parts



(4a) (4b)

Fig. 4.4 Body weight (g) of the amputated *E. fetida* in respect to days post amputation. (a) anterior parts (b) posterior parts.

4.3.1.3 *Histological observations*

Since A4 demonstrated the most successful development process in terms of gaining net body weight following amputation, the tissue level dynamics during regeneration of A4 amputated part with the aid of histological techniques was studied. It was revealed from the histological section of one-day- post- amputation (1dpa) A4 part that, all the major tissues like epidermal layer, circular muscle layer, longitudinal muscle layer, gut wall epithelium show anatomical variation than the normal tissue organization. Proliferating cells appeared in all the above mentioned tissues (Fig. 4.5 IIa). Neoblast cells were clearly observed to originate from de-differentiating longitudinal muscle layer and the chloragogue tissue in the coelomic face of those tissues (Fig. 4.5 IIc). The longitudinal muscle cells delaminated in the coelomic space, proliferated and formed neoblast cells.

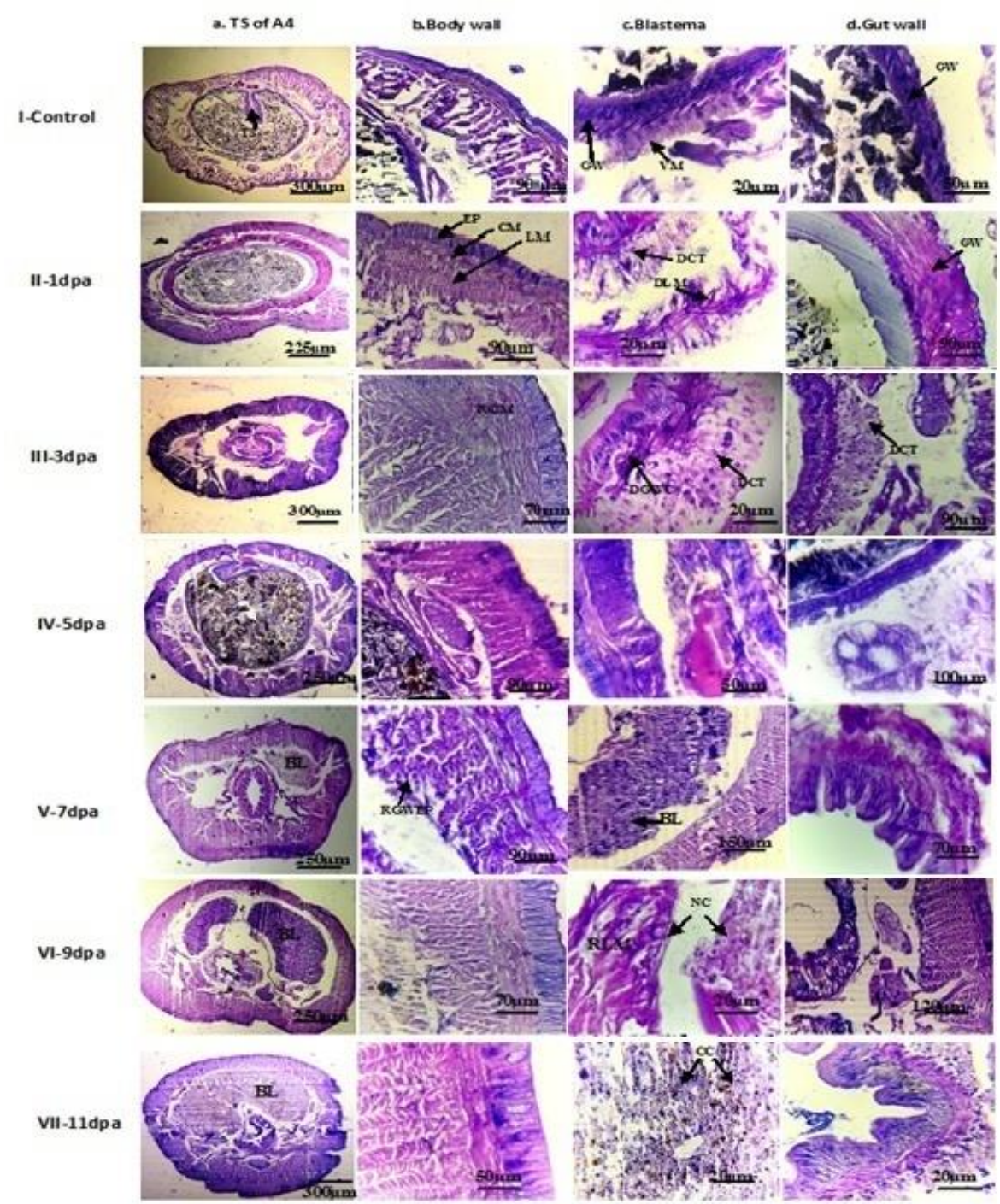


Fig.4.5: Photographs of histological preparations of tissues in regenerating region of A4 amputee up to 11 days post amputation and corresponding regions from control *E. fetida*.
(BL-Blastema, CC-Chloragogenous Cells, CM- Circular muscle, CT-Visceral mesoderm, DCT-De-differentiating chloragogue tissue, DGWC- De-differentiating gut wall cells, DLM- De-differentiating longitudinal muscle, EP- Epithelium, GW- Gut wall, LM- Longitudinal muscle, NC- Neoblast Cells, RCM- Re-differentiating circular muscle, RGWEP- Re-differentiating gut wall epithelium, RLM- Re-differentiating longitudinal muscle, TS-Typhlosole)

The chloragogue tissue layer grew thicker than the normal (Fig. 4.5 Ic, IIc & IIIc) and multiple layers of proliferating neoblast cells could be found. By day 3 (3dpa), all the major tissues were observed to develop growth by the neoblast cells. Regeneration blastema (cells in mass) was developed in the coelomic space (Va, VIa, VIIa). The muscle layers of the body wall were found to get re-established by 7dpa. The gut wall epithelium reformed by 7-9dpa. The blastema cells were found to contribute the re-differentiating longitudinal muscle layer (Fig. 4.5 VIc). (iv) Since histological surveillance was continued even after 10dpa, the complete differentiation of the mid dorsal groove of intestine (typhlosole) was found to be taking place by 11dpa. (v) Interestingly, blastema tissue was located and photographed to be formed not only from de-differentiating longitudinal muscle cells (Fig5 IIc) but also from chloragogen cell mass (Fig. 4.5 IVc). The other organs viz. excretory segmental nephridia were also found to develop fully by day 11 (Fig 4.5 VIIa).

Eisenia is the simplest eucoelomate resembling basic body plan of the higher groups of animals. The digestive system in *Eisenia* resembles a straight tube with various degrees of regional specialization at the anterior end (Fig. 4.2b). Below the epidermis there are thin circular and thick longitudinal muscle layers. The longitudinal muscle layer is bounded internally by the parietal peritoneum (Fig. 4.5 Ia). Mouth leads to a short buccal cavity, muscular pharynx, and esophagus. The posterior esophagus bears a crop, and one muscular gizzard. The rest of the digestive tube is a straight intestine leading to a short proctodeal hindgut and anus located on the pygidium (Brusca & Brusca, 2003). In the present study, it was observed that, presence of the anterior organs of the digestive system was vital for the survival of the severed worms. Prior to the present study, linear correlations between the survival frequencies of amputated earthworms with the differential length of the amputated segments were observed. Xiao *et al.*, (2011) reported that greater are the chances of survival when lesser segments were amputated in *E. fetida*. In our study, we have demonstrated that absence of vital organs in the amputated segment (A1) did not support survival vis-a-vis regeneration of the amputee *E. fetida* worms. Hence, anterior segments except the A1 amputee survived (survival rate >70%). This might be due to intactness of the anterior gut portions (vital for feeding and uptake of nutrition from the ingested food which the amputees draw after recovery from the severe

shock). It was observed that all the live anterior amputees lost their body weight to some extent during the initial period (1st to 6th week), but from 7th week onward gradual increase in body weights were noted in case of A3 and A4 amputees. The A2 amputee survived (for > 8 weeks) but failed to grow and regenerate. This may be because of the fact that they did not have sufficient intestine and associated tissue like typhlosole that increases the surface area of the intestine in *Eisenia* (Fig. 4.5 Ia). Many layers of chloragogen cells remain associated with the intestinal wall and typhlosole. These chloragocytes contain greenish, yellowish, or brownish globules and are reported to serve as producer and store-house of glycogen and lipids along with other important functions. The blastema tissue from 9 dpa was observed to be mainly occupied by the chloragogue tissue (Fig. 4.5 VIIc). Hence, the function of this tissue in the later stage of development seems to serve nutrition to cellular growth and proliferation of the regenerating part. All the severed parts containing the posterior region survived the amputation shock better than the counterpart A1 amputee but failed to gain weight. Few posterior fragments survived for as long as 60 days, presumably by slow exhaustion of nutrition which was stored in the chloragogenous tissue before amputation.

4.3.2. Dynamics of indigenous *Bacillus* species in coelomic fluid of the amputated body parts of *E. fetida*

Study of coelomic fluid at different intervals (at day1, 3, 7 and 11), from the survived body part (A4) of *E. fetida*, reveals a population fluctuation of the major Bacillus spp. An adventitious pathogen, *B. thuringiensis* is observed at very small number compared to other *Bacillus* spp. in the amputated A4 segment from day up to Day7. This pathogen is completely out populated by other indigenous *Bacillus* species by day11. The population of different indigenous *Bacillus* species is considerably lowered than the normal bacilli count in control animals. The number of *B. cereus* is significantly reduced on Day 7 ($p < 0.01$), which is slightly recovered on Day 11. Population of *B. megaterium* seems to have the least reduction ($p < 0.05$) compared to other bacilli. This outcome indicates that *B. megaterium* is favoured in the coelomic fluid than other bacilli (Fig. 4.6).

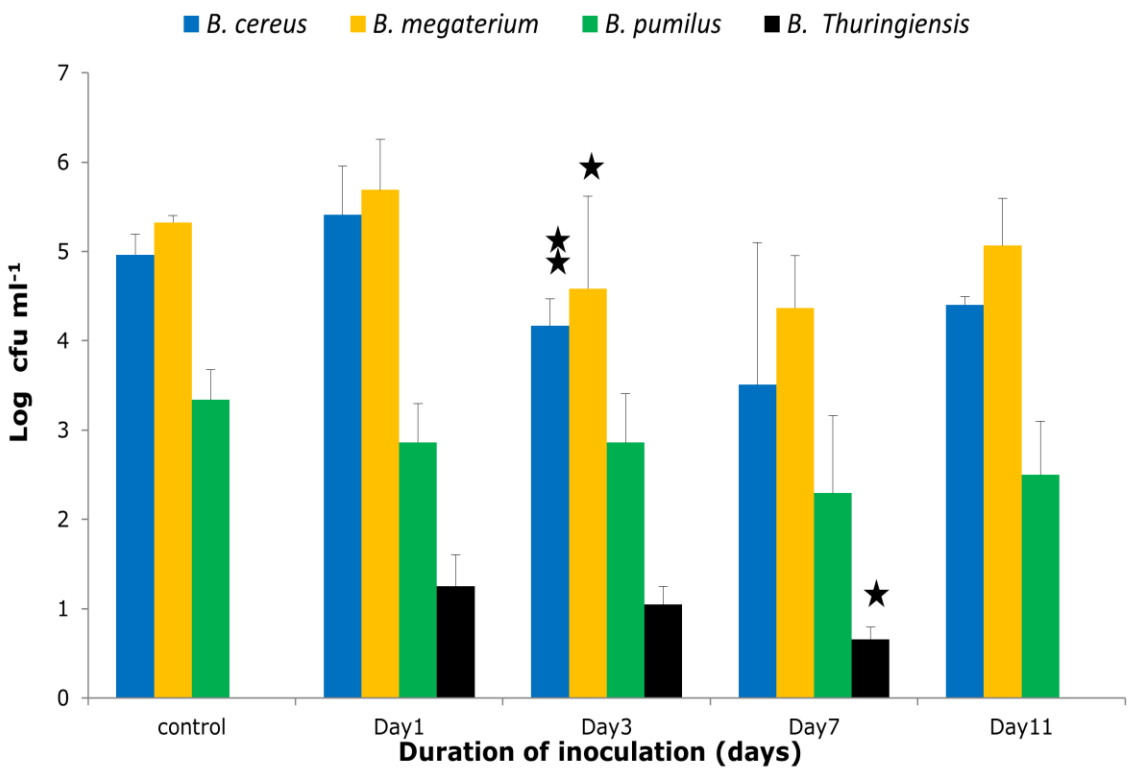


Fig 4.6: Dynamics of indigenous *Bacillus* species in coelomic fluid at different intervals (at day1, 3, 7 and 11), from the survived body part (A4) of *E. fetida*

4.3.3. Compositional changes of coelomic fluid relating to coelomocytes

4.3.3.1 Differential count (DC) of coelomocytes

Light microscopic studies of coelomocytes at different intervals (at day1, 3, 7 and 11), from the survived body part (A4) of *E. fetida* revealed modulation of the four group of coelomocytes, the normal count of which are- amoebocytes (or hyaline amoebocytes)(23±9%), large granulocytes (or basophils)(18±7%), eleocytes (or chloragocytes) (6±3%) and small granulocytes (or granular amoebocytes) (51±8%).

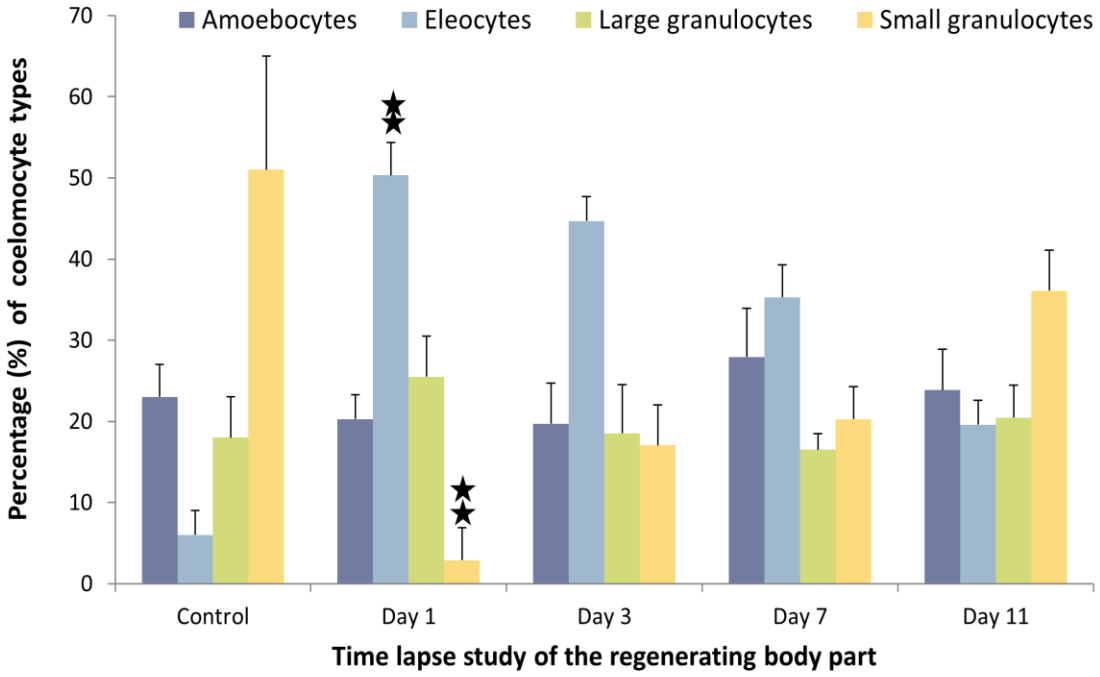


Fig 4.7: Differential count (DC) of coelomocytes at different intervals (at day1, 3, 7 and 11), from the survived body part (A4) of *E. fetida*

The population of eleocytes was significantly increased and the population of small granulocytes was significantly decreased on day 1 after amputation of the A4 fragment in comparison to the control animals. This deviation of the differential count of coelomocytes is minimized to some extent by day 11 (Fig. 4.7).

