Comparative Metagenomics To Study The Impact Of Methyl Parathion On Farmland Bacterial Populations

by
Meron Kidanemariam

A thesis submitted to the Department of Engineering Technology,
College of Technology
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCES

in Engineering Technology

Chair of Committee: Dr. Rupa Iyer

Committee Member: Dr. Moges Mequanint

Committee Member: Dr. Abdul Latif Khan

University of Houston
December 2021

Copyright 2021, Meron Kidanemariam

ACKNOWLEDGMENTS

I would like to recognize Dr. Rupa Iyer for giving me this tremendous opportunity to work in her research laboratory. I would also like to thank Professor Brian Iken and Dr.Moges Mequanint, who provided suggestions and advice throughout this process. I offer special gratitude to my mother, who supported and understood me; I could never reach this accomplishment without my mother. I'm thankful to my professors for participating in my thesis committee. I'm also grateful for my friends Rumana Mustafa and Hao Chau, who played a big role in my academic success. Thank you all for your constant support.

ABSTRACT

Methyl parathion is a broad-spectrum organophosphate insecticide and acaricide that has been classified as a Class I insecticide with restrictions for use by the U.S. Environmental Protection Agency. It is commonly applied as a dust, powder, concentrate, emulsion, or granules to inhibit boll weevils and other insects found in agricultural crops. Studies showed that methyl parathion (MP) is generally hydrolyzable and rapidly breaks down in water and sediments via photolysis and biodegradation by various microorganisms. This study uses metagenomics to compare the effect of MP on farmland soil microorganisms population. It was found that the unamended soil sample contained an abundant phylum of proteobacteria (45%) which are known to degrade MP. However, upon amending the soil with MP a net reduction of microbial species was observed, including in species harboring degradation biomarkers. Based on bioinformatics data collected from MG-RAST, the methyl parathion degradation (MPD) pathway is likely incomplete, effectively blocked with the formation of p-nitrophenol (PNP). PNP is also poisonous and significantly reduces bacterial metabolism and growth. p-nitrophenol genes are primarily found on mobile elements that were not present within these soil samples. It is therefore likely that upon inoculation into MP , the underlying bacterial population degraded MP to PNP but were unable to proceed further and slowly died off. The putative MPD species found for this study is the Rhizobium sp. MTP-10005, an alphaproteobacterium, since it is the one that has the biodegradation capability via the maleylacetate and hydroxyquinol of the methyl parathion degradation pathway.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

In a natural ecosystem, a symbiotic relationship exists between plants and microorganisms. Microbes help plants solubilize nutrients, obtain water, protect from pathogens and pests, break down harmful compounds, and prevent loss of nutrients. In turn, plants supply the microbes with compounds they need to survive through plant secretions at the roots. However, the introduction of chemicals that are thought to enhance the growth of plants disrupt the normal functioning of the microorganisms. Compounds such as fertilizer, pesticides, or herbicides could destroy or cause the microbes to mutate [1]. Measuring the effect of chemical compounds on the soil microbes is an important step to help restore the healthy functioning of the microbe-plant ecosystem.

1.1 Methyl parathion: An organophosphate compound

Organophosphates (OPs) are chemical compounds derived from the esterification reaction of phosphoric acid and alcohols. OPs irreversibly bind to Acetylcholine esterase (AChE) preventing ACh from breaking down into choline and acetate. This condition will then cause overstimulation of both the nicotinic and muscarinic receptors causing damages to the central nervous system, respiratory, cardiovascular, and gastrointestinal systems. For these reasons, OPs are commonly used as major components of pesticides, herbicides, and insecticides, as well as in nerve gas. They are highly toxic and acute or prolonged exposure to them has been documented to cause serious damage to plants, animals, insects, and humans [2]. This study will focus on the effect of an OP - *Methyl Parathion* - on the bacterial load of soil samples.

Methyl parathion (MP) - *O*,*O*-dimethyl *O*-p-nitro6262phenyl phosphorothioate - also known as *metaphos,* is a broad-spectrum OP insecticide and acaricide [3,4]. The U.S. Environmental Protection Agency (EPA) has classified MP as a *Class I* insecticide with use restrictions. MP is

commonly applied as a dust, concentrate, wettable powder, emulsion, or granules. It is mostly used to inhibit boll weevils and other insects found in agricultural crops such as alfalfa, cotton, barley, sorghum, corn, sunflower, wheat, and soybeans. It exists as a white crystalline substance in its pure form. It is typically sold as a technical grade compound containing 80 percent pure MP. MP's authorized application is for outdoor use under a licensed applicator. However, there have been several reported cases of illegal MP applications that ranged from indoor household pesticides to use as a suicidal poison [5].

1.1.1 Methyl Parathion toxicity

Exposure to MP could happen in several ways. Manufacturers and farm applicators are the most commonly affected although there are reported cases of MP exposures through water and food. The unauthorized use of MP as a pesticide or insecticide is risky for handlers that are not trained on the proper way to handle it. Those living near the disposal sites of hazardous wastes are also in danger of high–level exposure. As a properly applied insecticide, MP rapidly breaks down in water and sediments via photolysis and biodegradation by microorganisms [3,5]. It can be transported to the surrounding areas by wind, fog, and rain. Owing to its highly hydrolyzable nature, MP has little chance of contaminating the food chain as long as the prescribed curing period is observed. The half–life of MP is generally affected by temperature: at 0^{o}C, the half–life is 48 to 57 days; at 20^{o}C, the half–life is from 9.2 to 10.5 days; at 40^{o}C, its half–life is from 1.3 to 1.5 days [5].

Once MP enters the human body through skin contact or inhalation, it is either activated or detoxified. It goes into the bloodstream and the liver, brain, heart, and other organs. Activation occurs when the liver converts some of MP into *methyl paraoxon* which is extremely potent. Methyl paraoxon is responsible for the inactivation of AChE through the process of

phosphorylation of the enzyme resulting in the accumulation of the neurotransmitter. After several hours or days, both chemicals, MP and methyl paraoxon are excreted through the urine. At best, MP is not considered a carcinogenic substance [5].

1.1.2 The MP Degradation Pathway

The MP degradation pathway for this study involves the enzyme methyl parathion hydrolase (MPH) which was found present in the farmland soil BLASTx analysis. MPH was found to hydrolyze a wide range of OP, including lactone substrates, methyl parathion, and aryl esters. It also has the demonstrated ability to have *metal-ion dependent selectivity patterns* [6]. Its catalytic proficiency to hydrolyze MP was found to be high at K_{cal}/K_M of 10^6/M.s [6].

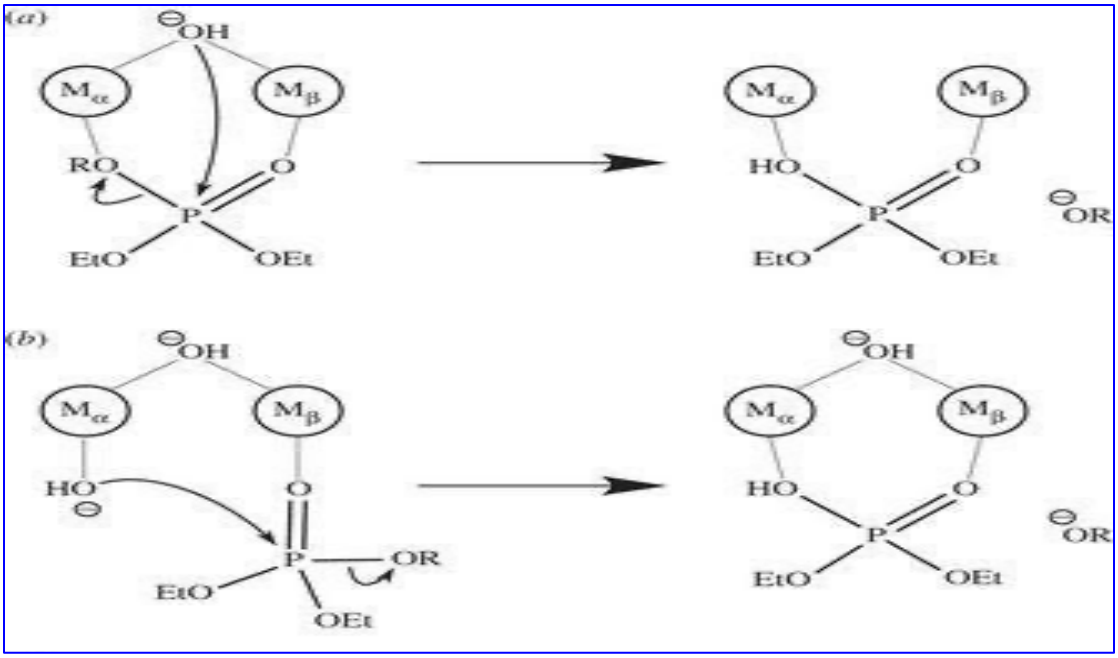


Figure 1.1: Methyl Parathion Hydrolase-catalyzed degradation pathway [6]

Figure 1.1 shows the comparison of two probable catalytic pathways for methyl paraoxon hydrolysis by MPH. The first mechanism (a) involved the "μ-bridging hydroxide ion as a nucleophile" [6]. The two enzymes' active sites were similar and possessed *binuclear active sites*. Both of the two enzymes also have three specific hydrophilic pockets. The mechanism

showed that due to steric limitations, it was not feasible for the substrate to attach to the active site cavity of the MPH that would make it vulnerable to in-line attack by the bridging hydroxide ion. The second mechanism (b) was an alternate pathway which involved the nucleophilic attack of a terminal hydroxide ion attached to the α-metal ion on a phosphate radical that has a monodentate relationship with the β-metal ion via the P=O bond, with similar cleavage of the P-O bond to the leaving aryl group [6].

1.2 MPH Enzyme structure and MP Catalysis

The MPH enzyme structure is composed of two identical protein sub-units - a *homodimer* - with two independent mixed β-sheets that have three α-helices which are exposed to solvent. Each of the sub-unit contains the *β-lactamase-like domain* [7], that has two metal ions as nucleus. The two-metal nucleus is flanked by two β-sheets with two αβ-loops around [7], as in Figure 2. The structure is colored red from the C terminal to blue at the N terminal. The two metals comprising the nucleus are Zinc (silver-colored) and Cadmium (gold-colored) [7]. The molecular weight of MPH is 37 kilodalton (kDa) and its turnover number is 114.70 +/- 13.19 per second [39].

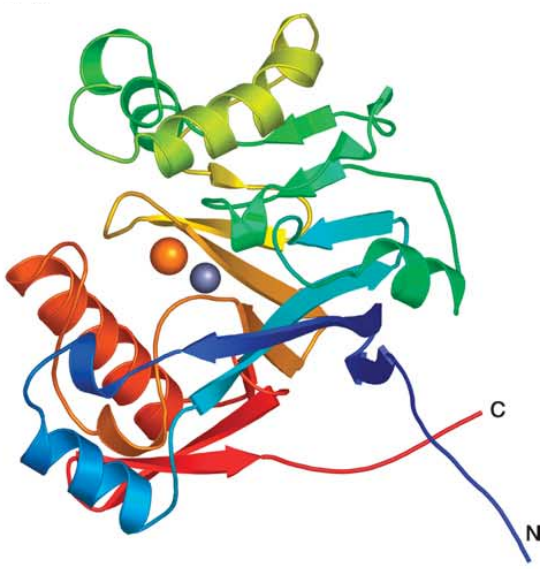


Figure 1.2: Methyl Parathion Hydrolase Structure [7]

The primary role of metal ions in metallo-β-lactamase (MBL) family, such as the MPH, was to act as an activator for a metal-ion-bond hydroxide ion that would perform a nucleophilic attack. The secondary role was to act as *Lewis acids* that would stabilize the reaction and the transition states of the intermediate compounds involved [6].

1.3 Project rationale, objectives, and hypothesis

The focus of this research project is to use comparative metagenomics to determine the impact of an OP pesticide which is methyl parathion (MP) to farmland soils. The old method of measuring the effect of chemicals on soil microbes involved isolating first the microbes from the soil samples using culture-dependent methods, then subjecting the isolated microorganisms to a sequence of tests to analyze and identify the microbes. This process could only culture one percent of the soil microbes [31]. To solve this problem, scientists developed alternative methods that are culture-independent but based on molecular biology, one of which is metagenomics. Metagenomics refers to the direct analysis of genetic materials (genomes) from a mixed community of organisms such as an environmental sample. It is usually used when it is difficult to separate a microbe from the others [9]. Through metagenomics, the functional component of the genes of the microbial population could be derived. Metagenomics could also provide genetic information on new enzymes or biocatalysts, as well as evolutionary profiles of microbial structure and community function [8].