4.3.3.2 Study of the phagocytic behaviour of coelomocytes

Both amoebocytes and granulocytes were frequently found with inclusion bodies including nano-carbon particles. These coelomocytes were frequently observed to accumulate towards the activated carbon particles, which indicate their increased efficiency of chemotaxis and phagocytosis. Active phagocytic cell increased significantly ($p < 0.001$) on day 1 after amputation of the A4 fragment in comparison to the control animals. This difference in the number of active phagocytic cells remained significant (p

< 0.05) till day11 (Fig. 4.6). The fraction of amoebocytes and granulocytes with visible inclusion bodies in respect to the total number of coelomocytes was considered to be the phagocytic percentage.

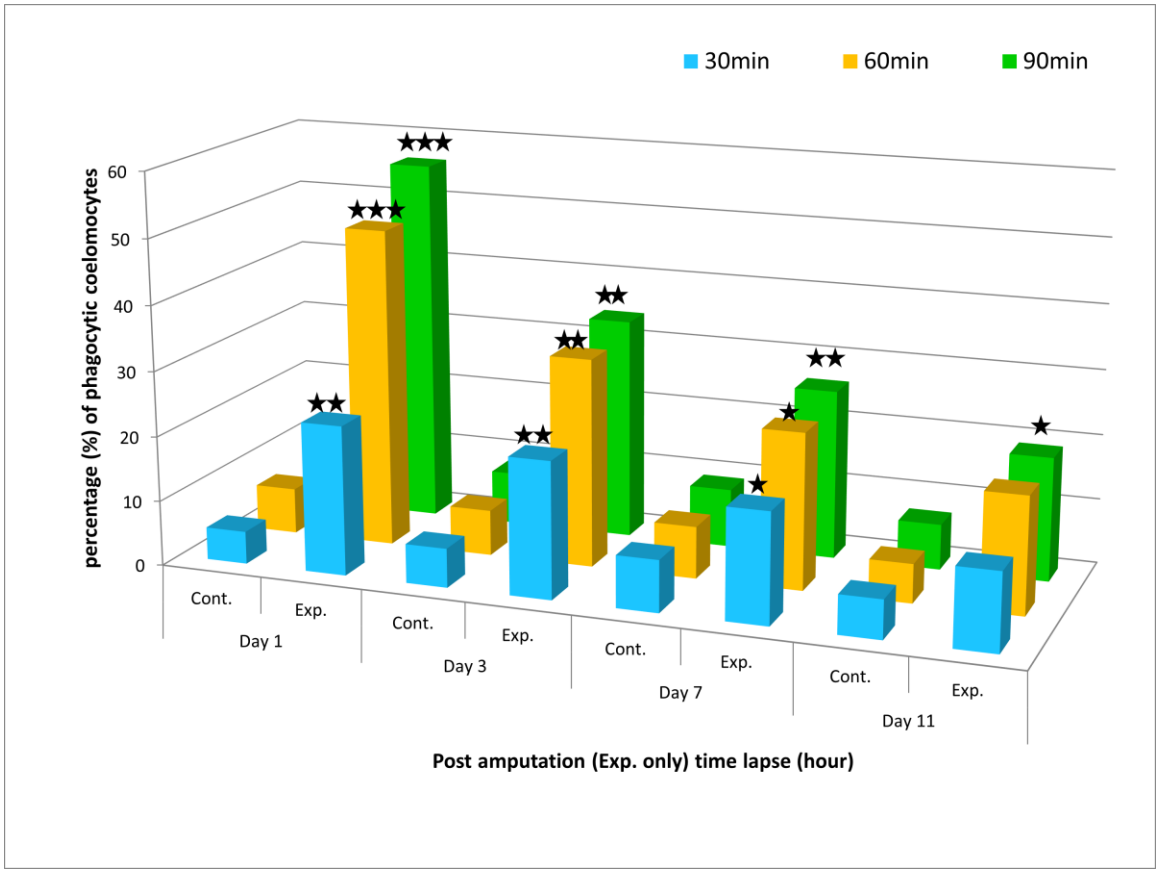


Fig.4.8: Bar graph showing percentage of coelomocytes with phagocytic inclusion bodies in the experimental (A4) and control (without previous amputation) *E. fetida* at different time intervals.

GENERAL DISCUSSION AND SUMMARY

Eisenia fetida is a rich source of novel microorganisms. Altogether 25 unique colonies (8 from the cast, 3 from the posterior gut and 5 from the anterior gut, 5 from coelomic fluid and 4from the whole gut in anaerobic condition) were isolated. Phylogenetic trees constructed using MEGA ver 7.0 for all the 20 gut-strains depict their unique phylogenetic positions in three major phyla- *Proteobacteria*, *Firmicutes* and *Actinobacteria* (Fig. 1.13). Whole genome sequencing has been carried out for four unique strains. These are ET03^T isolated from the cast (Saha *et al.,* 2018) ; EPG1 from the posterior gut; EAG2 and EAG3 from anterior gut of *E. fetida*. Earthworms, the so-called 'ecosystem engineers', play a key role in nutrient cycling by interacting with microorganisms particularly responsible to the turnover of organic matter in soil systems (Cao *et al.,* 2015). As detritivores, earthworms ingest and digest a mixture of dead organic matter and microorganisms, like animal manures. Manure type though does not significantly influence the taxonomic and phylogenetic composition of the cast, the earthworm produces. Manures strongly differed in their taxonomic and phylogenetic composition, but these differences were markedly reduced by the earthworm gut. The core earthworm cast microbiome comprised of the phyla *Proteobacteria*, *Actinobacteria*, *Firmicutes* and *Bacteroidetes* as found in our metagenomic data. Our results suggest that earthworms build up their cast microbiome by selecting from the pool of ingested bacteria.

The biological process of the nitrogen cycle is a complex interplay among many microorganisms catalyzing different reactions, where nitrogen is found in various oxidation states ranging from +5 in nitrate to -3 in ammonia. Reduction pathways are assimilatory nitrate reduction (MD:M00531) and dissimilatory nitrate reduction (MD:M00530) both for conversion to ammonia, and denitrification (MD:M00529). In denitrification nitrate or nitrite is reduced to gaseous nitrogen compounds (N_2, NO and N_2O) as a terminal electron acceptor at low oxygen or anoxic conditions and liberated to the atmosphere. Genes of nitrogen metabolism with significant hit were nitrite reductase (NO-forming), nitrate reductase catalytic subunit, *cyn*T, can; carbonic anhydrase , *nap*A; periplasmic nitrate reductase NapA , nitrate reductase (NADH) , *nir*A; ferredoxin-nitrite reductase , *nir*B; nitrite reductase (NAD(P)H) large subunit , *nor*B; nitric oxide reductase

subunit B , *nor*F; nitric-oxide reductase NorF protein and *nos*Z; nitrous-oxide reductase. Their effects on the nitrogen cycle are highlighted (Fig. DS-1).

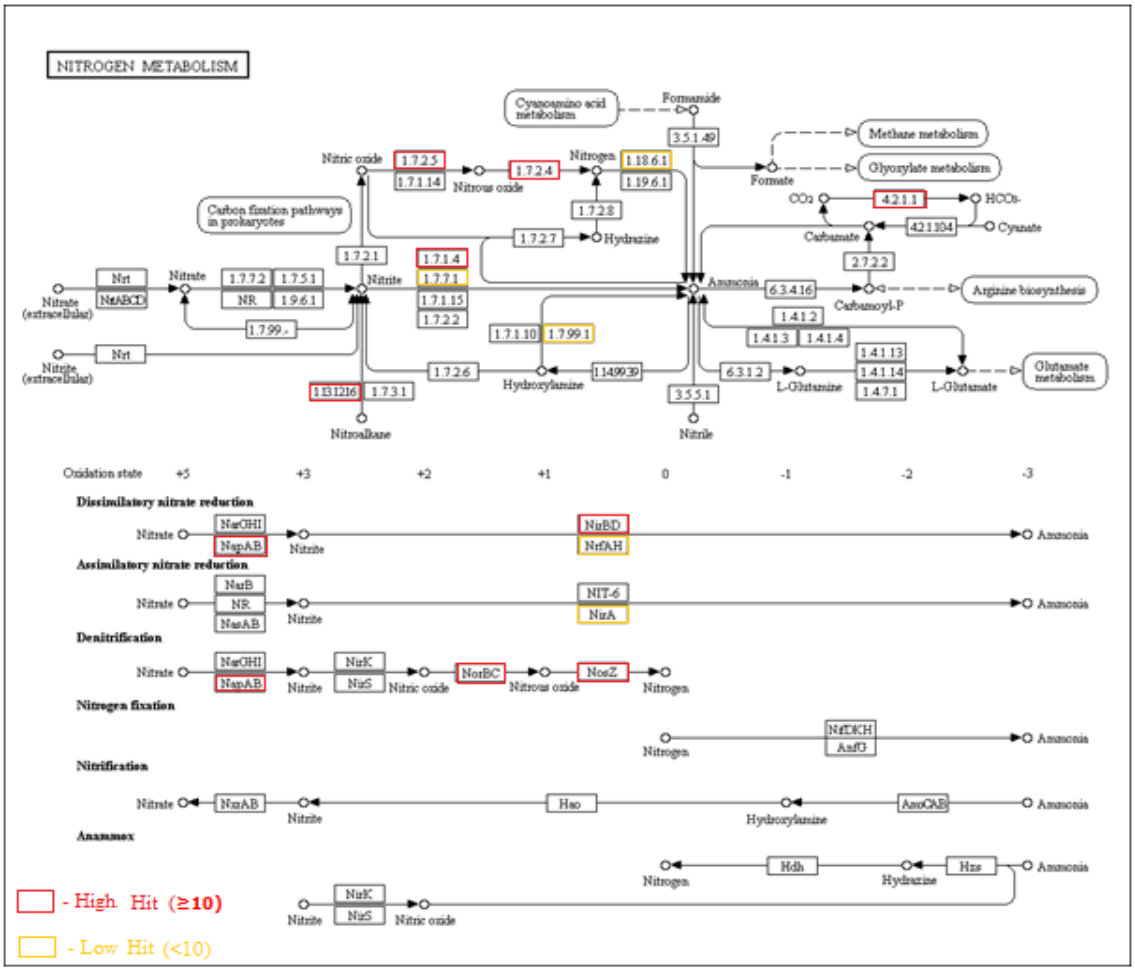


Fig. DS-1: Contribution of gut microbes of *E. fetida* on Nitrogen metabolism pathway/ Nitrogen cycle (The KEGG pathway was adopted from Kanehisa Laboratories, Japan).

Sulfur is an essential element for life and the metabolism. Sulfur occurs in various oxidation states ranging from +6 in sulfate to -2 in sulfide (H_2S). Sulfate reduction can occur in both an energy consuming assimilatory pathway and an energy producing dissimilatory pathway. In the assimilatory pathway, found in a wide range of organisms, sulfur compounds are reduced to compounds which ultimately are used for the biosynthesis of S-containing amino acids. In the dissimilatory pathway, which is restricted to obligatory anaerobic bacterial and archaeal lineages, sulfate (or sulfur) is the terminal electron acceptor of the respiratory chain producing large quantities of inorganic sulfide. The SOX (sulfur-oxidation) system (MD:M00595) is found in both

photosynthetic and non-photosynthetic sulfur-oxidizing bacteria. Genes of sulfur metabolism with significant hit were sulfate adenylyltransferase (*cys*D), adenylylsulfate kinase (*cys*C), phosphoadenosine phosphosulfate reductase (*cys*H), sulfite reductase (NADPH) hemoprotein beta-component (*cys*I), sulfite reductase (NADPH) flavoprotein alpha-component (*cys*J), sulfate adenylyltransferase subunit 1 (*cys*N) and bifunctional enzyme CysN/CysC (*cys*NC). Some other genes of sulfur metabolism with less hit (<10) were sulfite oxidase (SUOX), adenylylsulfate reductase, subunit A (*apr*A), adenylylsulfate kinase (*cys*C) and sulfite reductase (ferredoxin) (Fig. DS-2)

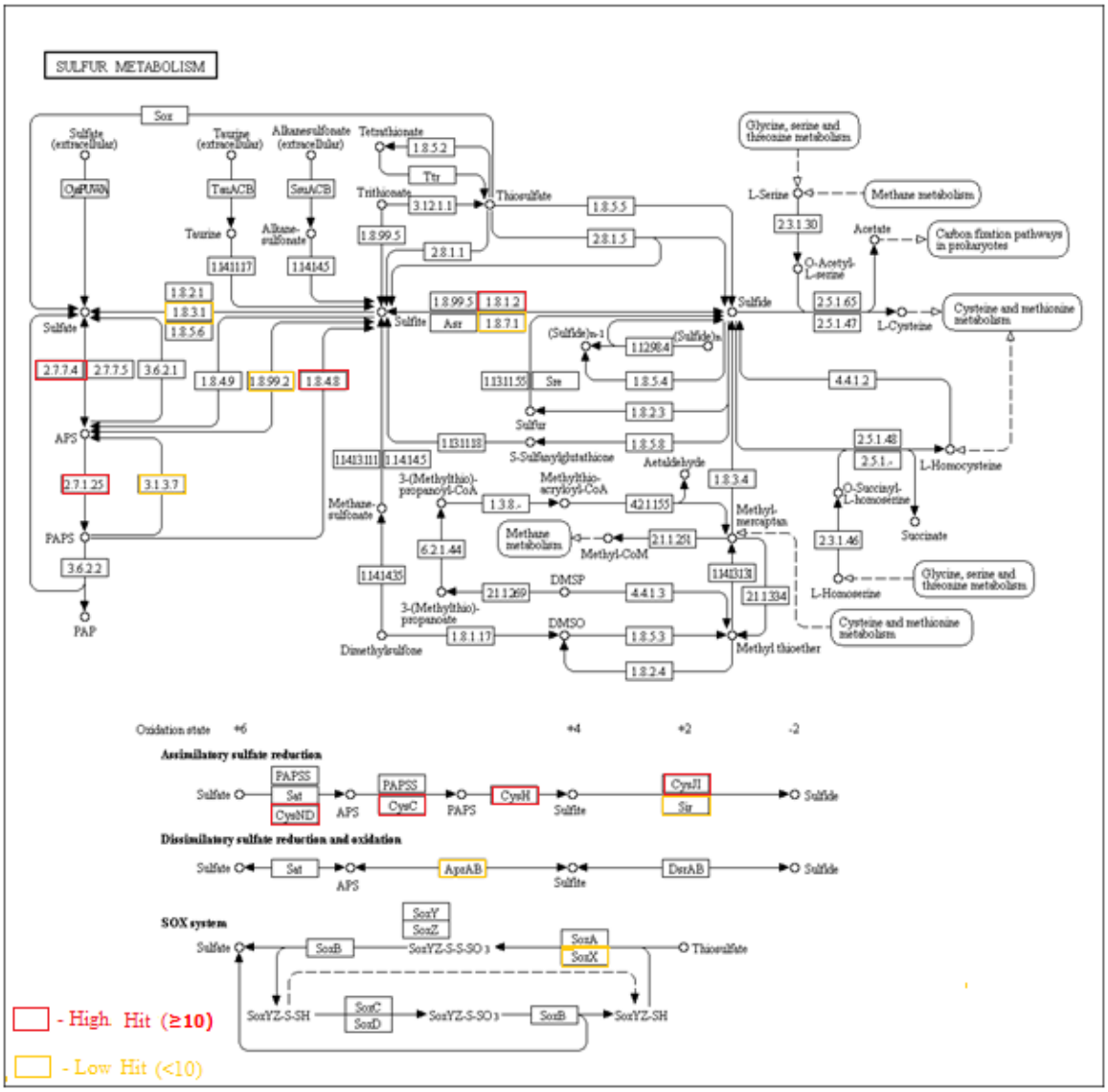


Fig. DS-2: Participation of gut microbes of *E. fetida* in sulfur metabolic pathway (The KEGG pathway was adopted from Kanehisa Laboratories, Japan).