For this study, metagenomics would be used to quantify the effect of MP on soil microorganisms. The results of this study would be useful in determining ways to treat excessive MP soil contamination.

Following are the objectives for this research study:

- ✧ Conduct metagenomic sequencing of soil samples before and after exposure to MP to know the change in their genetic makeup.
- ✧ Analyze the results of the metagenomic data and determine the effect of MP on the microbial population of the soil samples.
- ✧ Identify the degradation biomarkers of methyl parathion. These are the microorganisms that can break down the MP and propagate themselves in the process.

The hypothesis for this study is that the introduction of MP in the soil will have different effects on the soil microbial community. It would eliminate some, but would propagate those that are capable to degrade it, as MPD genes.

II. REVIEW OF LITERATURE

2.1 Chemodynamics of Methyl Parathion

The large-scale application of MP to crops is commonly done through spraying, either ground or airborne with the use of an aircraft. Once released into the atmosphere, MP deposits onto plants and soil by forming bond residues that restrict its movement to other areas. The MP soil adsorption is generally influenced by the amount of organic matter present and its *cation exchange capacity* (CEC). It was found that MP's leaching capacity is highly dependent on the pH of the soil. In general, the chemodynamics of pesticides could be controlled by increasing the soil adsorption area and making the structure microporous [3].

The investigation conducted on the adsorption capacity of MP revealed the following results:

2.1.1 Effect of pH

It was observed that there was an increase in the soil adsorbent capacity when the pH is increased from 4.0 to 7.0, showing the maximum efficiency at pH 7.0. As the pH was continuously increased to 10.0, there was already a decline in absorption capacity. At pH 7, the MP’s removal efficiency was 83 percent for agricultural soil and 82 percent for barren soil [3].

2.1.2 Effect of MP concentration

The capacity of soil to adsorb MP increased with increasing MP concentration for the agricultural soil sample. However, for the barren soil sample, a different behavior was observed. Table 2.1 shows the results of the adsorption percentage against MP concentration [3].

Table 2.1: Soil Adsorption with Changes in MP Concentration [3]

MP Concentration	Adsorption of Agricultural Soil	Adsorption of Barren Soil
10 µg/ L	72 %	78 %
30 µg/ L	78 %	66 %
50 µg/ L	83 %	80 %

2.1.3 Effect of adsorption time

MP adsorption of soils increased with elapsed time. It was observed that the adsorption rate was not constant throughout the cycle time. It was rapid from the start and attained an equilibrium, where it showed a constant rate. The initial faster rate could be explained by the presence of higher free binding sites at the start. Agricultural soil showed the fastest increase within the first six hours and gradually slowed down until it reached equilibrium after ten hours. For barren soil, the adsorption remained slower until equilibrium after ten hours also [3].

2.2 Ecotoxicity of Methyl Parathion

The research was carried out to quantify the effect of MP on humans and some animals. The following results were reported [10]:

- For humans, the minimum dosage that is considered lethal was 100 mg. The LD_{50} via contact dermatitis was 20 mg/ kg body weight, whereas it is 9 mg/ kg body weight if ingested orally. It was also found that once MP gets into the skin for at least fifteen minutes, it could no longer be washed away with soap and water.
- For rats, there was a difference between the male and female values. For male rats, the LD_{50} was 14 mg/ kg body weight while it was 24 mg/ kg body weight for female rats. When MP was given orally to the rats after six to eight minutes, it was already detected in the blood

plasma of rats. When taken intravenously, it was detected in the rats' brain tissues after 90 seconds.

- For chickens exposed to MP at a dosage of 14 percent of its LD_{50}, it was found to exist in the blood, kidney, liver, gastrointestinal tissues, and brain.
- The LD_{50} for bees was 0.17 mg.
- The LD_{50} for birds was between 3 to 8 mg/ kg body weight, depending on the species.
- For fish, the LD_{50} was 6 to 25 ppm.

2.3 Measures of the toxicity effect of pesticides on soil communities

When applied at the recommended dosage, pesticides have often transient effects on soil quality [11]. But prolonged use or at a higher dosage could result in a toxic soil environment, whether as low-level contamination or a high-level one, as in the case of accidental spills. There are now several ways to measure the degree of contamination of pesticides on soils. One method involves the examination of a bacterial population. As microorganisms, bacteria have a large surface area available for contact with other matters in the soil. This makes them ideal bioindicators to fluctuations in soil quality due to pesticide applications [11].

The study [11] discussed four approaches to quantify the effects of the bacterial exposure to pesticides:

2.3.1 Cultivation-dependent approaches

The cultivation-dependent method includes the mineralization and respiration tests, measuring the microbial biomass, use of standardized plates to evaluate growth and use of substrates, and physiological profiling at the community level. The microorganisms' metabolic activities can

also be measured through enzyme assays and carbon dioxide production levels. The research found that the effect of pesticides on soil bacteria is influenced by the agricultural practices in that area, the physicochemical characteristics of the soil, and the hydrophobic property of the pesticides [11].

2.3.2 Cultivation-independent approaches

The cultivation-independent method mostly uses DNA sequencing techniques. Another method involves the fatty acid evaluation which could give some key profiles of the bacterial population such as the population structure, responses to physiological stress, viable biomass, and nutritional status. Another method involves protein and metabolic analysis of the soil samples. Still, another method makes use of the polymerase chain reaction (PCR) which includes DNA fingerprinting using amplified ribosomal restriction analysis, length polymorphism of terminal restriction fragments, and denaturing gradient gel electrophoresis [11].

2.3.3 Metagenomically-derived bacterial diversity in polluted soil

Metagenomics has already been used to investigate various soil systems including sites with pesticides contamination. It has also been used to examine patterns of gene transcription that indicate how genes express their functional characteristics as a response to environmental stimulation. Some information could be obtained from metagenomics such as the metabolic pathways, phylogenic affiliation of selected genes, and their adaptive responses. An estimate of the bacterial biodiversity in soils revealed that one gram of soil contains from 10^4 to 10^7 phylotypes. The population becomes much lower in heavily contaminated areas [11].

2.3.4 Pollutant induced community tolerance (PICT) approach

Originally developed as a tool to study macro-organisms and algae, PICT has been used to detect

minor effects of contaminants in soil communities, providing a link to contaminant level and soil microbial population profile. The underlying principle for using PICT is that measuring the degree of bacterial tolerance to contaminants could indicate the degree of environmental disturbances [11].

2.4 Effect of MP on the soil bacterial respiration

A laboratory study was conducted to determine the effect of MP on the respiration of the soil bacteria. The concentrations of MP used were 1, 50, and 100 ppm. The amount of carbon dioxide (CO_2) generated was measured as the corresponding level of bacterial respiration. The 1 ppm culture showed the same trend as the control sample, i.e., increasing CO_2 level with each passing day, although comparatively lower in the measured value. The 50 and 100 ppm cultures showed decreasing trends for the first seven days, but gradually increased thereafter. It was observed that there was a remarkable increase in the soil microbial activity for all MP concentrations after 14 days. The explanation given for this was that during the first 14 days, the bacterial energy was used up in cell maintenance due to the disturbance experienced with the introduction of the MP solution. It was only when the environment stabilized after 14 days that the bacteria resumed their growth activities, hence the increase in CO_2 evolution. [12].

2.5 The transformation and biodegradation pathway of methyl parathion

2.5.1 In the air and water systems

Direct photolysis of MP occurs in the air resulting in the transformation of some portion of it into methyl paraoxon [13]. In natural water systems, MP degrades through hydrolysis. It was found that hydrolysis occurs faster in either alkaline (pH 8) or acidic (pH 5) environments compared to

neutral (pH 7) environments. The study showed the hydrolysis half-life of MP in freshwater to be 72 to 89 days at 25ºC at pH<8. At 40ºC and a pH>8, the half-life is much lower at 4 days. Table 2.2 shows the comparative half-lives of MP for four water samples and at 6ºC and 40ºC. It will be shown that the degradation of MP increases with temperature and pH [13].

Table 2.2: MP Degradation in water samples [13]

Water Sample	pH	MP Degradation, Days	
		@ 6 ºC	@ 22 ºC
Ultrapure Water	6.1	237	46
River Water	7.3	95	23
Filtered River Water	7.3	173	18
Sea Water	8.1	233	30

Some research conducted on the degradation products of MP in water systems revealed that the bacteria within the aquatic ecosystem, *aufwuchs communities* (microorganisms found in sediments, plants, and mats), were responsible for degrading the MP. Degradation products detected in water systems were 4-nitrophenol, dimethyl thiophosphoric acid, and some methyl aminoparathion [13].

2.5.2 In soil and sediments

Studies found some evidence that MP degrades faster in anaerobic soil (average half-life was 7 days) compared to aerobic soil (average half-life was 64 days). In an anaerobic sample from a flooded area, no degradation was observed at 6ºC, but at 25ºC, there was a formation of methyl aminoparathion. An additional 4-nitrophenol metabolite was observed at 35ºC. Generally, in

moist soil, MP degrades by undergoing reduction of the nitro group, producing 4-nitrophenol and 4-aminophenol as metabolites. These metabolites underwent a further transformation into CO_2 by the soil microbiota. The MP biodegradation is highly dependent on soil pH, temperature, and moisture [13].

2.6 Metagenomics of Topsoil Microbiome

Bahram et al conducted a study to determine the relationship between phylogenetic, genetic, and taxonomic abundance and diversity in 189 samples of topsoil representing the earth's terrestrial regions and biomes. They used soil chemistry, metagenomics, biomass assessment, and DNA metabarcoding in their study. Metagenomics was used to construct the soils' catalogue of genes, identifying around 160 million unique genes dominated by bacteria and fungi [41].

The results revealed that the soil microbial compositions were determined by niche differentiation and environmental filtering. It was found that the two factors that affected most the global distribution of soil fungi and bacteria were soil pH and degree of precipitation in such a way that the conditions favor either a bacteria- or fungi-driven cycling. Further, it was also found that the presence of *antibiotic resistant genes* (ARGs) affect the structuring of the microbial population. Lastly, the researchers found supporting evidence that the global climate change could affect the microbial community composition and functional characteristics [41]

III. MATERIALS AND METHODS

3.1 Site description and sample collection

Three soil samples were collected from three separate vegetable farms: 1) *Hope Farms* in 10401 Scott Street, Houston, Texas (29.646914° N, 95.3701409° W); 2) *Plant It Forward Farm* in 4030 Willowbend Boulevard, Westwood Park, Houston, Texas (29.6650135° N, 95.4433233° W), and 3) *Atkinson Farms* in 3217 Spring Cypress Road, Spring, Texas (30.059920°N, 95.465012°W). To obtain good representative samples, the collection was done by digging holes from three spots on each farm. The samples were collected at a depth of 8 to 12 cm and placed in labeled amber bottles to prevent degradation due to light exposure. The samples were then transported and stored at room temperature in their amber containers.

3.2 Rationale for collection of soil samples from different locations

Microbial communities found in farm soils are greatly diverse. They play important roles in maintaining the ecological cycles and helping plants be productive. But extensive agricultural practices (e.g. pesticide application) invariably alter the soil microbial population [14] to such an extent that they greatly decrease, reducing the capacity of the soil to degrade the chemicals, resulting in *contaminated soil*. Thus, it is increasingly becoming important to evaluate the microbial diversity of different soil samples to understand how a commonly applied pesticide - methyl parathion - affect the different microbial composition of each soil sample before and after the addition of MP.

3.3 Materials used

First, the Carbon Selective Media (CSM) was prepared by adding 945 mL of Milli-Q H_2O, 2 mL of 1 M Nitrilotriacetic acid (disodium salt), and 1 mL of 20% (w/v) of $MgSO_4.7H_2O$. The solution was then autoclaved and cooled to room temperature under the biosafety cabinet. The

following were added later: 1 mL of 4% (w/v) $Ca(NO_3)_2.4H_2O$, 1 mL of 0.5% (w/v) $FeSO_4.6H_2O$, and 50 mL of sterile phosphate buffer. Methyl parathion, with a concentration of 20 mg/ml, was sourced from Sigma-Aldrich (USA).

3.4 Experimental Method

The collected soil samples were homogenized at the laboratory. The amber bottles containing the samples were shaken vigorously to homogenize the contents. One gram each of the three farmland samples were taken from each bottle into a weigh boat. The three one-gram samples were mixed thoroughly in a weigh boat until homogenized, and placed in a falcon tube. Then, the prepared culture was incubated at 30^{o}C and 200 rpm for one week. This was the sample without MP. The procedure was repeated to prepare a second homogenized sample.

For the amended sample, the following were mixed in a 50-mL centrifuge tube: the second homogenized sample above + 5 mL of the prepared CSM + 5 μL of 20 mg/mL methyl parathion. Next, the prepared culture was incubated at 30^{o}C and 200 rpm for twenty days. This was the sample with MP.

3.5 Metagenomics Method

The two samples were sent for whole genome sequencing at Genewiz (South Plainsfield, NJ). Genewiz conducted all related procedures such as DNA isolation, library construction, and paired-end shotgun metagenomic sequencing of the samples. A PureLink Genomic DNA extraction kit was used for Illumina shotgun metagenomic sequencing. Quantification was arrived through a Nanodrop spectrophotometer and Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA). 50 ng of the farmland soil sample was taken for quality control purposes and ran out on a 0.6% w/v agarose gel. The associated protocols and reagents prescribed by the manufacturer (Illumina, San Diego, CA) were used throughout the sequencing process.

The Nextera DNA Flex Library prep kit was used to construct the DNA libraries, and analysis was done on the Agilent Tape Station (Agilent Technologies, Palo Alto, CA). Quantification was carried out through real time PCR (Applied Biosystems, Carlsbad, CA). Next, multiplexing was carried out on the DNA libraries which were loaded onto an Illumina Miseq instrument (Illumina, San Diego, CA).

Sequence annotation was done through the *metagenome rapid annotation using subsystems technology* server (MG RAST, Version 4.0.3) which is "an open-source system based on the SEED framework for comparative genomics" [16]. The raw sequence data were submitted to the MG RAST server in both forward and reverse *fasta* format. The system then processed and normalized the sequence and generated automatic summaries. The system allows each user to set and alter the parameters [16].

MG RAST uses the following databases in its sequence annotation: COG, eggNOG, and KO. COG stands for *clusters of orthologous groups* which conduct phylogenic categorization of proteins encoded in whole genomes. The 2019 release consists of 4,877 COGs that include complete genomes of 122 archaea and 1,187 bacteria [18].

The *evolutionary genealogy of genes: non-supervised orthologous groups* (eggNOG) contains resources for simultaneous analysis of thousands of genomes to establish orthology relationships among all their genes. The current eggNOG release is based on 2,502 viruses and 5.090 organisms.

The *Kyoto Encyclopedia of Genes and Genomes* (KEGG) *Orthology* (KO) is an integrated resource consisting of sixteen databases that are categorized into genomic information, health information, chemical information, and systems information. It facilitates understanding of

utilities and high-level functions "of the biological system such as the cell, the organisms, and the ecosystem, from genomic and molecular-level information [20].

Next, the annotated sequence was loaded onto NCBI's *basic local alignment search tool* (BLAST) which locates locally similar regions among the sequences. The system works by comparing sequences of protein nucleotides to database sequences, then computes the statistical significance of identified matches. BLAST is commonly used to identify evolving and functional relationships among sequences and also to determine members of gene families [17].

IV. RESULTS AND DISCUSSION

4.1 Source hits distribution

Figure 4.1 reflects the number of hits in the examined set of data for soil without MP that were annotated by various listed databases which include protein databases, rRNA databases, and protein databases with information on functional hierarchy. The colored bars represent the annotated reads by selected ranges of e-values. It will be seen that the different databases did not have the same number of hits and could also have different kinds of annotation data.

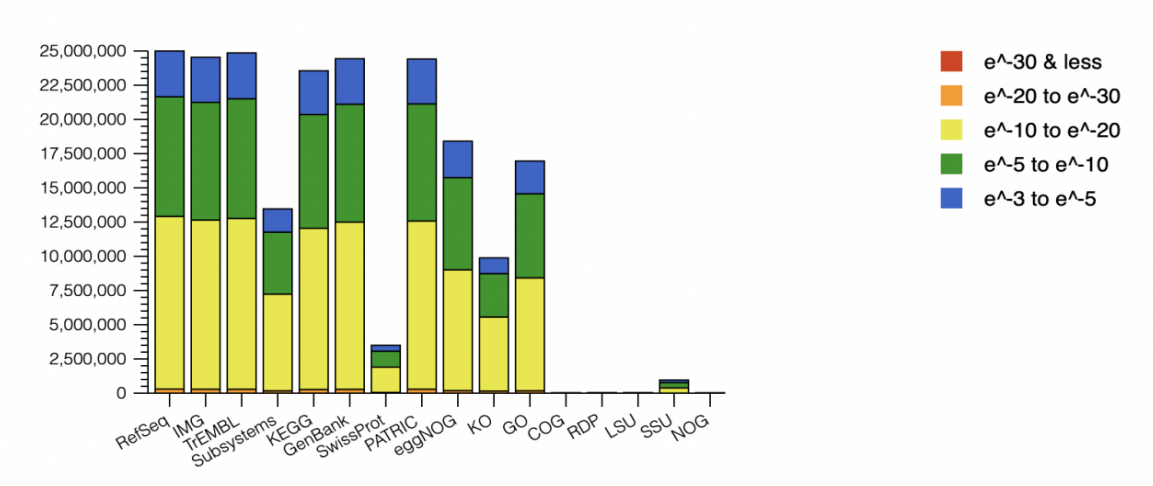


Figure 4.1: Source hits distribution for soil sample without methyl parathion

4.2 Rarefaction curve

Figure 4.2 depicts the rarefaction curve of the richness of the annotated species. It shows the plot of the total number of annotated distinct species against the number of sampled sequences. The steep curve on the left means that there are a large fraction of diverse species that has to be discovered. The flatter curve on the right means that a good number of individual species is sampled and further sampling would probably yield a few additional species. This rarefaction curve was derived from the table of abundant species.

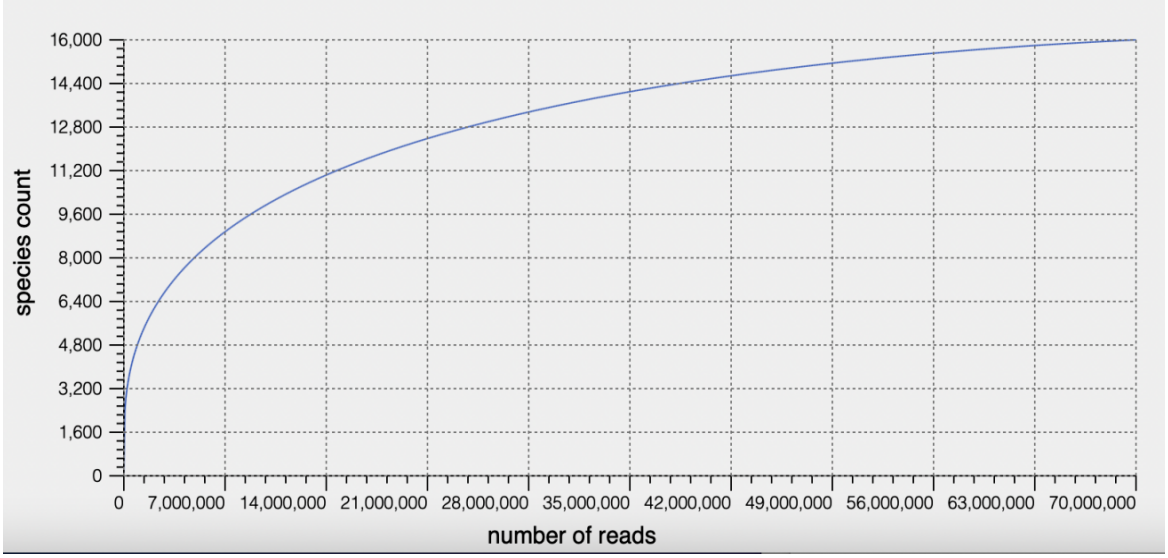


Figure 4.2: Rarefaction curve for soil sample without methyl parathion

4.3) Taxonomic Hits Distribution

Figure 4.3 represents the taxa distribution using a *contigLCA* algorithm in locating a single consensus taxonomic entity for all features on every individual sequence. Taxonomy and protein content classification was conducted using the default mode of Diamond program version 0.9.29 [34]. It was evident that the majority (98.82% and 98.74%) of the microorganisms were bacteria

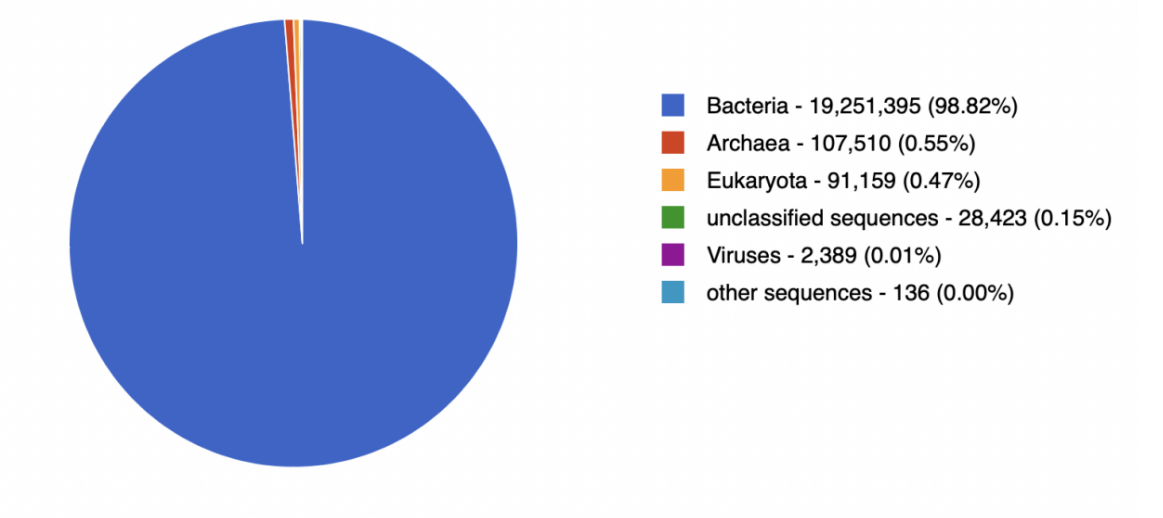


Figure 4.3: Taxonomic hits distribution for soil sample without methyl parathion

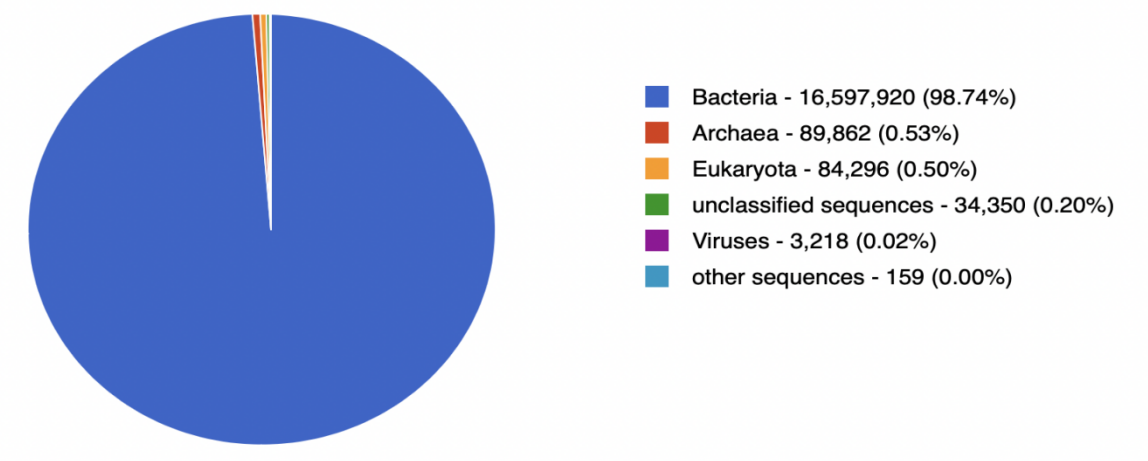


Figure 4.4: Taxonomic hits distribution for soil sample with methyl parathion

Table 4.1: Taxonomic hits distribution, comparison of with and without MP

Taxa	Without MP	With MP	Variance	Remarks
Bacteria	19,251,395	16,597,920	-2,653,475	Decreased by 13.8%
Archaea	107,510	89,862	-17,648	Decreased by 16.4%
Eukaryota	91,159	84,296	-6,863	Decreased by 7.5%
Viruses	2,389	3,218	829	Increased by 34.7%
Unclassified	28,423	34,350	5,927	Increased by 20.9%
Others	136	159	23	Increased by 16.9%
TOTAL	**19,481,012**	**16,809,805**	**-2,671,207**	**Decreased by 13.7%**

As shown in Table 4.1, there was a decrease in the sequences of bacteria, archaea, and eukaryota; while the sequences of viruses, unclassified, and others showed an increase. The overall effect of methyl parathion in farm soil was that it caused a decrease in identified sequences by 13.7%.