Methane is metabolized principally by methanotrophs and methanogens. Methanotrophs consume methane as the only source of carbon, while methanogens produce methane as a metabolic byproduct. Methylotrophs can obtain energy for growth by oxidizing one-carbon compounds, such as methanol and methane. Genes of methane metabolism with significant hit were tetrahydromethanopterin S-methyltransferase subunit A (*mtr*A),

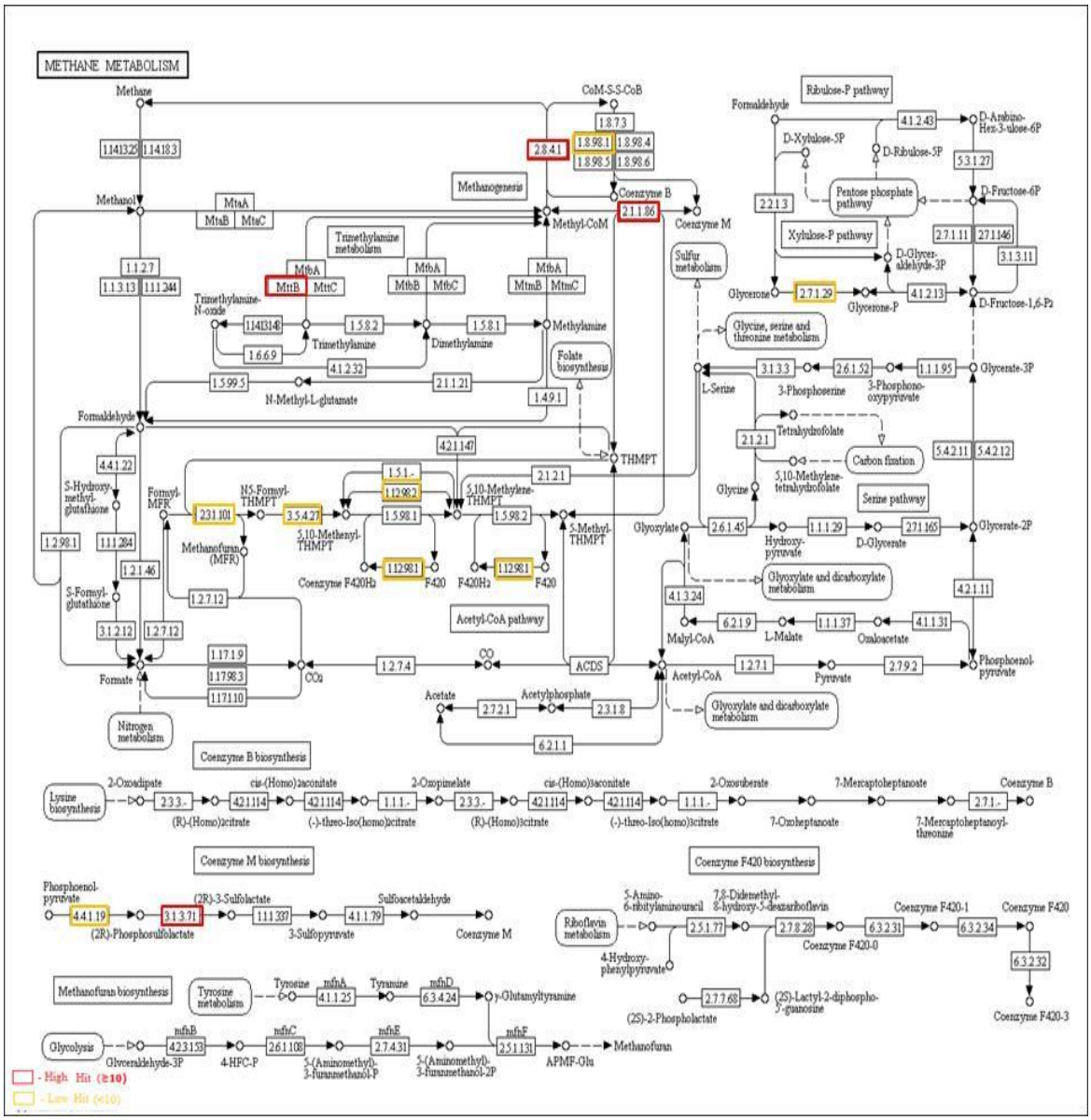


Fig. DS-3: Influence of gut microbes of *E. fetida* on methane metabolic pathway (The KEGG pathway was adopted from Kanehisa Laboratories, Japan).

2-phosphosulfolactate phosphatase (*com*B), methyl-coenzyme M reductase alpha subunit beta subunit (*mcr*A and *mcr*B) and tetrahydromethanopterin S-methyltransferase subunit

C, D, E and B (*mtr*C, *mtr*D, *mtr*E and *mtt*B). Some other genes of methane metabolism with less hit (<10) were formylmethanofuran-tetrahydromethanopterin N-formyltransferase (*ftr*), dihydroxyacetone kinase (DAK1, DAK2), phosphosulfolactate synthase (*com*A), coenzyme F420 hydrogenase beta subunit and gamma subunit (*frh*B, *frh*G), heterodisulfide reductase subunit A (*hdr*A), 5,10 methenyl-tetrahydro-methanopterin hydrogenase (*hmd*) and methenyltetrahydromethanopterin cyclohydrolase (*mch*)(Fig. DS-3).

Earthworms are constantly exposed to pathogens due to their detritivorous mode of life. Consequently, they have evolved various immuno-defense mechanisms which are assigned to coelomocytes, localized in the coelomic cavity. Light microscopic studies of coelomocytes in the *Eisenia fetida* coelom revealed the following morphologically distinct groups - amoebocytes (23±9%), granulocytes(18±7%) eleocytes(6±3%). Other than that small spherical cells (51±8%) has also been observed. Amoebocytes with phagocytotic activity can be identified from the size and shape of the nucleus. The cells tend to be smaller than, as little as 8 µm, but occasionally may be as large as 15 µm with fewer granules. Granulocytes have cytoplasm completely filled with small granules and have diameter of 15-22 µm. The eleocytes are large cells with 30-60 µm diameter. These cells possess a natural fluorescence like activity. The fourth group of cells in the coelomic fluid may be the transitory cells which are progenitors of the distinct cell lineages at different stages of maturation. We have differentially interrogated the growth patterns of *B. megaterium* and *B. thuringiensis* in Luria Broth and formulated broth that mimics the coelomic fluid of *Eisenia* in composition. It has been observed that both the species have similar growth kinetics in the Luria broth, but *B. megaterium* has distinctly faster growth rates in the coelom mimicking broth. Thus, the symbiosis like co-relation of *B. megaterium* and *Eisenia fetida* is evident. Further, closely related bacteria can secrete a wide array of antibacterial compounds such as bacteriocins when competing with other bacteria for the same resources. This may be the sources of future research to elucidate this bacterial competition. The bacterial growth in the coelomic fluid is controlled by the phagocytic and antimicrobial activity of coelomocytes. Expression of defence molecules like coelomic cytolytic factor (CCF) in the gut epithelial tissue is high due to continuous flow of immune response to an incessant flow of microbes with

ingested food. The coelomocyte population in response to forced introduction of a pathogen (*B. thuringiensis*) in the test and control *E. fetida* was compared. When there is a possibility of huge upsurge due to mobilization of pathogenic bacteria, pathogens in mass are captured or in other words encapsulated by multicellular entities produced by amoebocytes and eleocytes. The plausible immunomodulatory function of riboflavin (stored in chloragosomes) was demonstrated by establishing it as chemoattractant for coelomocyte-taxis. Riboflavin (vitamin B2) synthesis efficiency of the coelomic bacterial isolates was measured in a time-dependent manner during growth in a medium formulated on the basis of composition of the coelomic fluid. Chemotaxis of the environment bacteria towards coelomic fluid was discovered. Chemotaxis was studied by an improvised technique where bacterial mobilization to the CF contained in a capillary tube was quantified in a time-dependent manner. The notable findings of the present study are: (i) *Bacillus* population in the CF is up to the magnitude of 10^6 cells/ml (~68% of total bacterial load). (ii) The three major species recognized are- *Bacillus megaterium* (representative strain Ah4), *B. cereus* (representative strain BCR) and *B. pumilus* (representative strain BP). (iii) The number of coelomocytes particularly chloragocytes (or eleocytes) was found to increase from 12h of pathogen challenge. Since *E. fetida* is unable to produce riboflavin but compulsorily required for immune functions, chloragocytes store considerable amount of riboflavin in their chloragosomes (possibly derived from the symbionts). Strain Ah4, isolated from CF, has shown the fastest chemotactic mobility towards CF and at the same time its riboflavin production rate was highest among the *Bacillus* species that survive in the coelome. The basis of host-bacterial symbiosis (give-n-take phenomenon) was thus revealed.

Regeneration or re-growth of lost body parts after amputation,is restricted and sparsely found in few discrete groups of kingdom Animalia. *Eisenia fetida*, an annelid, besides being well-known for its usage in composting, has drawn special interest to the biologists because of its regenerative property. The present study is undertaken to describe tissue reorganization after amputation in *E. fetida.* Transverse amputation of adult *Eisenia fetida* at different regions of body followed by survival and development studies revealed that anterior fragments can regenerate missing posterior tail regions when amputations are done at least beyond the clitella. Tail regeneration from the 66th

metameric ring occurs with highest survival rate (89%). Normal nutritional activity and defecation restart from 9th day post amputation (dpa). Internal tissue reorganization and formation of the major tissues are complete on 11th dpa. The amputated segment's growth in respect to body weight gain was noted from 42 dpa. Histological studies reveal that neoblast cells originate not only from dedifferentiation of the longitudinal muscle cells, but also by multiplication of basal epithelial cells and visceral peritoneal cells. The blastemal mass comprising of chloragogue tissue has also been observed which might serve as reservoir for nutrition during regeneration. This histological study attempts to reveal the origin of blastemal cells during earthworm's wound healing and posterior regeneration. From 1 day post-amputation, neoblast cells could be observed to riginate from the de-differentiating longitudinal muscle layer of the body wall facing the coelomic side (Fig. 4.5 Ic). Similar observations were also reported by Park *et al.* 2013. These de-differentiating cells after proliferation and re-differentiation can rebuild the longitudinal muscle layer (Fig. 4.5 Vc). There was no existing report on de-differentiating cells from tissues other than longitudinal muscle cells. Here, it has been reported for the first time that dedifferentiating cells have also originated from- (i) the chloragogue tissue (Fig. 4.5 Ic), and (ii) intestinal tissue especially the basal cells of the typhlosole region (Fig. 4.5 IIe). The dedifferentiated chloragouge cells formed the blastema mass on the visceral side of the coelom which was not observed in the de-differentiated longitudinal muscle cells. From 5dpa onwards, the neoblast chloragogue cells could be found to accumulate in the coelomic space and by 11dpa it occupied most of it (Fig. 4.5 VIIa). This blastema might act as a nutrition-reservoir for the regenerating tissue. Most of the segmental structures viz. epithelium, circular and longitudinal muscle layers, wall of intestine were reformed by 11dpa. The release of cast through the regenerated pygidium by A4 amputee at 9dpa was also observed. Muscle-cells were previously thought to regenerate from cells embedded in the muscle layer itself as precursor cells (Bely, 1999). In contrst, we have observed during this histological study that, longitudinal muscle-cells loose-offfrom each other possibly by losing intra cellular adhesion molecules and divide resulting into the separation of nucleated part and contractile part, thus forming the de-differentiated neoblast cells. Cell-tracking studies by Tweeten & Reiner, 2012 conducted on earthworm *Lumbriculus,* also confirm that old gut tissue contribute to regenerate new gut tissue. In

Planarian regeneration, stem cells distributed in different regions of the body help to regenerate lost tissue structures (Tanaka & Reddien, 2011). Whether or not the earthworms follow the same rule has long been a debated issue. Blastema which is reported to be a homogeneous mass of neoblast (undifferentiated) cells is formed during regeneration of the fish fin and amphibian limb. Body part or limb regeneration experiments which were conducted in some vertebrates support lineage specific origin of blastema cells with unipotent functional attributes. Actually, the blastema in such cases was argued to be a heterogeneous mass of lineage-restricted cells (King and Newmark, 2012). The present findings suggest that during regeneration in *E. fetida* the origin of neoblast cells are tissue restricted. The neoblast cells originated through proliferation of the de-differentiated muscle cells and the chloragogue cells in respective tissue layers. These neoblast cells re-differentiated into tissues in the regenerating region of the amputated region of the earthworm (Fig. 4.5 IIIb, Vb, VIc). There could be cellular regulatory mechanisms which control the process of regeneration. Deep investigations at molecular level on model organisms may identify key molecules required to initiate and modulate the regenerative events. Some reports suggest that nerve cord may have cardinal role during annelid regeneration (Bely, 2014). Absence of the nerve cord at the amputed region of *Eisenia* inhibits ectodermal and mesodermal growth, but, endodermal growth remains undisturbed. The deep cellular events of de-differentiation are still not well understood. So, it is important to characterize how the extracellular environment is altered during regeneration, and how different extracellular factors may shape cellular activity in the blastema. The present histological guide is a preliminary advancement which offers an opportunity to the molecular biologists to use the anatomical events following amputation to link gene function during regeneration.

The molecular level characterization vows to our hypothesis that the gut of *E. fetida* houses novel bacteria and these unique bacteria might have an important biotechnological application in the field of complex sugar digestion or as the source of novel antibiotic or antifungal agents. Study of the *Eisenia fetida* coelomocytes' gene expression under bacterial challenge at genome-wide-transcriptomic level also has produced significant data having implications in both academic and potential biotechnological applications.

BIBLIOGRAPHY

Achal V, Pan X. Characterization of urease and carbonic anhydrase producing bacteria and their role in calcite precipitation. *Curr. Microbiol.* 2011. 62: 894-902

Adler J. A Method for Measuring Chemotaxis and Use of the Method to Determine Optimum Conditions for Chemotaxis by Escherichia coli. *J Gen Microbiol.*1973; 74:77-91.

Aira M, Monroy F, & Domı´nguez J. Microbial biomass governs enzyme activity decay during aging of worm-worked substrates through vermicomposting. *J Environ Qual* 2007; 36:448–452.

Aira M, Monroy F, Domínguez J (2009) Changes in bacterial numbers and microbial activity of pig manure during gut transit of epigeic and anecic earthworms. *J Hazard Mat* 162: 1404–1407.

Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215:403–410.

Altschul SF, Madden TL, Scha¨ffer AA, Zhang J, Zhang Z *et al*. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25:3389–3402.

Anderson RC, Majak W, Rassmussen MA, Callaway TR, Beier RC, *et. al.* Toxicity and metabolism of the conjugates of 3- Nitropropanol and 3-Nitropropionic acid in forages poisonous to livestock. *J Agric Food Chem* 2005;53:2344-2350.

Andre F. 1963. Contribution a l'analyse experimental de la reproduction des lombriciens. *Bull Biol Fr Belg* 97:1–101.

Bartzatt R. & Follis ML. Detection and Assay of Riboflavin 2 (Vitamin B2) Utilizing UV/VIS Spectrophotometer and Citric Acid Buffer. *J. Sci. Res. Rep*. 2014; 3:799-809.

Bassler BL, Gibbons, PJ & Roseman S. Chemotaxis to chitin oligosaccharides by *Vibrio furnissii*, a chitinivorous marine bacterium. *Biochem. Biophys. Res. Com.* 1989; 161:1172–1176.

Bely AE & Nyberg, KG. Evolution of animal regeneration: re-emergence of a field. *Trends Ecol. Evol.* 2010; 25:161-170.

Bely AE, **Sikes JM.** Latent regeneration abilities persist following recent evolutionary loss in asexual annelids. *Proc. Natl. Acad. Sci.* USA 107: 1464-1469, 2010.

Bely AE, Zattara1 EE & Sikes JM. Regeneration in spiralians: evolutionary patterns and developmental processes. *Int. J. Dev. Biol.* 2014; 58:623-634.

Bely AE. Decoupling of fission and regenerative capabilities in an asexual oligochaete. *Hydrobiologia.* 1999; 406:243-251.

Bely AE. Distribution of segment regeneration ability in the Annelida. *Integr. & Comp. Biol*. 46: 508- 518, 2006.

Bely AE. Early events in annelid regeneration: a cellular perspective. *Integr. Comp. Biol.* 2014; 54:688-699.

Bentley DR, Balasubramanian S, Swerdlow HP, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nat.* 2008;456(7218):53–59.

Berrill NJ. Regeneration and budding in worms. *Biol. Rev.* 1952; 27:401-438.

Bilej M, Brys L, Beschin A, Lucas R, Vercauteren E, Hanusova R, De Baetselier P. Identification of a cytolytic protein in the coelomic fluid of Eisenia foetida earthworms. *Immunol Lett* 1995;45(1–2):123–8.

Bilej M, De Baetselier P, & Beschin A. Antimicrobial defense of the earthworm. *Folia Microbiol* 2000; 45(4):283–300.

Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 1959; 37: 911–917.

Bode H, Berking S, David CN, Gierer A, Schaller H & Trenkner E. Quantitative analysis of cell types during growth and morphogenesis in Hydra. *Roux's Arch.* 1973; 171:269-285.

Bonnert TP, Garka KE, Parnet P, Sonoda G, Testa JR, Sims JE. The cloning and characterization of human MyD88: a member of an IL-1 receptor related family, *FEBS Lett*. 1997; 402(1):81–84.

Bouché, M.B. 1972. *Lombriciens de France: écologie et systématique*. Paris: Institut National des Recherche Agronomique.

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976; 72:248–54.

Brown AE. Benson's Microbiological Applications Lab Manual, The McGraw−Hill. 2001.

Brusca RC & Brusca GJ. Sinauer Associates. Sunderland. 2nd Edition. Invertebrates. 2003.

Bundya JG, Osbornb D, Weeksb JM, Lindona JC & Nicholsona JK. An NMR-based metabonomic approach to the investigation of coelomic fluid biochemistry in earthworms under toxic stress. *FEBS Lett*. 2001; **500:** 31-35.

Burke Janice M. Wound healing in *Eisenia fetida* (Oligochaeta). *Cell and Tissue Res.* 1974; 154 (1): 61-82.

Chakraborty R & Roy P. Chemotaxis of chemolithotrophic *Thiobacillus ferrooxidans* towards thiosulfate. *FEMS Microbiol Lett*. 1992; **98**:9–12.

Chauvel A, Grimaldi M, Barros E, Blanchart E, Sarrazin M & Lavelle P. Pasture degradation by an Amazonian earthworm. *Nat*. 1999; 389:32-33.

Chun J, Oren A, Ventosa A, Christensen H, Arahal DR. *et al.* Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol*. 2018; 68:461-466.

Cikutovic MA, Fitzpatrck LC, Goven AJ, Venables BJ, Giggleman MA, Cooper EL. Wound healing in earthworm *Lumbricus terrestris*: A cellular-based biomarker for assessing sublethal chemical toxicity. *Bull. Environ. Contam. Toxicol.* 1999; 62: 508-514.

Cikutovic MA, Fitzpatrck LC, Venables BJ, Goven AJ. Sperm count in earthworm (*Lumbricus terrestris*) as a biomarker for enviromental toxicology: Effect of cadmium and chlordane. *Environ. Pollut*. 1993; 81: 123-125.

Clark ME. Later stages of regeneration in the Polychaete, Nephtys. In: *Regeneration in Lower Vertebrates and Invertebrates*. 1972; 3: 75-102.

Claus D. A standardized Gram staining procedure. *World J Microb Biot.* 1992; 8:451–452.

Cooper EL, Roch P. Second-set allograft response in the earthworm Lumbricus terrestris: Kinetics and characteristics. *Transplant.* 41: 514- 520, 1986.

Cooper EL, Roch P. The Capacities of Earthworms to Heal Wounds and to Destroy Allografts are Modified by Polychlorinated Biphenyls (PCB). *J Invert Pathol.*1992; 60:59-63.

Cooper EL. Earthworm coelomocyte immunity. In: Contemporary Topics in Immunobiology Invertebrate Immunology Plenum Press. 1974; 4:91-107.

Cornec JP, Cresp J, Delye P, Hoarau F, Reynaud G. Tissue responses and organogenesis during regeneration in the oliogochaete Limnodrilus hoffmeisteri (Clap.). *Can J Zool* 1987; 65:403-414.

Cotuk A, Dales RP. Lysozyme activity in the coelomic fluid and coelomocytes of the earthworm Eisenia foetida SAV in relation to bacterial infection. *Comp Biochem Physiol* 1984; 78A:469–74.

Cui, D. Atlas of Histology with Functional and Clinical Correlations. Lippincott Williams & Wilkins, Baltimore. 2011.

Dales RP, Kalac Y. Phagocytic defence by the earthworm Eisenia foetida against certain pathogenic bacteria. *Comp Biochem Physiol.*1992; 101A:487–490.

Dhainaut A & Scaps P. Immune Defense and Biological Responses induced by toxics in Annelida. *Can J Zool*. 2001; 79:233-253.

Dvořák J, Roubalová R, Procházková P, Rossmann P, Škanta F & Bilej M. Sensing microorganisms in the gut triggers the immune response in *Eisenia andrei* earthworms. *Dev Comp Immunol*. 2016; 57:67-74.

Edwards CA & Bohlen, PJ. Biology and ecology of earthworms. London: Chapman and Hall. 1996.

Edwards CA & Lofty JR. *Biology of Earthworms*. Chapman and Hall, London. 1977:129-147.

El-Naggar NE-A, El-Ewasy SM. Bioproduction, characterization, anticancer and antioxidant activities of extracellular melanin pigment produced by newly isolated microbial cell factories *Streptomyces glaucescens* NEAE-H. *Sci Rep*. 2017;7:42129. doi: 10.1038/srep42129.

Engelmann P, Cooper EL, Opper B, Németh P. Earthworm innate immune system. In: Biology of Earthworms. Ed. Karaca A. Springer Verlag, Heidelberg: 2011; 229–245.

Engelmann P, Opper B & Németh P. Interactions of intracellular calcium and immune response in earthworms. *ISJ* 2011; 8:78-84.

Eyambe GS, Goven AJ, Fitzpatrick LC, Venables BJ, Cooper EL. A non-invasive technique for sequentia1 collection of earthworm *(Lumbricus rerresfris)* leukocytes during subchronic immunotoxicity studies. *Lab Anim.* 1991; 25: 61-67.

Fan H, Ives AR, Surget-Groba Y, Cannon CH. An Assembly and Alignment-Free Method of Phylogeny Reconstruction from next-Generation Sequencing Data. *BMC Genomics*. 2015; 16(1):522.

Faria VG, Martins NE, Magalhães S, Paulo T F, Nolte V, Schlötterer C, Sucena É & Teixeira L. *Drosophila* Adaptation to Viral Infection through Defensive Symbiont Evolution. *PLoS Genet*. 2016; 12:e1006297.

Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 1981; 17:368-376.

Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution.* 1985; 39(4):783-791

Fitch WM. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool* 1971; 20:406–416.

Fitzharris TP & Lesh GE. Gut and nerve-cord interaction in sabellid regeneration. In: *Regeneration in Lower Vertebrates and Invertebrates Vol 3* ed. V.B. Eichler, MSS Info. Corp. New York. 1972; 60-74.

Fitzpatrick LC, Goven AJ, Venables BJ & Cooper EL (1990) Earthworm immunoassays for evaluating biological effects of exposure to hazardous materials, in In situ Evaluation of Biological Hazards of Environmental Pollutants, Sandhu SS (Ed.), Plenum Press, NY. 1990; 119-129.

Francesca S, Agniswamy J, Yuan H, Vercammen K, Pelicaen R, *et. al.* The combined structural and kinetic characterization of a bacterial nitronate monooxygenase from *Pseudomonas aeruginosa* PAO1 establishes NMO Class I and II. *J Biol Chem* 2014;289: 23764-23775.

Francis K, Smitherman C, Nishino SF, Spain JC, Gadda G. The biochemistry of the metabolic poison propionate 3-nitronate and its conjugate acid, 3-nitropropionate. *IUBMB Life* 2013;65:759-768.

Fraune S & Bosch TCG. Why bacteria matter in animal development and evolution. *Bioessays.* 2010; 32:571–580.

Furlong MA, Singleton DR, Coleman DC, Whitman WB. Molecular and culture-based analyses of prokaryotic communities from an agricultural soil and the burrows and casts of the earthworm *Lumbricus rubellus*. *Appl Environ Microbiol.* 2001; 68:1265–1279.

Garcia-Angulo VA. Overlapping riboflavin supply pathways in bacteria. *Crit. Rev. Microbiol.* 2017; 43:196-209.

Gilbert JA, O'Dor R, King N, Vogel TM. The importance of metagenomic surveys to microbial ecology: or why Darwin would have been a metagenomic scientist *Microb. Inform. Exp.* 2011; (1).5.

Go´mez-Brando´n M, Aira M, Lores M, Domı´nguez J. Epigeic Earthworms Exert a Bottleneck Effect on Microbial Communities through Gut Associated Processes. *PLoS ONE* 2011; 6(9):e24786.

Gooneratne R and Muangphra P. Toxicity of Commercial Neem Extract to Earthworms (Pheretima peguana). *Applied and Environmental Soil Science***.** *Hindawi Publishing Corporation.*2011.

Gordon RE, Barnett DA, Handerhan JE, Pang, CH. Nocardia coeliaca, Nocardia autotrophica, and the nocardin strain. *Int J Syst Bacteriol*. 1974; 24:54–63.

Goven AI, Fitzpatrick LC, Venables BJ, Cooper EL. An invertebrate model for analyzing effects of environmental xenobiotics on immunity. *Clinical Ecology.* 1987; 4:150-4.

Govindarajan B & Prabhaharan V. Gut microflora of earthworms: A review. *Am. J. Biol. Pharm. Res.* 2014; 1:125-30.

Grimaldi A, Banfi S, Bianchi C, Greco G, Tettamanti G, Noonan DM. The leech: a novel invertebrate model for studying muscle regeneration and dseases. *Curr. Pharm. Des.* 2010; 16: 968-977.

Gupt KK, Rai Aneja K, & Rana D. Current status of cow dung as a bioresource for sustainable development. *Bioresour Bioprocess* 2016; 3:28.

Gupta RS, Lo B and Son J. Phylogenomics and Comparative Genomic Studies Robustly Support Division of the Genus Mycobacterium into an Emended Genus Mycobacterium and Four Novel Genera. Front. *Microbiol.* 2018; 9:67.

Halder KR. Annelida: Oligochaeta: Earthworm. in The Director, Zoological Survey of India (Eds) Fauna of West Bengal, Part-10, ZSI, Calcutta. 1999; 17-93.

Hamed SS, Kauschke E, Cooper EL. Cytochemical properties of earthworm coelomocytes enriched by percoll. *Int. J. Zoo. Res.* 2005; 1(1):74-83.

Heimpel AM & Angus TA. The taxonomy of insect pathogens related to *Bacillus cereus* Frankland and Frankland. C an. *J. Microbial.* 1958; 4: 531-541.

Heller K. Tsetse Flies Rely on Symbiotic Wigglesworthia for Immune System Development. *PLoS Biol.* 2011; 9:e1001070.

Hemberger S, Pedrolli, DB, Stolz J, Vogl C, Lehmann M & Mack M. RibM from *Streptomyces davawensis* is a riboflavin/roseoflavin transporter and may be useful for the optimization of riboflavin production strains. *BMC Biotechnol* 2011; 11:119.

Hill SD. Origin of the regeneration blastema in polychaete annelids. *Am. Zool.* 1970; 10: 101-112.

Homa J, Klimek M, Kruk J, Cocquerelle C, Vandenbulcke F, Plytycz B. 2010. Metal-pecific effects on metallothionein gene induction and riboflavin content in coelomocytes of Allolobophora chlorotica. *Ecotoxicol Environ Saf* 73:1937–1943.

Hyman LH (1940) Aspects of regeneration in annelids. *Am Nat* 74:513–527.

Ihssen, J., M. A. Horn, C. Matthies, A. Go¨ßner, A. Schramm, H. L.Drake, 2003. N_2O-producing microorganisms in the gut of the earthworm *Aporrectodea caliginosa* are indicative of ingested soil bacteria. *Appl. Environ.Microbiol.*, 69: 1655–1661.

Jaehme M & Slotboom DJ. The diversity of membrane transport proteins for vitamins in bacteria and archaea. *Biochim Biophys Acta.* 2015; 1850:565–76.

Joskova R, Silerova M, Prochazkova P & Bilej M. Identification and cloning of an invertebratetype lysozyme from *Eisenia andrei. Dev. Comp. Immunol.* 2009; 33:932-938.

Kale RD, Mallesh BC, Kubra B, Bhagyaraj DJ. (1992) Influence of vermicompost application on available micronutrients and selected microbial populations in paddy field. Soil *Biol and Biochem* 1992; 24: 1317-1320.

Kale RD, Sunitha NS. Efficiency of Earthworms (*E. Eugeniae*) in Converting the Solid Waste from Aromatic Oil Extraction Industry into Vermicompost. *Journal of IAEM.* 1995; 22: 267-269.

Kale RD. Earthworms: Nature's Gift for Utilization of Organic Wastes. St. Lucie Press, NewYork. 1998.

Kämpfer P. Order XIV. Streptomycetales ord. nov. In Goodfellow M. *et al.* (editors) *Bergey's Manual of Systemic Bacteriology 2nd ed Vol V The Actinobacteria Part A.* New York. Springer; 2012 Pp1446-1804. DOI 10.1007/978-0-387-68233-4.

Kaur I, Das AP, Acharya M, Klenk HP, Sree A *et al.*, *Planococcus plakortidis* sp. nov., isolated from the marine sponge *Plakortis simplex* (Schulze). *Int J Syst Evol Microbiol* 2012; 62:883–889.