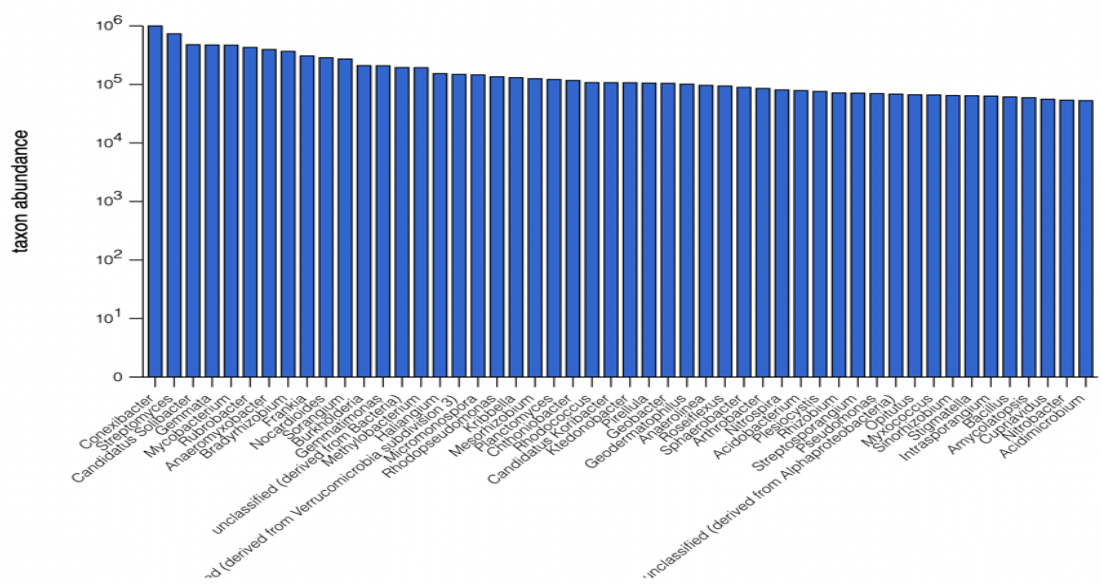


Figure 4.5 Taxonomic abundance for soil sample without methyl parathion

Figure 4.5 depicted the taxonomic abundance of the soil sample for the fifty most abundant genera. A taxonomic abundance profile provides an information of the different species that will most likely be observed in succeeding analyses [40]. It will be seen that the top two most abundant genera are *conexibacter* and *streptomyces* which are both nitrogen-fixating bacteria and could be used for soil bioremediation. This suggests that the unamended farmland soil samples already contain microorganisms capable of natural degradation. The effect of MP on such microorganisms will be the subject of this study.

4.4 MG RAST Results

4.4.1 MG RAST Analysis Statistics

Using the Illumina sequencing method, the initial count for the soil sample without MP was 9,658,805,600 base pair (bp) yielding a sequence count of 63,965,600 and a mean sequence

length of 151 bp. After quality control, the count was reduced to 8,975,952,926 bp, which was only 92.9% of the initial count. The post-QC sequence count was also reduced to 59,662,581, which was 93.3% of the initial sequence. The post-QC mean sequence length was reduced a little at 150 $\pm$ 3 bp. The identified protein features were counted at 20,815,709 which was only 36.6% of the predicted protein features. The identified ribosomal ribonucleic acid (rRNA) was 22,939 which was only 34.7% of the predicted rRNA. The reason for the overall reduction in counts was that during QC, MG RAST provided screening which included filtering, removal of duplicates, length-based read trimming, and quality-based read trimming.

4.4.2) For sample without MP

A) COG - served as the platform for functional annotation of newly-sequenced genomes.

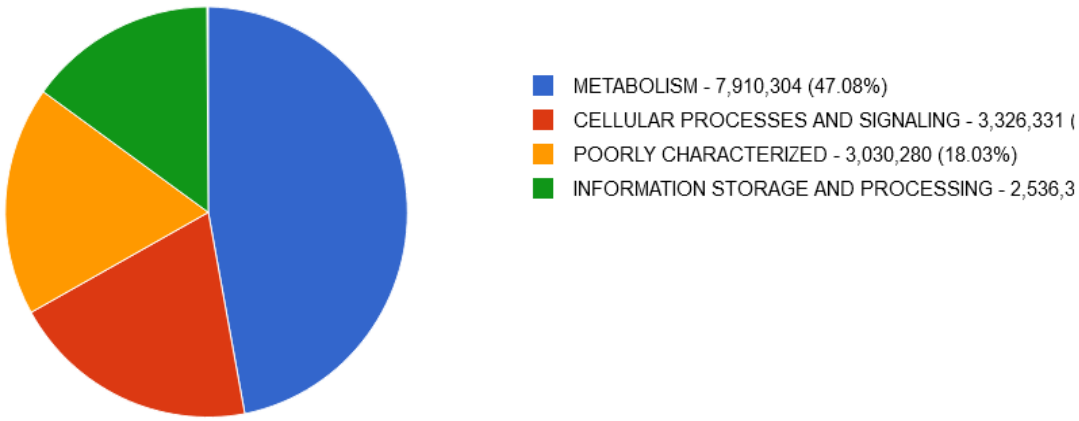


Figure 4.6: COG-based Profile of Soil Sample without Methyl Parathion

The COG database sequenced 7,910,304 genomes for metabolism functioning, 2,536,300 for information storage and processing, and 3,326,331 for cellular processes and signaling.

B) NOG

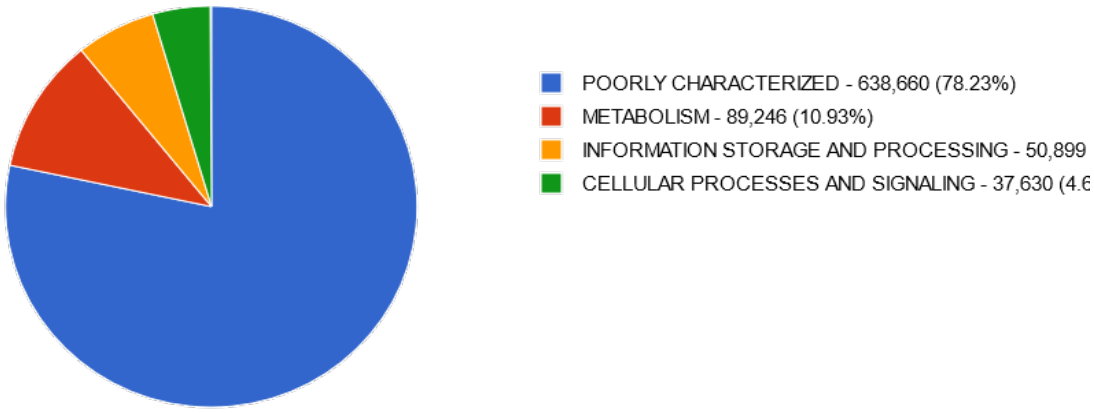


Figure 4.7: NOG-based Profile of Soil Sample without Methyl Parathion

The eggNOG database sequenced 89,246 genomes for metabolism, 50,899 for information storage and processing, and 37,630 for cellular processes and signaling.

C) KO

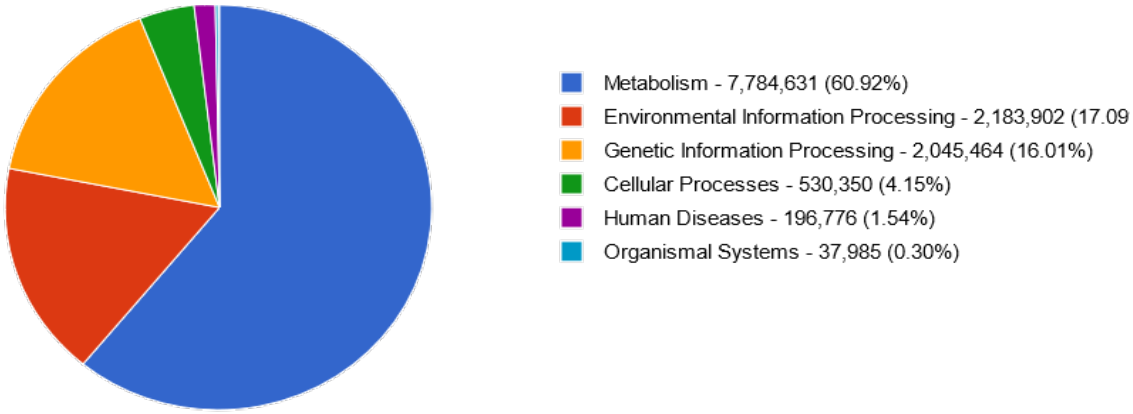


Figure 4.8: KO-based Profile of Soil Sample without Methyl Parathion

The KO database sequenced 7,784,631 genomes for metabolism functioning, 2,045,464 for genetic information processing, 2,183,902 for environmental information processing, and 530,350 for cellular processes, 196,776 for human diseases, and 37,985 for organismal systems.

D) Phyla

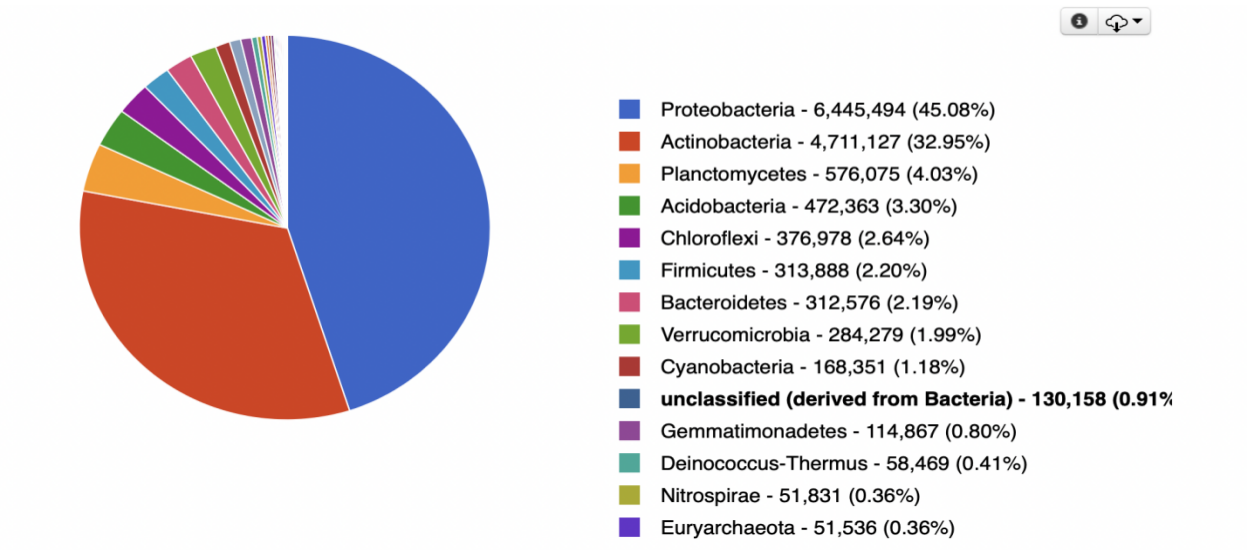


Figure 4.9: Top Bacterial Phyla Identified in Soil Sample without Methyl Parathion

The most abundant phylum of bacteria identified in the unamended farm soil was *proteobacteria* (45.08%) which consist of several groups of aerobic bacteria that have good degradation abilities. They are known to initiate nitrogen fixation in some plants, oxidizing ammonium into nitrite radicals. They are currently used in the bioremediation of industrial wastes, increasing the nitrogen availability to plants and at the same time limiting the fixation of carbon dioxide [27]. The second most abundant phylum of bacteria was *actinobacteria* (32.95%) which are saprophytes that are capable to produce several enzymes that can aid in the degradation of plant and animal polymers such as cellulose, lignin, chitin, and others. In this way, they play important

parts in soil development and biogeochemical cycling [28].

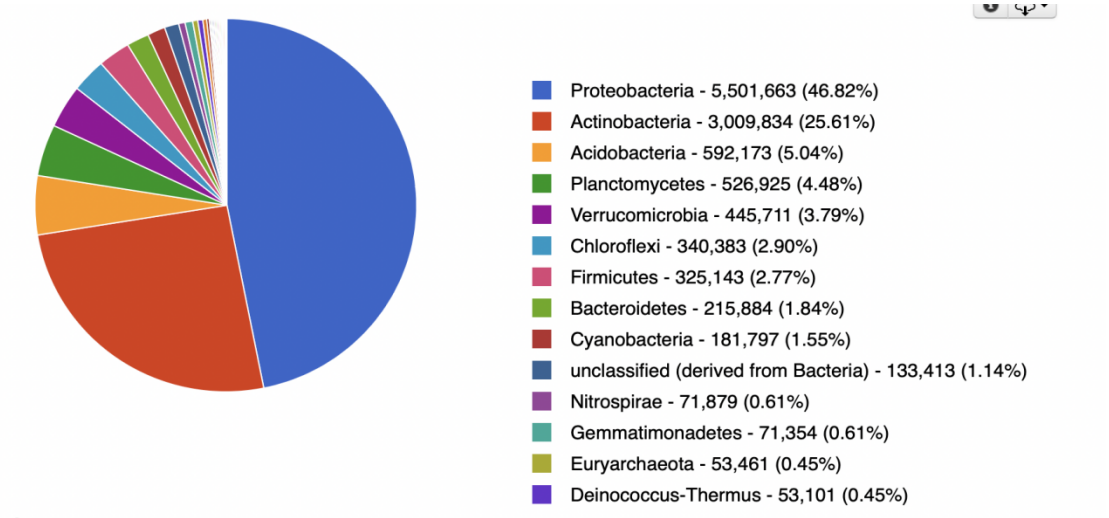


Figure 4.10: Top Bacterial Phyla Identified in Soil Sample with Methyl Parathion

Table 4.2 summarizes the phyla sequences of farm soil with and without MP. There was a significant decrease in the top two bacteria, *proteobacteria* (14.6%) and *actinobacteria* (36.1%). Other phyla that showed a decrease in population were *planctomycetes, chloroflexi, bacteroidetes, gemmatimonadetes,* and *deinococcus-thermus.* The rest showed an increase in population, most notable of which are the *acidobacteria* (25.4%), *verrucomicrobia* (56.8%), and *nitrospirae* (38.7%). The effect of MP on bacterial phyla is an overall decrease in the bacterial population by 18.1%.

The research of Kumar et al [31] mentioned the general degradation effect on pesticides by *actinobacteria, bacteroidetes, firmicutes, proteobacteria,* and *cyanobacteria*. In this study, it appears that the *firmicutes* and *cyanobacteria* have the desired degradation effect on MP, as indicated by the increased population upon the inoculation of the soil with MP. It is assumed that an increase in bacterial population after the application of MP on the soil suggests that the specific bacteria fed on the MP, contributing to its degradation

Table 4.2: Bacterial Phyla, comparison of with and without MP

Taxa	Without MP	With MP	Variance	Remarks
Proteobacteria	6,445,494	5,501,663	-943,831	Decreased by 14.6%
Actinobacteria	4,711,127	3,009,834	-1,701,293	Decreased by 36.1%
Planctomycetes	576,075	526,925	-49,150	Decreased by 8.5%
Acidobacteria	472,363	592,173	119,810	Increased by 25.4%
Chloroflexi	376,978	340,383	-36,595	Decreased by 9.7%
Firmicutes	313,888	325,143	11,255	Increased by 3.6%
Bacteroidetes	312,576	215,884	-96,692	Decreased by 30.9%
Verrucomicrobia	284,279	445,711	161,432	Increased by 56.8%
Cyanobacteria	168,351	181,797	13,446	Increased by 8.0%
Gemmatimonadetes	114,867	71,354	-43,513	Decreased by 37.9%
Deinococcus-Thermus	58,469	53,101	-5,368	Decreased by 9.2%
Nitrospirae	51,831	71,879	20,048	Increased by 38.7%
Euryarchaeota	51,536	53,461	1,925	Increased by 3.7%
Unclassified	130,158	133,413	3,255	Increased by 2.5%
TOTAL	**14,067,992**	**11,522,721**	**-2,545,271**	**Decreased by 18.1%**

E) Class

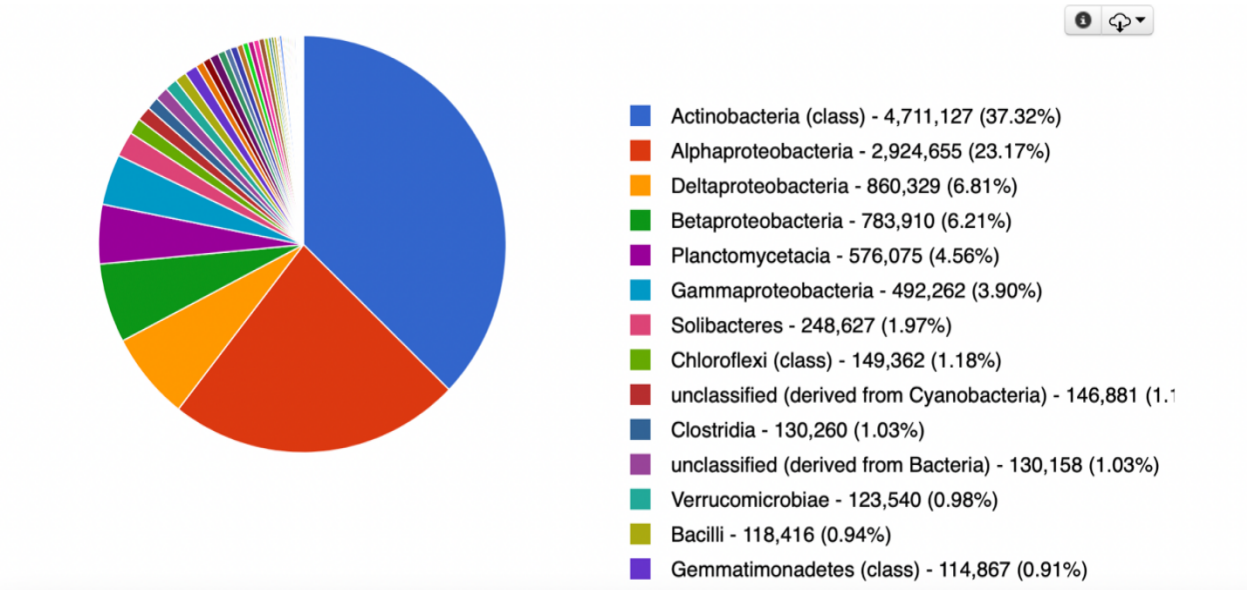


Figure 4.11: Top Bacterial Classes Identified in Soil Sample without Methyl Parathion

The most abundant class of bacteria identified were *actinobacteria* (37.32%) which possess many desirable properties in soil bioremediation. They play important roles in organic matter cycling by decomposing complex polymers in dead animals and plants resulting in the production of enzymes that are beneficial to crops. They also inhibit the growth of some plant pathogens in the rhizosphere. Studies have confirmed that they could improve the availability of minerals and nutrients in the soil, as well as enhance the production of metabolites, thereby promoting plant growth regulators [25].

The second most abundant class of bacteria identified was *alphaproteobacteria* (23.17%) which is one of the proteobacteria that are capable of surviving in an environment with low nutrients such as in deep soils. Some genera are known to initiate nitrogen fixation in some plants [27].

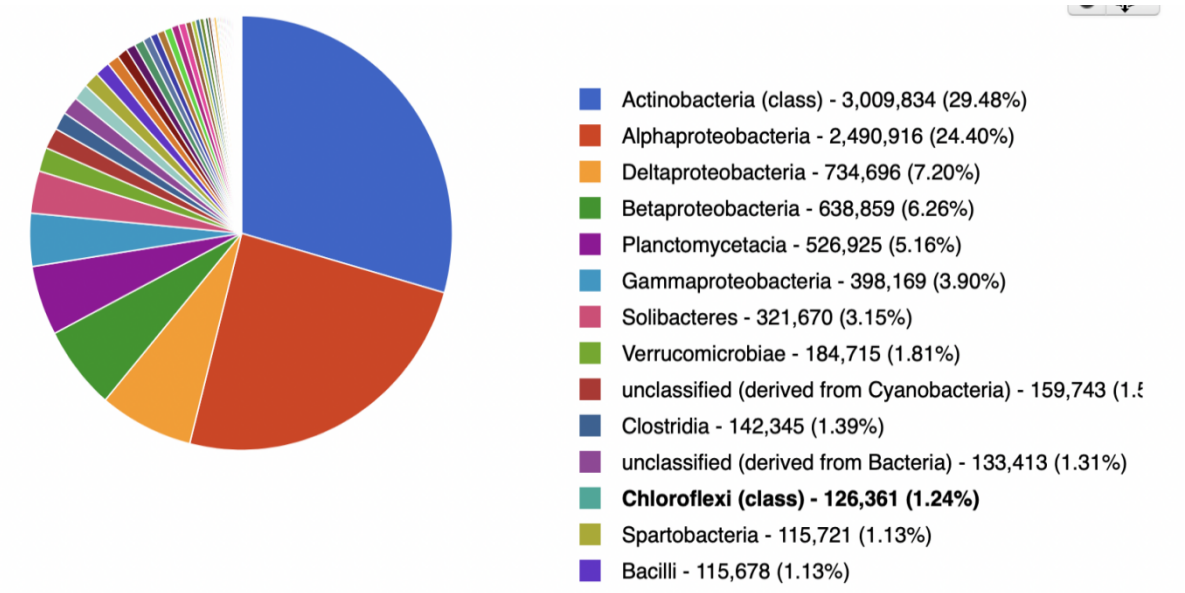


Figure 4.12: Top Bacterial Classes Identified in Soil Sample with Methyl Parathion

Table 4.3: Bacterial Classes, comparison of with and without MP

Class	Without MP	With MP	Variance	Remarks
Actinobacteria	4,711,127	3,009,834	-1,701,293	Decreased by 36.1%
Alphaproteobacteria	2,924,655	2,490,916	-433,739	Decreased by 14.8%
Deltaproteobacteria	860,329	734,696	-125,633	Decreased by 14.6%
Betaproteobacteria	783,910	638,859	-145,051	Decreased by 18.5%
Planctomycetacia	576,075	526,925	-49,150	Decreased by 8.5%
Gammaproteobacteria	492,262	398,169	-94,093	Decreased by 19.1%
Solibacteres	248,627	321,670	73,044	Increased by 29.4%
Chloroflexi	149,362	126,361	-23,001	Decreased by 15.4%
Clostridia	130,260	142,345	12,085	Increased by 9.3%
Verrucomicrobiae	123,540	184,715	61,175	Increased by 49.5%
Bacilli	118,416	115,678	-2,738	Decreased by 2.3%
Gemmatimonadetes	114,867	-	-	Wiped out
Spartobacteria	-	115,721	-	Flourished
Unclassified	277,039	293,156	16,117	Increased by 5.8%
TOTAL	**11,510,469**	**9,099,045**	**-2,411,424**	**Decreased by 20.9%**

The introduction of MP to the soil caused the overall bacterial classes to decrease by 20.9%. All variants of proteobacteria (alpha-, beta-, delta-, and gamma-) decreased by a range of 14.6% to 19.1%. The other classes that experienced decrease in population are *actinobacteria* (36.1%), *planctomycetacia* (8.5%), *chloroflexi* (15.4%), *gemmatimonadetes* (100%), and *bacilli* (2.3%). Those that increased are *solibacteres* (29.4%), *clostridia* (9.3%), *spartobacteria* (100%) and *verrucomicrobiae* (49.5%).