Kauschke E and Mohrig W. (1987) Comparative analysis of hemolytic and hemagglutinating activities in the coelomic fluid of Eisenia foetida and Lumbricus terrestris (Annelida~ Lumbricidae). *Dev Com Immunol.* 1987; 2: 331-341.

Keng LB. Phil. Trans. Roy. Soc. London Ser. B 1895; 186:383-400.

Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014; 64:346–351.

King, R.S. & Newmark, P.A. The cell biology of regeneration. *J. cell Biol.*2012; 196 (5): 553.

Kiyasudeen SK, Ibrahim MH, Quaik S & Ismail SA. Microbial Ecology Associated with Earthworm and Its Gut. In Prospects of Organic Waste Management and the Significance of Earthworms Springer International Publishing, AG 2017; 123-145

Knapp BA, Podmirseg SM, Seeber J, Meyer E, Insam H (2009) Diet-related composition of the gut microbiota of *Lumbricus rubellus* as revealed by a molecular fingerprinting technique and cloning. Soil Biol Biochem 41: 2299–2307.

Ko¨hlerova´ P, Beschinb A, Silerova M, De Baetselierb P, Bilej M. Effect of experimental microbial challenge on the expression of defense molecules in *Eisenia foetida* earthworm. *Dev Com Immunol*. 2004; 28: 701–711.

Kondo M, Akasaka K. Regeneration in crinoids. *Dev. Growth Differ*. 2010; 52:57-68.

Kramer CY. Extension of Multiple Range Tests to Group Means with Unequal Numbers of Replications *Biometrics* 1956; 12, 307-310.

Kumar A, Mukherjee S, Chakraborty R. Characterization of a novel trimethoprim resistance gene, dfrA28, in class 1 integron of an oligotrophic *Acinetobacter johnsonii* strain, MB52, isolated from River Mahananda, India. *Microb Drug Resist* 2010; 16:29–37.

Kumar S, Stecher G, Tamura K, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets *Mol Biol Evol* 2016; 33*:*1870–1874.

Kurek A, Plytycz B. Annual changes in coelomocytes of four earthworm species. Pedobiologia 2003;47:689–701.

Labeda DP, Dunlap CA, Rong X, Huang Y, Doroghazi JR, Ju K-S, Metcalf WW. Phylogenetic relationships in the family Streptomycetaceae using multi-locus sequence analysis *Antonie van Leeuwenhoek* 2017; 110:563–583. DOI 10.1007/s10482-016-0824-0.

Lacal J, Alfonso C, Liu X, Parales R E, Morel B, Conejero-Lara F, Rivas G, Duque E, Ramos J L & Krell T. Identification of a Chemoreceptor for Tricarboxylic Acid Cycle Intermediates Differential Chemotactic Response Towards Receptor Ligands. *J. Biol. Chem.* 2010; 285:23126–23136.

Lanyi B. Classical and rapid identification methods for medically important bacteria. *Methods Microbiol* 1987; 19:1–67.

Lassegues M, Roch P & Valembois P. Antibacterial activity of *Eisenia fetida andrei* coelomic fluid: evidence, induction and animal protection. *J Invert Pathol*. 1989; 53:1–61989 (1989).

Lebenko EV, Sekerina OA & Chemerilova VI. The Chemotactic Characteristics of the S and R Dissociants of *Bacillus thuringiensis*. *Microbiol*. 2005; 74:76-79.

LeBlanc JG, Milani C, de Giori GS, Sesma F, van Sinderen D & Ventura M. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Curr Opin Biotechnol*. 2013; 24:160–8.

Lee JY, YeopLee J, Jung HW, Hwang BK. *Streptomyces koyangensis* sp. nov., a novel actinomycete that produces 4-phenyl-3-butenoic acid. *Int. J. Syst. Evol. Microbiol*. 2005; 55: 257-262.

Lee SD. *Brevibacterium samyangense* sp. nov., an actinomycete isolated from a beach sediment. *Int J Syst Evol Microbiol* 2006;56:1889– 1892.

Liebmanm E. The coelomocytes of lumbricidae, *J. Morph*. 1942; 71(2):221-245.

Lund MB, Kjeldsen KU & Schramm A. The earthworm *Verminephrobacter* symbiosis: an emerging experimental system to study extracellular symbiosis. *Front Microbiol.* 2014; 5.

MacFaddin JF. Biochemical Tests for Identification of Medical Bacteria, 3rd Ed, Williams and Wilkins, Baltimore Md. 2000; 29.

Maciejewska M, Adam D, Naômé A, Martinet L, Tenconi E. *et al.* Assessment of the Potential Role of *Streptomyces* in Cave Moonmilk Formation. *Front. Microbiol.* 2017;8:1181. doi: 10.3389/fmicb.2017.01181.

Magee GM, and Ward AC. Genus I. Mycobacterium Lehmann and Neumann 1896; 363.

Mandel MJ, Schaefer AL, Brennan CA, Heath-Heckman EAC, DeLoney-Marino CR, McFall-Ngai MJ & Ruby EG. Squid-derived chitin oligosaccharides are a chemotactic signal during colonization by Vibrio fischeri. *Appl. Environ. Microbiol.* 2012; 78:4620–4626.

Manzano-Martin A & Latorre A. Settling down: the genome of *Serratia symbiotica* from the aphid *Cinara tujafilina* zooms in on the process of accommodation to a cooperative intracellular life. *Genome Biol Evol* 2014; 6:1683–98.

Marks DH, Cooper EL. *Aeromonas hydrophila* in the Coelomic Cavity of the Earthworms *Lumbricus terrestris* and *Eisenia foetida* .*J Invert Pathol*. 1977; 29:382-383.

Marmur JA. Procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J Mol Biol* 1961; 3:208–IN1.

Mayr, E. 1942. *Systematics and the origin of species*. New York: Columbia University Press.

Mazur AI, Klimek M, Morganb AJ & Plytycz B. Riboflavin storage in earthworm chloragocytes and chloragocyte-derived eleocytes and its putative role as chemoattractant for immunocompetent cells. *Pedobiologia* 2011; 54S:S37– S42.

Meier-Kolthoff JP, Klenk HP & Göker M. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol*. 2014; 64:352–356.

Minnikin DE, Odonnell AG, Goodfellow M, Alderson G, Athalye M *et al*., An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *J Microbiol Methods* 1984; 2:233–241.

Møller P, Lund MB & Schramm A. Evolution of the tripartite symbiosis between earthworms, *Verminephrobacter* and *Flexibacter*-like bacteria. *Front Microbiol.* 2015; 6:529 (2015).

Monroy F, Aira M, Domínguez J (2011) Epigeic earthworms increase soil arthropod populations during first steps of decomposition of organic matter. *Pedobiologia* 54: 93–99.

Müller MCM & Henning L. Ground plan of the polychaete brain: Patterns of nerve development during regeneration in *Dorvillea bermudensis* (Dorvilleidae). *J. Comp. Neurol.* 2004; 47149-58.

Müller MCM, Berenzen A & Westheide W. Experiments on anterior regeneration in *Eurythoe complanata* ("Polychaeta", Amphinomidae): reconfiguration of the nervous system and its function for regeneration. *Zoomorph.* 2003; 122: 95-103.

Müller MCM. Nerve development, growth and differentiation during regeneration in *Enchytraeus fragmentosus* and *Stylaria lacustris* (Oligochaeta). *Dev. Growth Differ*. 2004; 46:471-478.

Murray RGE, Doetsch RN, Robinow CF. Determinative and cytological light microscopy. In: Gerhardt P, Murray RGE, Wood WA, Krieg NR (editors). *Methods for General and Molecular Bacteriology* Washington, DC: ASM*;* 1994; 21–41.

Myohara M, Niva CC, Lee JM. Molecular approach to annelid regeneration: cDNA subtraction cloning reveals various novel genes that are upregulated during the large-scale regeneration of the oligochaete, Enchytraeus japonensis. *Dev. Dyn.* 2006; 235:2051-2070.

Myohara M. Differential tissue development during embryogenesis and regeneration in an annelid. *Dev. Dynam.* 2004; 231(2): 349–358.

Myohara M. What role do annelid neoblasts play? A comparison of the regeneration patterns in a neoblast-bearing and a neoblast-lacking Enchytraeid Oligochaete. *PLoS One* 2012; 7(5): e37319.

Nakazato T, Ohta T, Bono H. Experimental design-based functional mining and characterization of high-throughput sequencing data in the sequence read archive. *PLoS One*. 2013;8(10):e77910.

Nene YL. Utilizing traditional knowledge in agriculture. Traditional knowledge system of India and Sri Lanka. 1999; 32–38.

Newmark PA, Sanchez Alvardo A. Bromodeoxiuridine specifically labels the regenerative stem cells of planarians. *Dev. Biol*. 2000; 220:142-153.

Niv CC, Lee JM, Myohara M. Glutamine synthetase gene expression during the regeneration of the annekid Enchytraeus japonensis. *Dev. Genes evol*. 2008; 218:39-46.

Pan W, Liu S, Ge F & Zheng T. Reconfirmation of anti-microbial activity in the coelomic fluid of the earthworm *Eisenia fetida andrei* by colorimetric assay. *J Biosci.* 2003; 28:723-731.

Park SK, Cho SJ & Park SC Histological observations of blastema formation during earthworm posterior regeneration. *Inver. Repro. Dev.* 2003; 57(2):165-169.

Parle JN. Micro-organisms in the intestines of earthworms. *J. Gen. Microbiol*. 1963; 31:1-11.

Parthasarathy K, Ranganathan LS, Anandi V, Josef Z. Diversity of Microflora in the gut and casts of tropical composting earthworms reared on different substrates. *J. Environ. Biol*, 2007; 28:87-97.

Patruno M, Thorndyke MC, Carnevali MDC, Bonasoro F, Beesley PW. Growth factors, heatshock proteins and regeneration in echinoderms. *J. Exp. Biol.* 2001; 843-848.

Pavel AB & Vasile CI. PyElph - a software tool for gel images analysis and phylogenetics *BMC Bioinform.* 2012; 13.

Pearson WR. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* 1990; 183:63–98.

Pérez-Losada, M., Ricoy, M., Marshall, J.C. & Domínguez, J. 2009. Phylogenetic assessment of the earthworm *Aporrectodea caliginosa* species complex (Oligochaeta: Lumbricidae) based on mitochondrial

Plytycz B & Morgan AJ. Riboflavin storage in earthworm chloragocytes/eleocytes in an ecoimmunology perspective. *Inv. Surv. J.* 2011; 8:199- 201.

Procházková P, Σilerová M, Felsberg J, Josková R, Beschin A, De Baetselier P & Bilej M. Relationship between hemolytic molecules in Eisenia fetida earthworms. *Dev Comp Immunol.* 2006; 30:381–392.

Randhawa GK & Kullar JS. Bioremediation of pharmaceuticals, pesticides, and etrochemicals with gomeya/cow dung. *ISRN Pharmacol.* 2011.

Ratcliffe NA, Rowley AF, FitzgeraldS W, Rhodes CP. Invertebrate immunity: basic concepts and recent advances. *Int. Rev. Cytol.* 1985; 97:183e350.

Reddien PW, Alvardo AS. Fundamentals of palnarian regeneration. Ann. Rev. *Cell Dev. Biol.* 2004; 20:725-757.

Richter M, Rosselló-Móra R, Glöckner F & Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinform.* 2015; 32:929–931.

Roch P. Protein analysis of earthworm coelomic fluid: I. Polymorphic system of natural hemolysin of Eisenia fetida andrei. *Dev Comp Immunol* 1979; 3:599–608.

Rochfort S, Wyatt MA, Liebeke M, Southam AD, Viant MR & Bundy JG. Aromatic metabolites from the coelomic fluid of Eisenia earthworm species' *Eur J Soil Biol.* 2017; 78:17-19.

Rodionova IA, Li X, Plymale AE, Motamedchaboki K, Konopka AE, Romine MF, Fredrickson JK, Osterman AL & Rodionov DA. Genomic distribution of B-vitamin auxotrophy and uptake transporters in environmental bacteria from the Chloroflexi phylum. *Environ Microbiol Rep* 2014; 7:204–10.

Rodriguez-Grau I, Venables BI, Fitzpatrick LC, Goven AJ. Suppression of secretory rosete formation by PCBs in *Lumbricus terrestris:* an earthworm assay for humoral immunotoxicity of xenobiotics. *Environ Toxicol Chem.* 1989: 1201-7.

Ross MG, Russ C, Costello M, Hollinger A, Lennon NJ, Hegarty R, Nusbaum C and Jaffe DB. Characterizing and measuring bias in sequence data. *Genome Biol*. 2013;14(5):R51.

Saha T, Chakraborty B, Das S, Thakur N, Chakraborty R. *Chryseomicrobium excrementi* sp. nov., a gram-stain-positive rod-shaped bacterium isolated from an earthworm (*Eisenia fetida*) cast. *Int J Syst Evol Microbiol*. 2018; 68(6): doi: 10.1099/ijsem.0.002791.

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; 4:406–425.

Salem H, Bauer E, Strauss AS, Vogel H, Marz M & Kaltenpoth M. Vitamin supplementation by gut symbionts ensures metabolic homeostasis in an insect host. *Proc. R. Soc. B.* 2014; 281.

Sanchez Alvarado A. Regeneration in the metazoans: why does it happen? *Bioessays.* 2000; 22:578–590.

Sanger, F., Nicklen, S., Coulson, A.R.,. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* USA 1977; 74, 5463–5467.

Santocki M, Morgan AJ & Plytycz B. Differential time course of restoration of experimentally depleted coelomocytes and fluorophores in the earthworm *Eisenia andrei*. *Folia Biologica* (Kraków). 2016; 64:121-130.

Schindler A. The Innate Immune Response in *Eisenia fetida* to Microbial Challenges," *Journal of Undergraduate Research at Minnesota State University*, Mankato 2004; 4:14.

Schwieger F, Tebbe CC. Efficient and accurate PCR amplification and detection of a recombinant gene in DNA directly extracted from soil using the expand TM high fidelity PCR system and T4 gene 32 protein. *Biochemica.* 1997; 2: 21–23.

Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinform*. 2014; 30:2068–2069.

Sheppard PS.. Specific differences in cocoon and hatchling production in Eisenia fetida and E. andrei. In: Edwards CA, Neuhauser EF, editors. Earthworms in waste and environmental management. The Hague, Netherlands: SPB.1988. p 83–92

Shimazu, R., Akashi, S., Ogata, H., Nagai, Y., Fukudome, K., Miyake, K., Kimoto, M., 1999. MD-2, a molecule that confers lipopolysaccharide responsiveness on Toll-like receptor 4. *J. Exp. Med.* 189, 1777–1782.

Sims RW, Gerard BM. 1985. Earthworms. Keys and notes for identification and study of the species. Synopses of the British fauna No. 31. London, UK: The Linnean Society for the Estuarine and. Brackish-Water Sciences Association.

Smibert RM, Krieg NR. Phenotypic characterization In: Gerhardt P, Murray RGE, Wood WA, Krieg NR. (editors) *Methods for General and Molecular Bacteriology*, Washington, DC: ASM. 1994; pp 617, 620, 622.

Smirnoff WA, & Heimpel AM. Notes on the pathogenicity of Bacillus thuringiensis var. tkuringiensis Berliner for the earthworm Lumbricus terrestris Linnaeus. *J. Insect Pathol.* 1961; 3:403-408.

Sneath PHA & Sokal RR. Numerical taxonomy: the principles and practice of numerical classification. San Francisco: Freeman. 1973; 573.