F) Order

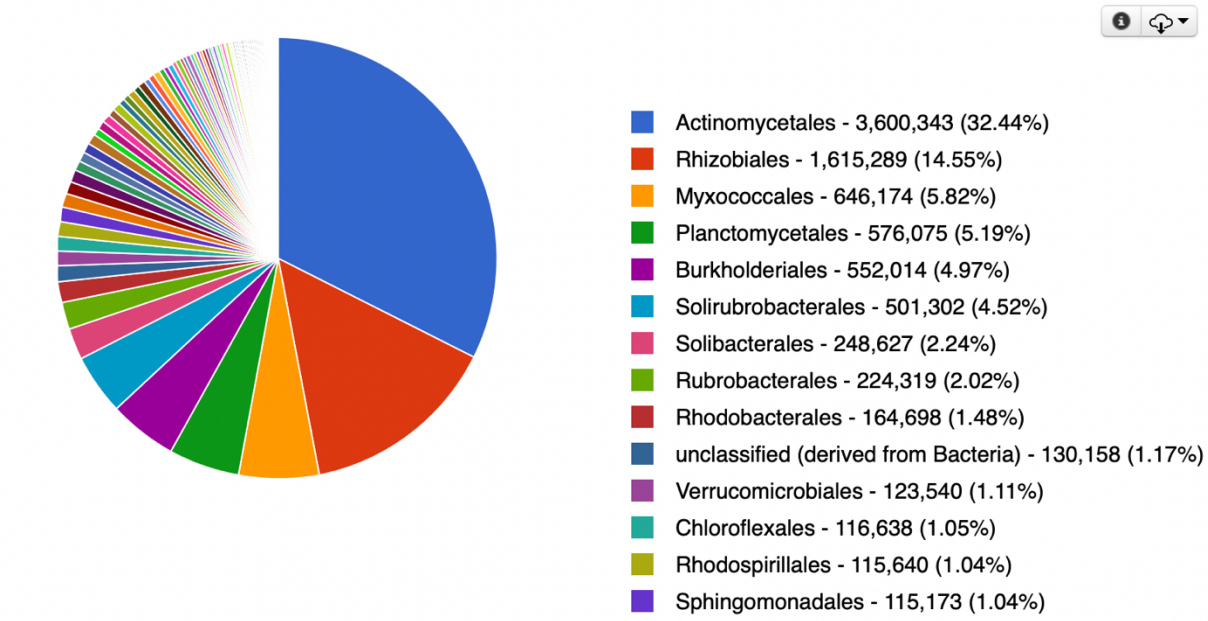


Figure 4.13: Top Bacterial Order Identified in Soil Sample without Methyl Parathion

The most abundant order of bacteria identified were *Actinomycetales* (32.44%) which possess many desirable properties in soil bioremediation. They play important roles in organic matter cycling by decomposing complex polymers in dead animals and plants resulting in the production of enzymes that are beneficial to crops. They also inhibit the growth of some plant pathogens in the rhizosphere. Studies have confirmed that they could improve the availability of

minerals and nutrients in the soil, as well as enhance the production of metabolites, thereby promoting plant growth regulators [25].

The next most abundant order of bacteria was the *rhizobiales* which have gained renown as plant growth promoters. They also can degrade organic pollutants and have shown resistance to heavy metals. Some species invade the roots of leguminous plants and form nodules that convert nitrogen from the air into ammonia, which is then given to the host plant, allowing the growth of plants even in the absence of oxygen [26].

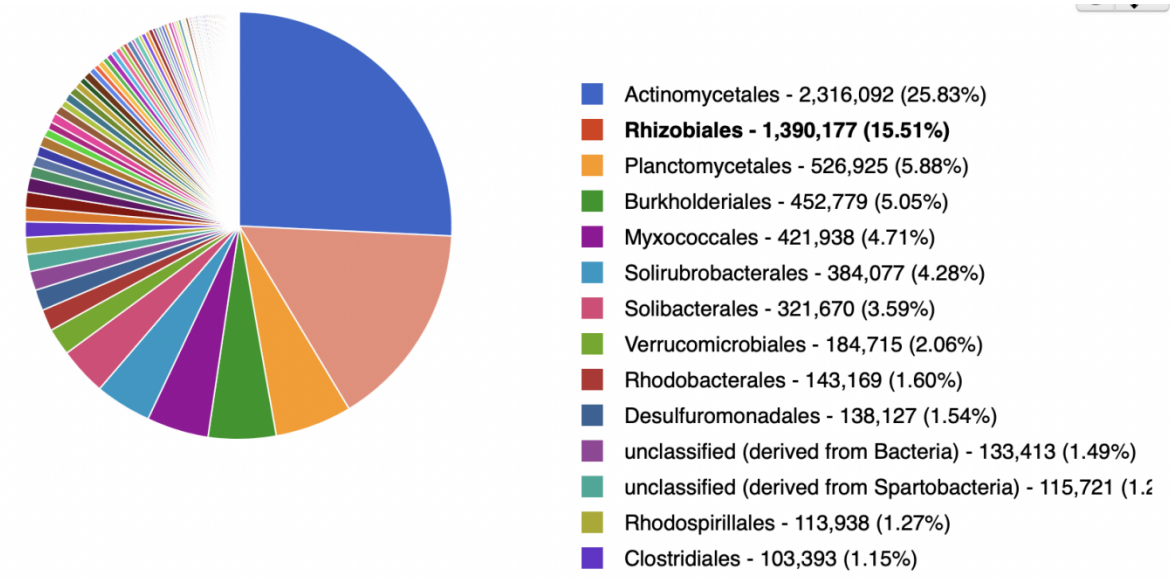


Figure 4.14: Top Bacterial Order Identified in Soil Sample with Methyl Parathion

Table 4.4: Bacterial Order, comparison of with and without MP

Order	Without MP	With MP	Variance	Remarks
Actinomycetales	3,600,343	2,316,092	-1,284,251	Decreased by 35.7%
Rhizobiales	1,615,289	1,390,177	-225,122	Decreased by 13.9%
Myxococcales	646,174	421,938	-224,236	Decreased by 34.7%
Planctomycetales	576,075	526,925	-49,150	Decreased by 8.5%
Burkholderiales	552,014	452,779	-99,235	Decreased by 18.0%
Solirubrobacterales	501,302	384,077	-117,225	Decreased by 23.4%
Rubrobacterales	224,319	-	-	Wiped out
Rhodobacterales	164,698	143,169	-21,529	Decreased by 13.1%
Verrucomicrobiales	123,540	184,715	61,175	Increased by 49.5%
Chloroflexales	116,638	-	-	Wiped out
Rhodospirillales	115,640	113,938	-1,702	Decreased by 1.5%
Sphingomonadales	115,173	-	-	Wiped out
Desulfuromonadales	-	138,127	-	Flourished
Unclassified	130,158	249,134	118,976	Increased by 91.4%
TOTAL	**8,481,363**	**6,321,071**	**-2,160,292**	**Decreased by 25.5%**

The application of MP on farm soil caused the general decrease of the bacterial order by 25.5%. Only two bacterial orders increased in number, *verrucomicrobiales* (49.5%) and *desulfuromonadales* (100%). There was not many studies linking these two bacteria to MP degradation.

G) Family

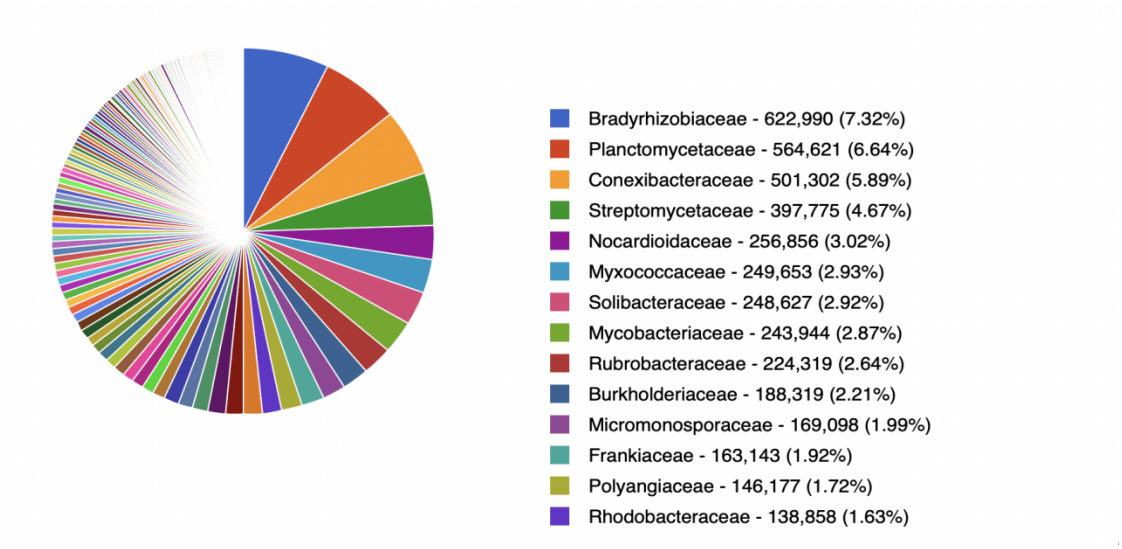


Figure 4.15: Top Bacterial Family Identified in Soil Sample without Methyl Parathion

The top bacteria family found in the original soil samples was *bradyrhizobiaceae* (7.32%) which possesses the capability to utilize different sources of nitrogen in its metabolism. This symbiotic fixation of nitrogen is considered one of the vital aspects of the sustainability of the agroecosystem. It was also found to have a role in the mitigation of greenhouse gases (GHG) emissions since it could be used in place of nitrogen fertilizers that are normally produced using fossil fuels [23].

The second top family of bacteria found was *planctomycetaceae* (6.64%) which are mostly anaerobic and can convert ammonia to dinitrogen without the presence of oxygen. This family of bacteria plays an important part in the global nitrogen cycle and is currently being considered as a tool for remediation of nitrogen-rich wastes [24].

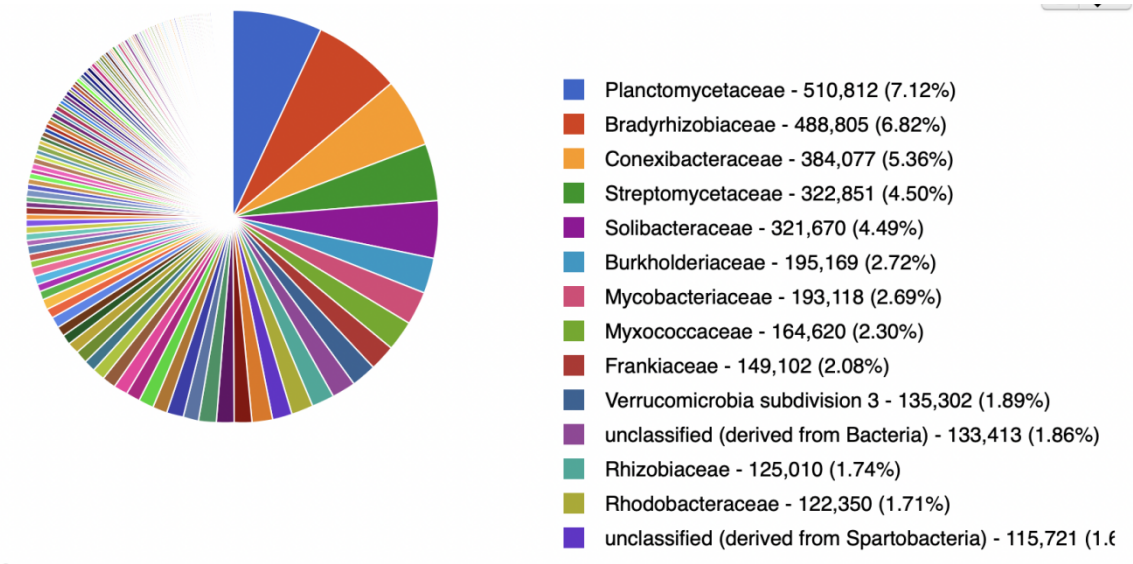


Figure 4.16: Top Bacterial Family Identified in Soil Sample with Methyl Parathion

Table 4.5: Bacterial Family, comparison of with and without MP

Family	Without MP	With MP	Variance	Remarks
Bradyrhizobiaceae	622,990	488,805	-134,185	Decreased by 21.5%
Planctomycetaceae	564,621	510,812	-53,809	Decreased by 9.5%
Conexibacteraceae	501,302	488,805	-12,497	Decreased by 2.5%
Streptomycetaceae	397,775	322,851	-74,924	Decreased by 18.8%
Nocardioidaceae	256,856	0		Wiped out
Myxococcaceae	249,653	164,620	-85,033	Decreased by 34.1%
Solibacteraceae	248,627	321,670	73,043	Increased by 29.4%
Mycobacteriaceae	243,944	193,118	-50,826	Decreased by 20.8%
Rubrobacteraceae	224,319	-		Wiped out
Burkholderiaceae	188,319	195,169	6,850	Increased by 3.6%
Micromonosporaceae	169,098	-		Wiped out
Frankiaceae	163,143	149,102	-14,041	Decreased by 8.6%
Polyangiaceae	146,177	-	-	Wiped out
Rhodobacteraceae	138,858	122,350	16,508	Decreased by 11.9%
Verrucomicrobia Sub3	-	135,302		Flourished
Rhizobiaceae	-	125,010		Flourished
Unclassified	-	249,134		Flourished
TOTAL	**4,115,682**	**3,466,748**	**-648,934**	**Decreased by 15.8%**

The application of MP on farm soil caused a decrease in the total microbial family population by 15.8%. Those that registered an increase were *solibacteraceae* (29.4%), *burkholderiaceae* (3.6%), *verrucomicrobia subdivision 3* (100%), and *rhizobiaceae* (100%). Other studies

confirmed the MP degradation capacity of *burkholderiaceae*, a betaproteobacterium, and *rhizobiaceae,* an alphaproteobacterium [15, 26].

H) Genus

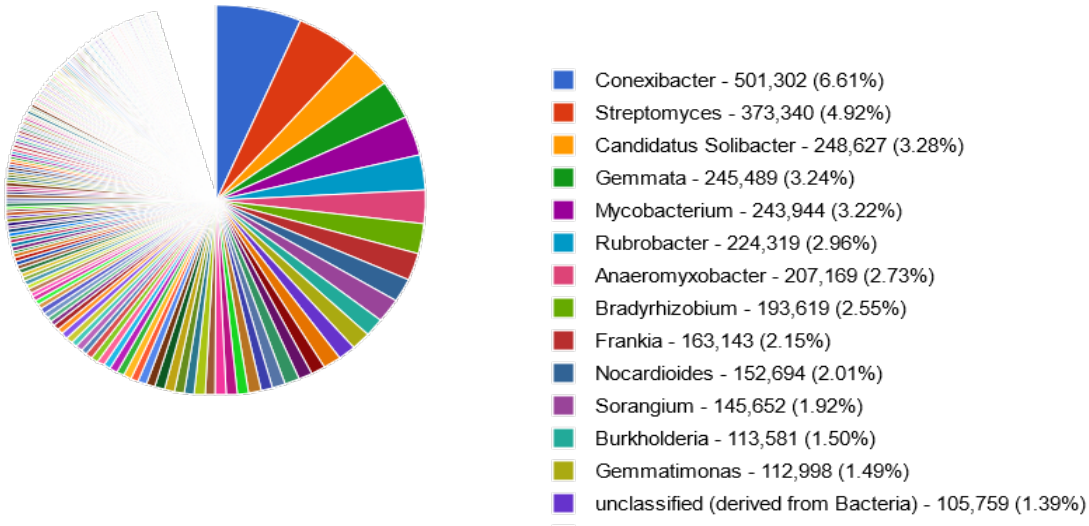


Figure 4.17: Top Genera Identified in Soil Sample without Methyl Parathion

The top genus identified was *conexibacter* (6.61%) which is an aerobic microbe and can reduce nitrate to nitrite, even with limited oxygen. It is being considered for bio-remediation due to its nitrate-reducing ability. It also plays a role in nitrogen and carbon cycling in soils [21].

The next top genus was *streptomyces* (4.92%) which was known to be direct promoters of plant growth by producing phytohormones, fixating nitrogen, scavenging of ferric iron from the soil, and producing *l-aminocyclopropane-l-carboxylate* (ACC) deaminase activity that suppresses stress in plants [22].

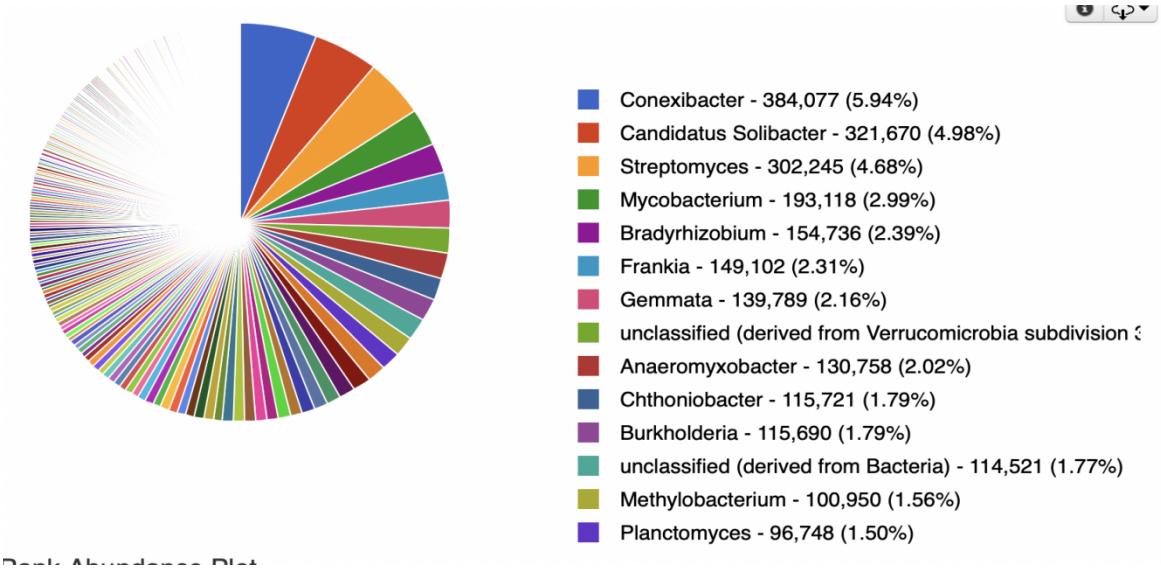


Figure 4.18: Top Genera Identified in Soil Sample with Methyl Parathion

Table 4.6: Bacterial Genera, comparison of with and without MP

Genera	Without MP	With MP	Variance	Remarks
Conexibacter	502,302	384,077	-118,225	Decreased by 23.5%
Streptomyces	373,340	302,245	-71,095	Decreased by 19.0%
Candidatus Solibacter	248,627	321,670	73,043	Increased by 29.4%
Gemmata	245,489	139,789	-105,700	Decreased by 43.1%
Mycobacterium	243,944	193,118	-50,826	Decreased by 20.8%
Rubrobacter	224,319	-		Wiped out
Anaeromyxobacter	207,169	130,758	-76,411	Decreased by 36.9%
Bradyrhizobium	193,619	154,736	-38,883	Decreased by 20.1%
Frankia	163,143	149,102	-14,041	Decreased by 8.6%
Nocardoides	152,694	-		Wiped out
Sorangium	145,652	-		Wiped out
Burkholderia	113,581	115,690	2,109	Increased by 1.9%
Gemmatimonas	112,998	-	-	Wiped out
Chthoniobacter	-	115,721		Flourished
Methylobacterium	-	100,950		Flourished
Planctomyces	-	96,748		Flourished
Unclassified	105,759	114,521		Increased by 8.3%
TOTAL	**3,032,636**	**2,319,125**	**-713,511**	**Decreased by 23.5%**

MP caused the reduction of the total bacterial genera by 23.5%. Those that registered an increase in population were *Candidatus solibacter* (19.0%), *Burkholderia* (1.9%), *Chthnoniobacter* (100%), *Methylobacterium* (100%), and *Planctomyces* (100%). *Burkholderia*, a betaproteobacterium, and *Methylobacterium*, an alphaproteobacterium, were previously associated with bioremediation of OPs [15, 31].

I) The MPD biomarker genes

Table 4.7 provides the identity of the MPD genes found in the unamended and amended farmland soil sample.

Table 4.7: The MPD biomarker genes

Category	MPD Gene	WGM-05 R1-001 Unamended # of Hits	WGM-06 R2-002 Amended # of Hits	Representative Species
MP	MBL Fold Hydrolase Metallo hydrolase	70	20	*Rhodanobacter sp. SCN*
MP	Methyl Parathion Hydrolase (MPH)	10	0	*Sphingomonas Sp DSP-2*
PNP	**1,2,4-trihydroxybenzene 1,2-dioxygenase benzoquinone reductase**	13	2	*Burkholderia sp. AK-5*
PNP	Hydroquinol 1,2-dioxygenase	47	10	*Rhizobium sp. MTP-10005*

Out of the four biomarkers identified, three were the most probable candidates as the MPD gene for the farmland soil from three locations in Texas. Although the count of the three biomarkers decreased in the amended soil sample, it could be explained as being due to the stress they experienced upon the introduction of the pesticide. The study conducted by Bindhya et al [12] revealed that upon introduction of MP to the soil sample, the organisms decreased in microbial activity at first, since they used their bacterial energy for cell maintenance, then after fourteen days, the microbial activity again began to increase, implying that they have resumed their normal activities, including propagation [12].

The first putative species is *Burkholderia sp. AK-5* which is a betaproteobacterium. It is known to degrade aromatic compounds like 4-aminophenol by utilizing it as the only source of nitrogen, carbon, and energy [37]. Since MP has no 4-aminophenol byproduct in its degradation pathway, Burkholderia sp. AK-5 will not be considered as an MPD enzyme for this study.

The second species is *Rhizobium sp. MTP-10005* which is an alphaproteobacterium [35] enzyme having a polypeptide chain composed of 351 amino acid residues, and a mass of 36,405. Rhizobium sp. MTP-10005 is a bioremediation organism capable of degrading β-resorcylate and γ-resorcylate via the methylacetate and hydroxyquinol pathway [36]. Figure 4.19 and 4.20 showed the presence of methylacetate as an intermediate product in the MP degradation via the HQ pathway. Thus, Rhizobium sp. MTP-10005 will be considered as a putative MPD species for this study.

The third species is *Rhodanobacter sp. SCN* which is a gammaproteobacterium. Little is known of this species in terms of degradation capability, so it will not be considered as an MPD gene for this study].

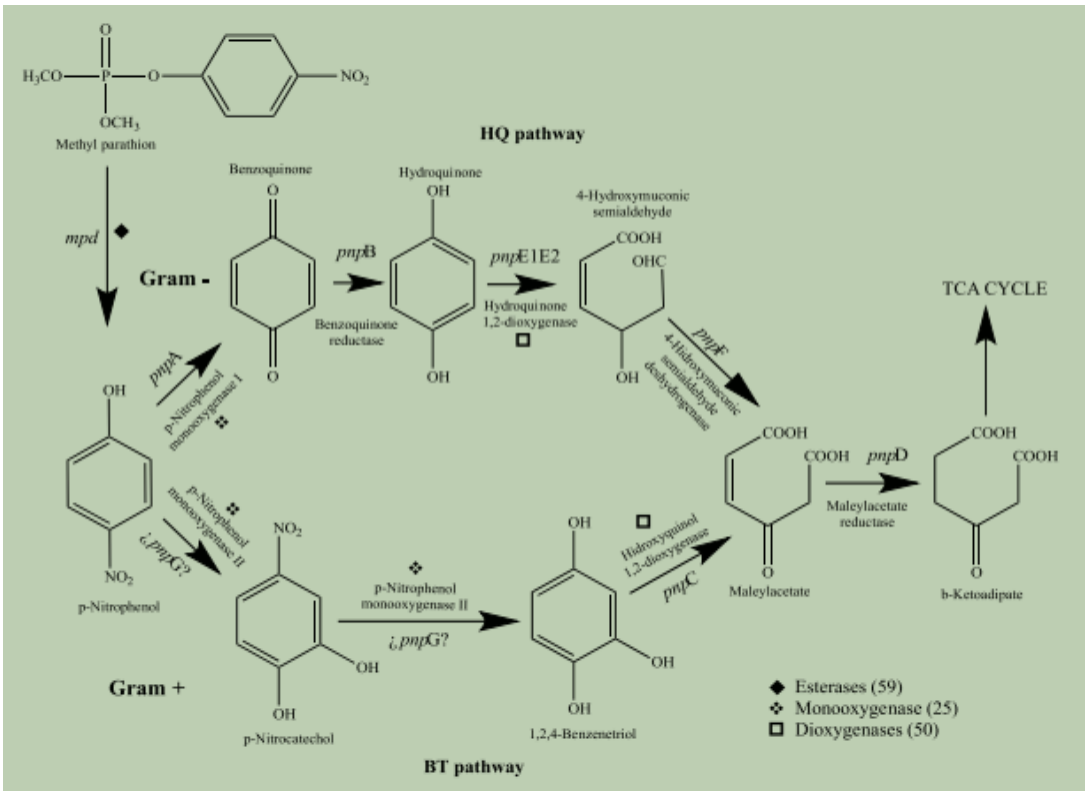


Figure 4.19: Possible metabolic pathways that could be used by the MPD gene for p Nitrophenol degradation and MP hydrolysis [15]