Stein E, Cooper EL. The role of opsonins in phagocytosis by coelomocytes of the earthworm, Lumbricus terrestris. *Dev Comp Immunol.* 1981;5(3):415–25.

Stephan-Dubois, F. Les neoblastes dans la regeneration posterieur des oaligochetes microdriles. *Bull. Biol. Fr. Belg.* 1954; 88: 182-247.

Stines-Chaumeil C, Talfowenier F, Branlant G. Mechanistic characterization of the MSDH (methylmalonate semialdehyde dehydrogenase) from *Bacillus subtlis*. *Biochem J* 2006;395: 107-115.

Stocker R. & Seymour JR. Ecology and Physics of Bacterial Chemotaxis in the Ocean. *Microbiol. Mol. Biol. Rev.* 2012; 76, 792–812.

Sugio M, Takeuchi K, Kutsuna J, Tadokoro R, Takahashi Y, Yoshidanoro C & Tochinai S. Exploration of embryonic origins of germline stem cells and neoblasts in *Enchytraeus japonensis* (Oligochaeta, Annelida). *Gene Exp. Patt.* 2008; 8:227-236.

Sugio, M., Yoshida-Noro, C., Ozawa, K. & Tochinai, S. Stem cells in asexual reproduction of *Enchytraeus japonensis* (Oligochaeta, Annelid): Proliferation and migration of neoblasts. *Dev. Growth Differ.* 2012; 54:439-450.

Sulik P, Klimek M, Talik P, Kruk J, Morgan AJ & Plytycz B. Searching for external sources of the riboflavin stored in earthworm eleocytes. *ISJ.* 2012; 9:169-177.

Sun ZJ. Purification, characteristics and induction of earthworm peptide. PhD Thesis, China Agricultural University. 1997.

Tadokoro R, Sugio M, Kutsuna J, Tochinai S & Takahashi Y. Early segregation of germ and somatic lineages during gonadal regeneration in the annelid *Enchytraeus japonensis*. *Curr. Biol.* 2006; 16:1012-1017.

Tamura K, Dudley J, Nei M & Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 2007; 24:1596-1599.

Tanaka, E.M. & Reddien, P.W. 2011. The cellular basis for animal regeneration. *Dev. Cell.* 2011; 21:172-185.

Thakuira D, Schmidt O, Finan D, Egan D, Doohan FM. Gut wall bacteria of earthworms: a natural selection process. *ISME J.* 2010; 4: 357–366.

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994; 22:4673–4680.

Thouveny Y, Tassava RA. Regeneration through phylogenesis. In Cellular and molecular basis of regeneration: from invertebrates to humans. Ferretti P, Geraudie J (eds) John Wiley & Sons, New York. 1998; 9-43.

Tindall BJ, Tomlinson GA, Hochstein LI. Transfer of *Halobacterium denitrificans* Tomlinson, Jahnke, and Hochstein to the genus *Haloferax* as *Haloferax denitrificans* comb. nov. *Int J Syst Bacteriol.* 1989; 39:359–360.

Tindall BJ. Qualitative and quantitative distribution of diether lipids in haloalkaliphilic archaebacteria. *Syst Appl Microbiol.* 1985; 6:243–246.

Tortoli, E. Phylogeny of the genus *Mycobacterium*: many doubts, few certainties. *Infect. Genet. Evol.* 2012; 12:827–831.

Tweeten, K. A. & Reiner A. Characterization of serine proteases of *Lumbriculus variegatus* and their role in regeneration. *Invert. Biol.* 2012; 131:322-332.

Undabarrena A, Ugalde JA, Castro-Nallar E, Seeger M, Cámara B. Genome sequence of *Streptomyces* sp. H-KF8, a marine Actinobacterium isolated from a northern chilean patagonian fjord. *Genome Announc. Am. Soc. Microbiol.* 2017;5:8–9. doi: 10.1128/genomeA.01645-16.

Valembois P, Roch P, Du Pasquier L. Degradation *in vitro* de proteines etrangeres par les macrophages du lombricien *Eisenia foetida* (Sav.). Comptes rendus des seances de l'Academie des Sciences, Paris, France ISerie DI 1973; 277:57-60.

Valembois P, Roch P, Lassegues M, Cassand P. Antibacterial activity of the haemolytic system from the earthworm *Eisenia fetida* andrei. *J Invertebr Pathol.* 1982;40:21–7.

Valembois P, Roch P, Lassegues M. Antibacterial molecules in annelids. In: Brehelin M, editor. Immunity in invertebrates. Berlin: Springer; 1986. p. 74–93.

Van der Sar AM, Stockhammer OW, van der Laan C, Spaink HP, Bitter W, Meijer AH. MyD88 innate immune function in a zebrafish embryo infection model, *Infect. Immun.* 2006; 74(4):2436–2441.

Ville P, Roch P, Cooper EL, Masson P, Narbone J-F. PCBs increase molecular-related activities (lysozyme antibacterial, hemolysis, proteases) but inhibit mecrophage-related functions (phagocytosis, wound healing) in earthworms. *J. Invert. Pathol.* 65: 217-224, 1995.

Visvanathan C, Trankler J, Jospeh K, Nagendran R. Vermicomposting as an Eco-Tool in Sustainable Solid Waste Management, Asian Institute of Technology, Annamalai University, Chidambaram. 2005.

Vitreschak AG, Rodionov DA, Mironov AA & Gelfand MS. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res.* 2002; 30:3141–3151.

Wayne LG, Kubica GP. Genus *Mycobacterium*. Sneath PHA, Mair NS, Sharpe ME, and Holt JG. *(editors)*. Bergey's Manual of Systematic Bacteriology, vol. 2. Williams & Wilkins, Baltimore, 1986; 1436–1457.

Wechter P, Williamson J, Robertson A, Kluepfel D (2003) A rapid, cost-effective procedure for the extraction of microbial DNA from soil. *World J Microbiol Biotechnol* 19:85–91.

Wee WY. Dutta A, and Choo SW. Comparative genome analyses of mycobacteria give better insights into their evolution. *PLoS ONE.* 2017;12:e0172831.

Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 1991; 173:697–703.

Wheaton S., M.D. Lambourne, A.J. Sarson, J.T. Brisbin, A. Mayameei, S. Sharif, Molecular cloning and expression analysis of chicken MyD88 and TRIF genes, *DNA Seq.* 2007; 18(6): 480–486.

Xiao N, Feng G & Edwards CA. The regeneration capacity of an earthworm, *Eisenia fetida*, in relation to the site of amputation along the body. *Acta Ecol. Sin.* 2011; 31:197–204.

Xing W, Li C, Zhenjun S. Differential expression of genes in the earthworm *Eisenia fetida* following exposure to *Escherichia coli* O157:H7. *Dev Comp Immunol.* 2011; 35:525–529.

Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W. *et al.*, Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014; 12(9).635-645.

Zattara EE, Bely AE. Evolution of a novel developmental trajectory: fission is distinct from regeneration in the annelid *Pristina leidyi*. *Evol. Dev.* 2011; 13:80-95.

Zhang GF, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L *et al.*, The oyster genome reveals stress adaptation and complexity of shell formation, *Nat.* 2012a; 490(7418):49–54.

Zhang S, Li CZ, Yan H, Qiu W, Chen GY, Wang PH, *et al.*, Identification and function of myeloid differentiation factor 88 in *Litopenaeus vannamei*, *PloS One*. 2012b; 7:e47038.

Zoran MJ, Martinez VG. *Lumbriculus variegatus* and the need for speed: A model system for rapid escape, regeneration and asexual reproduction. In: Annelids in Modern Biol. Shain DH (ed), John Wiley & Sons Inc. NY. 2009; 185-204.

INDEX

PUBLICATIONS

Thesis related publications:

Saha T, Chakraborty B, Das S, Thakur N, Chakraborty R. (2018). *Chryseomicrobium excrementi* sp. nov., a gram-stain-positive rod-shaped bacterium isolated from an earthworm (*eisenia fetida*) cast. *Int J Syst Evol Microbiol*.68(7): DOI: 10.1099/IJSEM.0.002791.

Tilak Saha, Ashim K. Chakravarty, Ranadhir Chakraborty Dynamics of Regeneration Process In *Eisenia Fetida* NBU Journal of Animal Sciences 2016 (10): 79-88. ISSN: 0975-1424

Other publications:

Ranjan VK, **Saha T,** Mukherjee S, Chakraborty R. Draft Genome sequence of a novel bacterium, Pseudomonas sp. MR02, capable of 2 pyomelanin production, isolated from River Mahananda at Siliguri, West Bengal, India Genome Announcements ASM 2018 6: e01443-17. https://doi.org/10.1128/genomeA .01443-17.

Denial Mahata, Malabendu Jana, Arundhuti Jana, Abhishek Mukherjee, Nibendu Mondal, **Tilak Saha**, Subhajit Sen, Chinmay K. Mukhopadhyay, Ranadhir Chakraborty, Golok B. Nando, Santi M. Mandal Lignin-*graft*-Polyoxazoline Conjugated Triazole a Novel Anti-Infective Ointment to Control Persistent Inflammation Scientific report (Nature Publishing Group). 2017. 7. Article Number 46412.

Rudra Prasad Roy, **Tilak Saha** and Ranadhir Chakraborty 2018 Contrasting observation in culturable aerobic and micro-aerophilic heterotrophic fish gut-bacteria: Intestine-breathing *Lepidocephalichthys guntea* (Hamilton Buchannan) versus gill breathing *Labeo rohit.Current Science*(Accepted)

Some selected Conference abstracts:

Tilak Saha and Ranadhir Chakraborty. Coelome microbiome and revelation of host-bacterial symbiosis in coelomic cavity of *Eisenia fetida.* in the international conference on "*Microbiology in the New Millennium: from Molecules to Communities*" held on October 27-29, 2017 at Bose Institute, Kolkata.

Tilak Saha and Ranadhir Chakraborty. Opsonization activity of amoebocytes of *Eisenia fetida* increases by exposure to *Bacillus thuringiensis.* in the national "*Zoological conference*" held on 26th March, 2017, at Department of Zoology, Raiganj University.

Tilak Saha, Ashim K. Chakravarty, Ranadhir Chakraborty. Tissue Reorganization During Regeneration In *Eisenia fetida*. in the "*West Bengal State Science and Technology Congress*" held on November 7-8, 2016 at Ananda Chandra College, Jalpaiguri.

INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

TAXONOMIC DESCRIPTION
Saha *et al.*, *Int J Syst Evol Microbiol*
DOI 10.1099/ijsem.0.002791

MICROBIOLOGY SOCIETY

Chryseomicrobium excrementi sp. nov., a Gram-stain-positive rod-shaped bacterium isolated from an earthworm (*Eisenia fetida*) cast

Tilak Saha,[1,2] Biswanath Chakraborty,[3] Sayak Das,[4] Nagendra Thakur[4] and Ranadhir Chakraborty[1,*]

Abstract

A Gram-stain-positive, rod-shaped, slightly halotolerant, nitrate-reducing bacterial strain, designated ET03^T, was isolated from the cast of an earthworm (*Eisenia fetida*) reared at the Centre of Floriculture and Agribusiness Management, University of North Bengal at Siliguri, West Bengal, India. On the basis of 16S rRNA gene sequence phylogeny, the closest relative of strain ET03^T was *Chryseomicrobium palamuruense* PU1^T (99.1 % similarity). The DNA G+C content of strain ET03^T was 42.9 mol%. Strain ET03^T contained menaquinone-8 as the most predominant menaquinone and phosphatidylethanolamine, phosphatidylinositol, phosphatidylinositol mannoside and phosphatidylglycerol as the main polar lipids. The diagnostic diamino acid was meso-diaminopimelic acid. Major cellular fatty acids were iso-$C_{15:0}$, $C_{16:1}\omega7c$ alcohol and iso-$C_{16:0}$. Other biochemical and physiological analyses supported genotypic and phenotypic differentiation of the strain ET03^T from its nearest taxonomic neighbours: *C. palamuruense*, *Chryseomicrobium amylolyticum*, *Chryseomicrobium imtechense*, *Chryseomicrobium aureum* and *Chryseomicrobium deserti*. The draft genome of strain ET03^T consisted of 2.64 Mb distributed in 14 scaffolds (N_{50} 894072). A total of 2728 genes were predicted and, of those, 2664 were protein-coding genes including genes involved in the degradation of polychlorinated biphenyl and several aromatic compounds. The isolate, therefore, represents a novel species, for which the name *Chryseomicrobium excrementi* sp. nov. is proposed. The type strain is ET03^T (=KCTC 33943^T=LMG 30119^T=JCM 32415^T).

The genus *Chryseomicrobium* was established by Arora *et al.* [1], within the family *Planococcaceae* of the order *Bacillales* and class *Bacilli* within the phylum *Firmicutes*, for some non-sporulating, non-motile, Gram-stain-positive rods. At present, five species isolated from diverse habitats, ranging from tropical soil, sewage sediment to desert soil, have been reported in this genus. These are *Chryseomicrobium imtechense* [1], *Chryseomicrobium amylolyticum* [2], *Chryseomicrobium aureum* [3], *Chryseomicrobium palamuruense* [4] and *Chryseomicrobium deserti* [5]. In the present study, we have characterized strain ET03^T, which represents a novel member of the genus *Chryseomicrobium*. The organism was isolated from the freshly liberated cast of an earthworm (*Eisenia fetida*) reared at the Centre of Floriculture and Agribusiness Management of University of North Bengal at Siliguri (26.7072° N, 88.3558° E), West Bengal, India.

For isolation of strain ET03^T, *Eisenia fetida* earthworms were washed several times with sterile distilled water and left on sterile tissue paper. Freshly liberated cast pellets were collected aseptically, serially diluted (in PBS, pH 7.2), plated on Luria agar (LA; M575, HiMedia) and incubated aerobically overnight at 37 °C. Single colonies were picked and purified by streaking on LA plates. Golden yellow coloured colonies that developed on LA plates were isolated and stored for taxonomic analyses. The strain was also capable of growing in other media such as nutrient broth (HiMedia) and tryptone soy broth (HiMedia). LA was used for

Author affiliations: [1]Department of Biotechnology, OMICS Laboratory, University of North Bengal, Siliguri 734 013, West Bengal, India; [2]Department of Zoology, Microbiology and Immunology Laboratory, University of North Bengal, Siliguri 734 013, West Bengal, India; [3]Central Instruments Laboratory, University of Kalyani, Kalyani 741235, West Bengal, India; [4]Department of Microbiology, Sikkim University, Gangtok, Sikkim, India.
***Correspondence:** Ranadhir Chakraborty, rcnbusiliguri@gmail.com
Keyword: *Chryseomicrobium* excrementi, ET03T, cellular fatty acids, 16S rRNA gene sequence, draft genome, COFAM.
Abbreviations: AAF, assembly and alignment-free; ANI, average nucleotide identity; COFAM, Center of Floriculture and Agribusiness Management; DDH, DNA–DNA hybridization; DSM, Deutsche Sammlung von Mikroorganismen; EMBL, European Molecular Biology Laboratory; FAME, fatty acid methyl esters; GGDC, Genome-to-genome direct comparison; JCM, Japan Collection of Microorganisms; KCTC, Korean Collection for Type Cultures; LMG, Laboratory of Microbiology- UGent; MK, menaquinone; NCBI, National Center for Biotechnology Information; ONPG, *ortho*-Nitrophenyl-β-galactoside; PE, phosphatidyl-ethanolamine; PI, phosphatidyl-inositol; PIM, phosphatidylinositol mannoside; PG, phosphatidyl-glycerol; TSB, tryptone soy broth; VP, Voges–Proskauer.
The GenBank accession number for the 16S rRNA gene sequence of ET03^T is KU230523.2. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number PCGR00000000. The version described in this paper is version PCGR01000000.
Three supplementary tables and four supplementary figures are available with the online version of this article.

maintenance of strain ET03^T and for the determination of phenotypic and chemotaxonomic characteristics.

Cell morphology and motility were determined with a phase-contrast microscope (Olympus CH2); details of the cell shape were ascertained with the help of a scanning electron microscope (EVO LS10, Zeiss). The Gram reaction was performed by the KOH lysis method [6] and further confirmed by the Gram-staining method of Claus [7]. The growth range of strain ET03^T was determined at 4, 10, 20, 28, 30, 37, 40 and 45 °C (±1 °C). For salt tolerance tests, 2, 4, 6, 8, 10, 15 and 20 % (w/v) NaCl was added to peptone–yeast extract medium (composition: 10 g peptone, 5 g yeast extract) devoid of NaCl or KCl. To assess growth at different pH levels, the pH of the sterile Luria–Bertani medium was adjusted from pH 3.0 to 12.0 by using either 0.1 M HCl or 0.1 M NaOH. Results were obtained after 48 h incubation at 37 °C. Catalase activity was examined by the production of oxygen bubbles after the addition of few drops of 3 % (v/v) H_2O_2. The ability to hydrolyse starch was determined by assessing the development of clear zones (after treatment with Gram's iodine solution) around the streaked culture on starch agar plates (nutrient agar 2.3 %; soluble starch 0.5 %; pH 7.2). The Voges–Proskauer test was performed by observation of colour development after the addition of alpha-naphthol and potassium hydroxide to the Voges–Proskauer broth of strain ET03^T culture. Urease test was performed by observing the development of a deep red colour at the periphery of the bacterial colonies on Christensen's urea agar plates [8]. Other biochemical characteristics such as the presence of amylase, cellulase, gelatinase, caseinase and phosphatase, tests such as H_2S production, indole test, methyl red, citrate utilization, ortho-nitrophenyl-β-galactoside test, nitrate reduction, ability to ferment glucose, lactose fermentation, and maltose fermentation ability were examined by following standard methods [9–11]. Phenotypic characterization of strain ET03^T was performed using the Biolog GEN III MicroPlate following the manufacturer's instructions. Briefly, bacterial suspension, prepared in a special 'gelling' inoculating fluid, was transferred onto the GEN III MicroPlate (100 µl per well). Incubation was carried out in an aerobic atmosphere for 48 h. Increased respiration due to the growth of bacteria using the single carbon source provided in each well caused reduction of the tetrazolium redox dye, forming a purple colour. The reactions were read using the fully automated OmniLog system. Carbon source utilization assays were also determined by using HiCarbohydrate kit parts A, B and C (HiMedia) according to the manufacturer's protocol. Antibiotic susceptibility (specific for oligotrophic bacteria) was determined according to the method described by Kumar *et al.* [12]. Susceptibilities to some of the antibiotics were also tested using the Biolog GEN III MicroPlate (columns 10–12) and the results were interpreted according to the manufacturer's instructions.

For the study of quinones and polar lipids, two-stage lipid extraction using methanol–hexane (2 : 1 v/v) described by Tindall *et al.* [13] was undertaken with modifications. Briefly, the menaquinone part was purified by running the hexane fraction on thin-layer chromatography (TLC) silica gel 60 F254 (Merck) using petroleum benzene: di-ethyl-ether (0.85 : 0.15) as the solvent. Further development of menaquinone components was performed using acetone–water (0.99 : 0.01) as the solvent and observed under UV light. The presence of *men* genes, coding enzymes for menaquinone biosynthesis, was identified in the genome of strain ET03^T to validate the results. Polar lipids were extracted from the methanolic phase using chloroform–methanol–0.3 % NaCl (1 : 2 : 0.8) as the extraction medium [14–16]. Polar lipids were separated by two-dimensional TLC on silica gel. In the first dimension, chloroform–methanol–water (65 : 25 : 4, v/v), and in the second dimension, chloroform–methanol– acetic acid–water (80 : 12 : 15 : 4, v/v) were used as the solvents. Lipid functional groups were identified using spray reagents specific for phospholipids (Mb-Blue), free amino groups (ninhydrin) and sugars (α-naphtol). The peptidoglycan structure was determined by using a hydrolysate of purified cell wall following methods described by Schleifer and Kandler [17] and subsequent TLC analyses as described by Staneck and Roberts [18], and Hancock [19]. The presence of genes coding for the specificity of Mur ligases (the enzymes responsible for the synthesis of the peptide stem of the peptidoglycan structure), were identified in the genome of strain ET03^T to validate the results. For analysis of fatty acids, fatty acid methyl esters were extracted from 36 h old (exponentially growing) cells grown in tryptone soy agar (M290; HiMedia) at 37 °C. They were then analysed by gas chromatography (Hewlett Packard 5890 II plus) and the Sherlock Microbial Identification System using version 4.10 of the TSBA40 library (Microbial ID).

The 16S rRNA gene of strain ET03^T was amplified from the genomic DNA prepared by the standard method [20], purified and sequenced according to Kumar *et al.* [12]. The obtained 16S rRNA gene sequence was compared with entries in the updated GenBank and EMBL databases by using the BLASTn program [21–23]. To determine the phylogenetic affiliation, the 16S rRNA gene sequence of strain ET03^{T}was aligned with the sequences of members of the genus *Chryseomicrobium* by using the CLUSTAL_W program [24]. Evolutionary relationships of members of genus *Chryseomicrobium* were inferred using three different tree-making algorithms (neighbour-joining [25], maximum-likelihood [26] and maximum-parsimony [27]) in MEGA 6.0 [28]. Phylogenetic analyses and the fidelity of the tree topologies were evaluated by bootstrap analysis with 1000 replicates [29, 30].

The genome of strain ET03^T was sequenced using NextSeq 500. Briefly, approximately 200 ng DNA was fragmented by covaris M220 to generate ~400 bp segments. End-

repaired products were size-selected by AMPure XP beads, PCR amplified with index primers and analysed in a 4200 Tape Station system (Agilent). After obtaining Qubit concentration, PE Illumina libraries were loaded onto the NextSeq 500 for cluster generation and sequencing. The copied reverse strands were then used to sequence from the opposite end of the fragment. Thus adapter-free data of 1.1 Gb was generated, which is required for the genome to be used for taxonomic purpose [31]. The high-quality reads were then *de novo* assembled by using the SPAdes genome assembler. Prokka [32] was used to predict the genes from final scaffolds. The NCBI annotation pipeline was used to annotate the whole genome sequence (WGS) of strain ET03^T. G+C content was estimated by using the genomic dataset. Since an earlier study has indicated that the G+C content varies no more than 1 % within species and by using 'logistic regression model based on all pairs of genomes, the probability of [digital DNA-DNA hybridization (DDH)] $\geq$70 % was 0.8443 for a 0 % difference in G+C content, 0.05 for 0.7271 %, 0.009 for 1 %, and virtually zero for 3 and 5 %' [33], DDH between ET03^T and other species of the genus *Chryseomicrobium* was not required because the difference in G+C content with the other five species is >5.0. An assembly and alignment-free (AAF) method [34] was used to reconstruct phylogeny from next-generation sequencing data. We calculated the BLAST-based average nucleotide identity (ANI) scores with the genome sequences of strain ET03^T and other related WGSs available in the databases using the JSpecies WS program with default parameters [35]. Analysis was also performed using all three equations in the Genome-to-Genome Distance Calculator (GGDC) online program, version 2.1 [33].

Cells of strain ET03^T were Gram-stain-positive, non-motile rods, measuring 1.5±0.5 µm long and 0.5 µm wide (Fig. S1, available in the online version of this article). Differential physiological and biochemical properties of strain ET03^T and its closest phylogenetic neighbours are shown in Table 1. Strain ET03^T contained menaquinone-8 (MK-8) as the most predominant menaquinone (Fig. S2). This phenotype corroborated with the genotype derived from WGS analyses of ET03^T. In the pathway of MK-8 biosynthesis, enzymes encoded by *men* genes (*men*F, *men*D, *men*H, *men*C, *men*E, *men*B and *men*A) and the *ubi*E (menG) gene have been duly annotated in the NCBI database. The enzymes engaged in MK-8 biosynthesis are: MenF – isochorismate synthase (accession no. PJK17315), MenD – 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (PJK17316), MenH – SHCHC synthase (PJK17317), MenC – o-succinyl benzoate synthase (PJK17077), MenE – o-succinylbenzoic acid-CoA synthase PJK17319), MenB – 1,4-dihydroxy-2-naphthoyl-CoA synthase (PJK17318), MenA – 1,4-dihydroxy-2-napthoate octaprenyl-transferase (PJK16692) and UbiE (MenG) – 2-methoxy-6-polyprenyl-1,4-benzoquinol methylase (PJK17772). Phosphatidyl ethanolamine, phosphatidyl inositol, phosphatidyl glycerol and phosphatidyl inositolmannoside were identified as the phosplolipids present in strain ET03^T (Fig. S3). The cell-wall peptidoglycan of strain ET03^T contained L-alanine, D-glutamic acid, and meso-diaminopimelic acid in the peptide stem. In the bacterial peptidoglycan structure, the variation of the peptide stem is due to the specificity of the Mur ligases, the enzymes responsible for its synthesis. The first, second and third amino acids of the peptide stem are added by the MurC, MurD and MurE ligases, respectively. In the genome of strain ET03^T, *mur*C, *mur*D, and *mur*E gene was analysed to code for UDP-N-acetylmuramate-alanine ligase (accession no. PJK17354), UDP-N-acetylmuramoylalanine-D-glutamate ligase (PJK17947) and UDP-N-acetylmuramoyl-L-alanyl-D-glutamate-2,6-diaminopimelate ligase (PJK166217 and PJK15783), respectively.

The major cellular fatty acids were 13-methyltetradecanoic acid or iso-$C_{15:0}$ (45 %), (9Z)-9-hexadecenoic acid or $C_{16:1}$ $\omega7c$ alcohol (13 %) and 14-methylpentadecanoic acid or iso-$C_{16:0}$ (11.8 %). Trace amounts of 12-methyltridecanoic acid or iso-$C_{14:0}$ (6.2 %), 12-methyltetradecanoic acid or anteiso-$C_{15:0}$ (3.4 %), 15-methylhexadecanoic acid or iso-$C_{17:0}$ (2.2 %) and 14-methylhexadecanoic acid or anteiso-$C_{17:0}$ (1.6 %) were present, which is typical of members of the genus *Chryseomicrobium*, but the proportions differed from those reported for other members of the genus (Table S1).

An almost-complete 16S rRNA gene sequence comprising 1517 bp was obtained, which was reliable for delineation of taxonomic hierarchy [36]. According to the comparison of 16S rRNA gene sequences (using NCBI, BLASTN), the closest relatives of strain ET03^T were *C. palamuruense* strain PU1^T (99.1 %), *C. aureum* strain BUT-2 (99.0 %), *Psychrobacillus psychrodurans* strain BAB-2243 (99.0 %), *C. amylolyticum* strain ID4 (98.9 %), *C. imtechense* strain HWG-A7 (98.4 %), *C. deserti* strain THG-T1 (96.3 %), *Planococcus rifietoensis* strain M8 (95.7 %), *Planococcus plakortidis* strain DSM 23997 (95.6 %) sequentially followed by other members of the Family *Planococcaceae*. A tree depicting the phylogenetic position of strain ET03^T within the genus *Chryseomicrobium* is shown in Fig. 1. Based on 16S rRNA gene sequence comparison, strain ET03^T forms a distinct subcluster with *C. palamuruense* PU1^T and other members of the genus *Chryseomicrobium*. Using BLASTn, it was revealed that the *gyr*B sequence of strain ET03^T produced maximum identity (83 %) with the *gyr*B gene sequence of *C. imtechense* MW10 (accession no. HM989964; the only *gyr*B sequence from any strain of the genus *Chryseomicrobium* available in the database) followed by *Planococcus* species PAMC 21323 (accession no. CP009129; 76 %) and *Planococcus maritimus* strain Y42 (accession no. CP019640; 74 %) showing greater sequence divergence, compared to the BLASTn results with the 16S rRNA gene sequence of ET03^T.

The draft genome of strain ET03^T includes 2 644 068 bp distributed in 14 scaffolds (N_{50} 894072). Based on whole-genome data, the DNA G+C content of ET03^T was calculated as 42.9 mol%. The difference of G+C content between strain ET03^T and the five previously described species was >5 %, which is well within the range consistent with

Table 1. Characteristics that differentiate strain ET03^T from other members of the genus *Chryseomicrobium*

Strains: 1, ET03^T; 2, *Chryseomicrobium imtechense* MW 10^T; 3, *Chryseomicrobium amylolyticum* JC16^T; 4, *Chryseomicrobium aureum* BUT-2^T; 5, *Chryseomicrobium palamuruense* PU1^T; 6, *Chryseomicrobium deserti* THG-T1.1B^T. +, Positive; –, negative. AL, unidentified aminolipid; DPG, diphosphatidylglycerol; GL, unidentified glycolipid; PE, phosphatidyl ethanolamine; PG, Phosphatidyl-glycerol; PI, phosphatidyl inositol; PIM, phosphatidyl inositolmannoside; PL, unidentified phospholipid; m-DAP, meso-diaminopimelic acid; ND, not detected.

Characteristic	1	2	3	4	5	6
Cell size (μm) (length×diameter)	1.5–2.2×0.5–0.6	1.7–2.9×0.3–0.7	2.0–3.0×1.0	1.5–2.0×0.5–0.86	1.6–2.0×0.6–0.7	2.4–2.7×0.5–0.7
Motility	Non-motile	Non-motile	Non-motile	Non-motile	Motile	Non-motile
Growth temperature range (optimum)	20–40 °C (35–37)	4–45 °C	25–40 °C (30–37)	20–35 °C	18–40 °C	20–35 °C (28–30)
pH range (optimum)	6–9	6–9	7–11 (7–8)	7–10	7 to 10	5–7 (7)
NaCl tolerance limit (%, w/v)	8	6	5	7	9	3
Catalase reaction	–	–	+	–	+	+
Nitrate reduction	+	–	–	–	–	–
Voges–Proskauer reaction	–	+	–	–	–	–
Urease activity	–	–	–	+	+	+
Oxidase	–	–	–	–	+	–
Hydrolysis of:						
Starch	–	–	+	–	+	–
Gelatin	+	–	–	+	–	- ?
Organic substrates utilized for growth:						
Citrate	–	+	–	–	+	+
Glycerol	–	+	+	–	+	–
Acid production from various carbohydrates:						
Glucose	+	+	–	–	+	–
Salicin	+	+	–	+	–	–
Mannose	–	–	+	–	+	+
Fructose	+	+	–	–	–	–
Maltose	+	+	–	–	+	+
Sucrose	–	+	–	–	–	+
Inulin	–	+	–	+	–	–
Trehalose	–	–	+	–	–	+
Melibiose	–	+	–	–	+	+
Cellobiose	–	+	–	–	+	–
Menaquinones (descending abundance)	MK-8, MK-7, MK-6	MK-7, MK-8, MK-7$_{H2}$, MK-6	MK-7, MK-8, MK-6	MK-7, MK-6, MK-8	MK-8	MK-7, MK-8, MK-6
Polar lipids	PE, PI, PIM, PG	DPG, PG, PE, PC, GL	DPG, PG, PE, AL, PL,	DPG, PG, PE, PL	DPG, PG, PE	DPG, PE, PG, GL, AL
Peptidoglycan type	m-DAP	L-Lys-D-Asp	L-Orn-D-Glu	L-Orn-D-Glu	L-Orn	L-Orn-D-Glu
DNA G+C content (mol%)	42.9	53.4	57.6	48.5	48.5	50.4

species belonging to the same genus (33). A total of 2728 genes were predicted, of which, 2664 were protein-coding genes including 121 *de novo* genes with no BLAST hit. Of the protein-coding genes, there were at least six genes coding for several dioxygenases [2-nitropropane dioxygenase (accession no. PJK17992), biphenyl 2,3-dioxygenase (PJK16435), glyoxalase/bleomycin resistance/estradiol dioxygenase (PJK16272) and three ring-cleaving dioxygenases (PKJ15803, PKJ15871 and PJK15941). These microbial genes are responsible for degradation of biphenyls such as polychlorinated biphenyl and several xenobiotic aromatic compounds in the environment. The phylogenetic tree reconstructed from next-generation sequencing data using the AAF method depicts the distinction of strain ET03^T from its taxonomic neighbours (Fig. S4). ANI scores generated during global comparisons of the genome sequence [37] of strain ET03^{T}with previously deposited WGSs in databases indicates sufficient distance from *Bhargavaea cecembensis* T14, *Jeotgalibacillus malaysiensis* D5, *Planomicrobium glaciei* UCD-HAM, *Solibacillus silvestris* StLB046, *Paenisporosarcin*a species HGH0030, *Planococcus kocurii* ATCC 43650, *Planococcus maritimus* MKU009, *Planococcus plakortidis* DSM 2399739, *Planococcus rifietoensis* M8 and *Sporosarcina psychrophila* DSM 6497, which supports the findings of 16S rRNA gene phylogeny (Table S2). The ANI

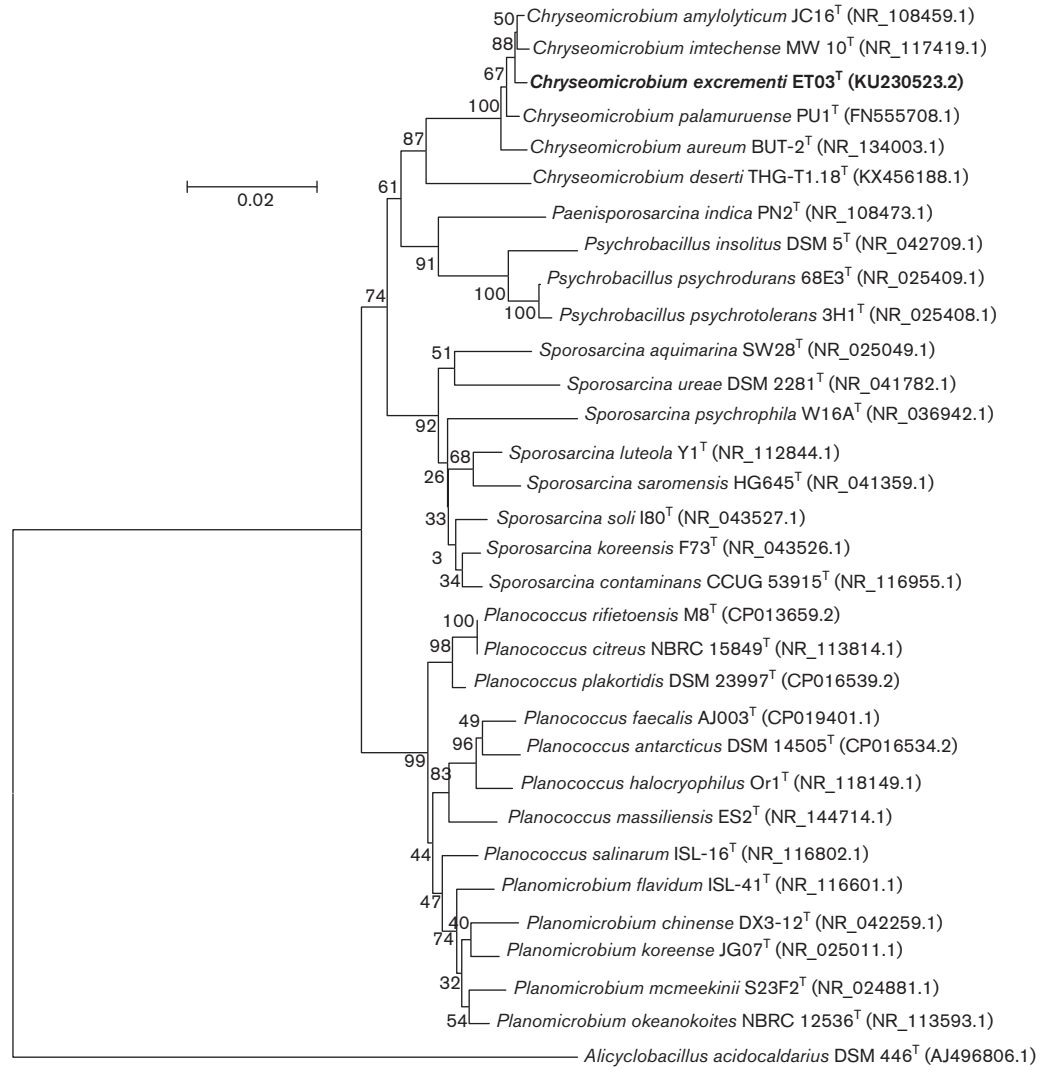


Fig. 1. Phylogenetic tree reconstructed by the neighbour-joining method based on 16S rRNA gene sequences showing the phylogenetic relationship between strain ET03^T and closely related species. Bootstrap percentages (based on 1000 replications) are shown at the nodes. The GenBank accession numbers for 16S rRNA gene sequences are shown in parentheses. *Alicyclobacillus acidocaldarius* strain DSM 446 (AJ496806.1) was used as an outgroup. Bar, 2 nt substitution per 100 nt.

score for strain comparisons between EAG3^T and *Paenisporosarcin*a species HGH0030 was the highest 68.6 % (coverage 38.8 %), which is far below 95–96 % cut-off value for novel species determination by this approach. The GGDC results with BLAST+ for strain EAG3^T and other taxonomic neighbours in the family *Planococcaceae* gave DNA–DNA homology values of <30 % by any of the three models used, assuring sufficient distance among the genomes taken for the analysis (Table S3).

Strain ET03^T can be readily differentiated from its closest relative, *C. palamuruense* PU1^T, with reference to its physiological and biochemical characteristics (e.g. strain ET03^T is negative for catalase, urease and oxidase but PU1^T is positive for those whereas ET03^T is positive for nitrate reduction but PU1^T is negative), cellular fatty acids, 16S rRNA gene sequence and DNA G+C content (Table 1). On the basis of the data obtained from our study using a polyphasic taxonomic approach, strain ET03^T merits recognition as a novel species of the genus *Chryseomicrobium*, for which we propose the name *Chryseomicrobium excrementi* sp. nov.

DESCRIPTION OF *CHRYSEOMICROBIUM EXCREMENTI* SP. NOV

Chryseomicrobium excrementi (ex.cre.men′ti. L. gen. n. *excrementi* - of excreta).

Cells stain Gram-positive and are non-motile, spore-forming rods, 1.5±0.5μm long and 0.5 μm wide. Colonies are golden yellow, flat, circular with a smooth margin when grown on LA plates for 16 h at 37 °C. Growth is observed at

20–40 °C (optimum, 35–37 °C), pH 6–9 (optimum, 7.0) and can tolerate concentrations of NaCl up to 8 % (optimum, 2 %). Negative for catalase and oxidase. Can utilize D-salicin, D-fructose, 3-methyl glucose, L-rhamnose, D-arabitol, myo-inositol, D-glucose-6-PO4, L-arginine, L-aspartic acid, L-glutamic acid, L-pyroglutamic acid, D-gluconic acid, quinic acid, L-malic acid, acetoacetic acid and acetic acid; but is unable to utilize trehalose, cellobiose, gentiobiose, sucrose, turanose, stachyose, raffinose, lactose, melibiose, methyl α-D-glucoside, *N*-acetyl-D glucosamine, *N*-acetyl-D galactosamine, *N*-acetyl neuraminic acid, D-mannose, D-galactose, D-fucose, L-fucose, Inosine, D-sorbitol, D-mannitol, glycerol, D-fructose-6-PO_4, D-aspartic acid, D-serine, gelatin, glycyl-L-proline, L-alanine, L-histidine, L-serine, pectin, D-galacturonic acid, D-glucuronic acid, glucuronamide, mucic acid, D-saccharic acid, methyl pyruvate, D-lactic acid methyl ester, L-lactic acid, citric acid, keto-glutaric acid, D-malic acid, bromo-succinic acid, Tween 40, amino-butryric acid, hydroxy- butyric acid, keto-butyric acid, propionic acid and formic acid. The strain is sensitive to fusidic acid, D-serine, troleandomycin, minocycline, lincomycin, guanidine HCl, Niaproof 4, vancomycin and tetrazolium blue. The strain is resistant to 1 % sodium lactate, nalidixic acid, lithium chloride, potassium tellurite, aztreonam and sodium butyrate.

Strain ET03^T contains MK-8 as the predominant menaquinone and phosphatidylethanolamine, phosphatidylinositol, phosphatidylinositol mannoside and phosphatidylglycerol as the main polar lipids. The diagnostic diamino acid is meso-diaminopimelic acid. The major cellular fatty acids are iso-$C_{15:0}$, $C_{16:1}\omega7c$ alcohol and iso-$C_{16:0}$.

The type strain, ET03^T (=KCTC 33943^T=LMG 30119^T =MCC378^T), was isolated from cast of an earthworm (*Eisenia fetida*) reared at the Centre of Floriculture and Agribusiness Management of University of North Bengal at Siliguri (26.7072° N, 88.3558° E), West Bengal, India. The DNA G+C content of strain ET03^T is 42.9 mol%.

Funding information
The work was partially supported by the UGC Innovative project grant entitled 'Development of high-resolution optical detector for growth monitoring of bacteria in liquid culture' (North Bengal University, India).

Acknowledgements
Dr Sadhan Ch. Das, Department of Diabetes Complications and Metabolism, Beckman Research Institute of City of Hope, Duarte, California, USA is duly acknowledged for literature support that helped us to write this manuscript.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Arora PK, Chauhan A, Pant B, Korpole S, Mayilraj S *et al.* *Chryseomicrobium imtechense* gen. nov., sp. nov., a new member of the family *Planococcaceae*. *Int J Syst Evol Microbiol* 2011;61:1859–1864.
2. Raj PS, Sasikala C, Ramaprasad EV, Subhash Y, Busse HJ *et al.* *Chryseomicrobium amylolyticum* sp. nov., isolated from a semi-arid tropical soil, and emended descriptions of the genus *Chryseomicrobium* and *Chryseomicrobium imtechense*. *Int J Syst Evol Microbiol* 2013;63:2612–2617.
3. Deng SK, Ye XM, Chu CW, Jiang J, He J *et al.* *Chryseomicrobium aureum* sp. nov., a bacterium isolated from activated sludge. *Int J Syst Evol Microbiol* 2014;64:2682–2687.
4. Pindi PK, Ashwitha K, Rani S. *Chryseomicrobium palamuruense* sp. nov., a bacterium isolated from activated sludge. *Int J Syst Evol Microbiol* 2016;66:3731–3736.
5. Lin P, Yan ZF, Li CT, Kook M, Wang QJ *et al.* *Chryseomicrobium deserti* sp. nov., isolated from desert soil in South Korea. *Int J Syst Evol Microbiol* 2017;67:4126–4131.
6. Murray RGE, Doetsch RN, Robinow CF. Determinative and cytological light microscopy. In: Gerhardt P, Murray RGE, Wood WA and Krieg NR (editors). *Methods for General and Molecular Bacteriology*. Washington, DC: American Society for Microbiology; 1994. pp. 21–41.
7. Claus D. A standardized Gram staining procedure. *World J Microbiol Biotechnol* 1992;8:451–452.
8. Macfaddin JF. *Biochemical Tests for Identification of Medical Bacteria*, 3rd ed. Baltimore MD: Williams and Wilkins; 2000.
9. Gordon RE, Barnett DA, Handerhan JE, Pang CH-N. *Nocardia coeliaca*, *Nocardia autotrophica*, and the Nocardin strain. *Int J Syst Bacteriol* 1974;24:54–63.
10. Lanyi B. Classical and rapid identification methods for medically important bacteria. *Methods Microbiol* 1987;19:1–67.
11. Smibert RM, Krieg NR. Phenotypic characterization. In: Gerhardt P, Murray RGE, Wood WA and Krieg NR (editors). *Methods for General and Molecular Bacteriology*. Washington, DC: American Society for Microbiology; 1994. pp. 617, 620, 622.
12. Kumar A, Mukherjee S, Chakraborty R. Characterization of a novel trimethoprim resistance gene, dfrA28, in class 1 integron of an oligotrophic *Acinetobacter johnsonii* strain, MB52, isolated from River Mahananda, India. *Microb Drug Resist* 2010;16:29–37.
13. Tindall BJ, Tomlinson GA, Hochstein LI. Transfer of *Halobacterium denitrificans* (Tomlinson, Jahnke, and Hochstein) to the genus *Haloferax* as *Haloferax denitrificans* comb. nov. *Int J Syst Bacteriol* 1989;39:359–360.
14. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 1959;37:911–917.
15. Minnikin DE, O'Donnell AG, Goodfellow M, Alderson G, Athalye M *et al.* An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *J Microbiol Methods* 1984;2:233–241.
16. Tindall BJ. Qualitative and quantitative distribution of diether lipids in haloalkaliphilic archaebacteria. *Syst Appl Microbiol* 1985;6:243–246.
17. Schleifer KH, Kandler O. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* 1972;36:407–477.
18. Staneck JL, Roberts GD. Simplified approach to identification of aerobic actinomycetes by thin-layer chromatography. *Appl Microbiol* 1974;28:226–231.
19. Hancock IC. Analysis of cell wall constituents of Gram-positive bacteria. In: Goodfellow M and O'Donnell AG (editors). *Chemical Methods in Prokaryotic Systematics. Modern Microbiological Methods*. Chichester, New York, Brisbane, Toronto, Singapore: John Wiley and Sons Ltd.; 1994. pp. 64–84.
20. Marmur J. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J Mol Biol* 1961;3:208–218.
21. Pearson WR. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol* 1990;183:63–98.
22. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215:403–410.
23. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–3402.

24. **Thompson JD, Higgins DG, Gibson TJ.** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–4680.
25. **Saitou N, Nei M.** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987;4:406–425.
26. **Felsenstein J.** Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 1981;17:368–376.
27. **Fitch WM.** Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool* 1971;20:406–416.
28. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 2013;30:2725–2729.
29. **Felsenstein J.** Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 1985;39:783–791.
30. **Tamura K, Dudley J, Nei M, Kumar S.** MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007;24:1596–1599.
31. **Chun J, Oren A, Ventosa A, Christensen H, Arahal DR *et al.*** Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461–466.
32. **Seemann T.** Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–2069.
33. **Meier-Kolthoff JP, Klenk HP, Göker M.** Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 2014;64:352–356.
34. **Fan H, Ives AR, Surget-Groba Y, Cannon CH.** An assembly and alignment-free method of phylogeny reconstruction from next-generation sequencing data. *BMC Genomics* 2015;16:522.
35. **Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J.** JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 2016;32:929–931.
36. **Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W *et al.*** Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014;12:635–645.
37. **Kim M, Oh HS, Park SC, Chun J.** Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–351.