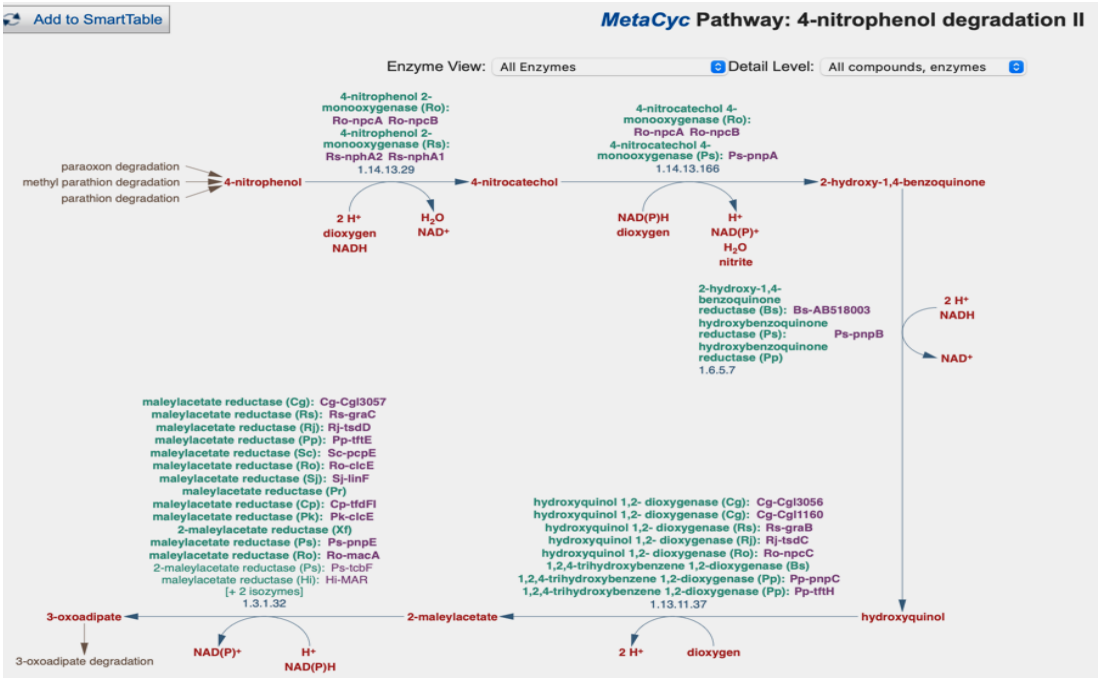


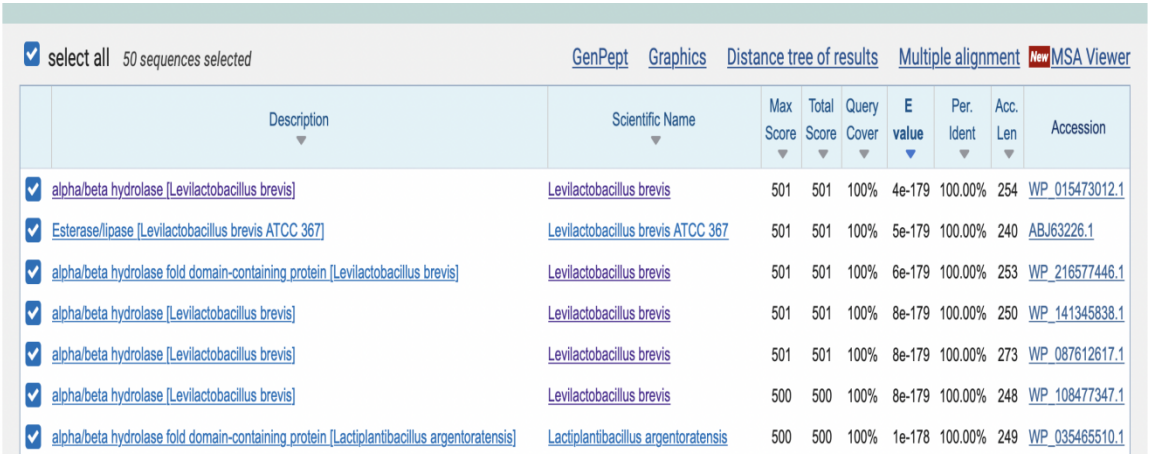
Figure 4.20: 4-Nitrophenol pathway from MP Degradation [38]

4.1.2 BLAST Analysis

The following six protein sequence shown in Figure 4.15 was used with BLASTx.

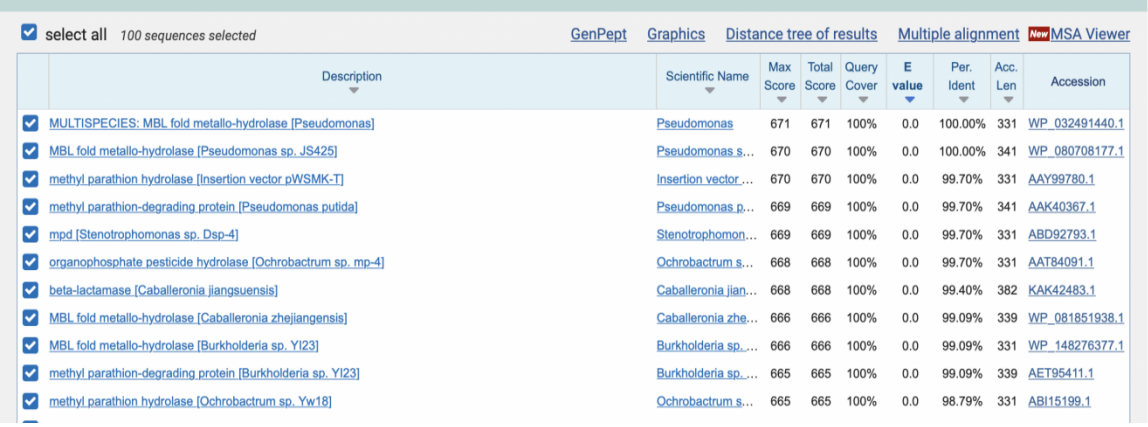
```
>tr|Q841S6|Q841S6_PSEWB Methyl parathion hydrolase OS=Pseudomonas sp. (strain WBC-3) OX=165468 GN=mpd PE=1 SV=1
MPLKNRLLARLSCVAAVVAATAAVAPLTLVSTAHAAAPQVRTSAPGYYRMLLGDFEITAL
SDGTVALPVDKRLNQPAPKTQSALAKSFQKAPLETSVTGYLVNTGSKLVLVDTGAAGLFG
PTLGRLAANLKAAGYQPEQVDEIYITHMHPDHVGGLMVGEQLAFPNAVVRADQKEADFWL
SQTNLDKAPDDESKGFFKGAMASLNPYVKAGKFKPFSGNTDLVPGIKALASHGHTPGHTT
YVVESQGQKLALLGDLILVAAVQFDDPSVTTQLDSDSKSVAVERKKAFADAAKGGYLIAA
SHLSFPGIGHIRAEGKGYRFVPVNYSVVNPK
```

Figure 4.21: Sequence query for soil sample without methyl parathion



select all 50 sequences selected — GenPept | Graphics | Distance tree of results | Multiple alignment | New MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
alpha/beta hydrolase [Levilactobacillus brevis]	Levilactobacillus brevis	501	501	100%	4e-179	100.00%	254	WP_015473012.1
Esterase/lipase [Levilactobacillus brevis ATCC 367]	Levilactobacillus brevis ATCC 367	501	501	100%	5e-179	100.00%	240	ABJ63226.1
alpha/beta hydrolase fold domain-containing protein [Levilactobacillus brevis]	Levilactobacillus brevis	501	501	100%	6e-179	100.00%	253	WP_216577446.1
alpha/beta hydrolase [Levilactobacillus brevis]	Levilactobacillus brevis	501	501	100%	8e-179	100.00%	250	WP_141345838.1
alpha/beta hydrolase [Levilactobacillus brevis]	Levilactobacillus brevis	501	501	100%	8e-179	100.00%	273	WP_087612617.1
alpha/beta hydrolase [Levilactobacillus brevis]	Levilactobacillus brevis	500	500	100%	8e-179	100.00%	248	WP_108477347.1
alpha/beta hydrolase fold domain-containing protein [Lactiplantibacillus argentoratensis]	Lactiplantibacillus argentoratensis	500	500	100%	1e-178	100.00%	249	WP_035465510.1



select all 100 sequences selected GenPept Graphics Distance tree of results Multiple alignment New MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
MULTISPECIES: MBL fold metallo-hydrolase [Pseudomonas]	Pseudomonas	671	671	100%	0.0	100.00%	331	WP_032491440.1
MBL fold metallo-hydrolase [Pseudomonas sp. JS425]	Pseudomonas s...	670	670	100%	0.0	100.00%	341	WP_080708177.1
methyl parathion hydrolase [Insertion vector pWSMK-T]	Insertion vector ...	670	670	100%	0.0	99.70%	331	AAY99780.1
methyl parathion-degrading protein [Pseudomonas putida]	Pseudomonas p...	669	669	100%	0.0	99.70%	341	AAK40367.1
mpd [Stenotrophomonas sp. Dsp-4]	Stenotrophomon...	669	669	100%	0.0	99.70%	331	ABD92793.1
organophosphate pesticide hydrolase [Ochrobactrum sp. mp-4]	Ochrobactrum s...	668	668	100%	0.0	99.70%	331	AAT84091.1
beta-lactamase [Caballeronia jiangsuensis]	Caballeronia jian...	668	668	100%	0.0	99.40%	382	KAK42483.1
MBL fold metallo-hydrolase [Caballeronia zhejiangensis]	Caballeronia zhe...	666	666	100%	0.0	99.09%	339	WP_081851938.1
MBL fold metallo-hydrolase [Burkholderia sp. YI23]	Burkholderia sp....	666	666	100%	0.0	99.09%	331	WP_148276377.1
methyl parathion-degrading protein [Burkholderia sp. YI23]	Burkholderia sp....	665	665	100%	0.0	99.09%	339	AET95411.1
methyl parathion hydrolase [Ochrobactrum sp. Yw18]	Ochrobactrum s...	665	665	100%	0.0	98.79%	331	ABI15199.1

Figure 4.22: Blast sequence results for soil sample without methyl parathion

Table 4.8: Top 10 Microorganisms with the highest BLASTx Total Score/ Max Score for Sample without MP

	Description	Scientific Name	Max Score	Total Score
1	Metallo Beta Lactamase (MBL) fold metallo-hydrolase	*Pseudomonas*	671	671
2	MBL fold metallo-hydrolase	*Pseudomonas sp. JS425*	670	670
3	Methyl parathion hydrolase (MPH)	*Insertion vector PWSMK-T*	670	670
4	Methyl parathion-degrading protein	*Pseudomonas putida*	669	669
5	MPD	*Stenotrophomonas sp Dsp-4*	669	669
6	OP pesticide hydrolase	*Ochrobactrum sp. Mp-4*	668	668
7	Beta-lactamase	*Caballeronia jiangsuensis*	668	668
8	MBL fold metallo-hydrolase	*Caballeronia zhejiangensis*	666	666
9	MBL fold metallo-hydrolase	*Burkholderia sp. Y123*	666	666
10	Methyl parathion-degrading protein	*Burkholderia sp. Y123*	666	666

The BLASTx result indicated the presence of MP hydrolase and MP degrading (MPD) protein that function as an MPD biomarker. The MPH is responsible for the hydrolysis of MP with the help of the MBL that acts as the activator for the nucleophilic attack to the MP bond. As the reaction progresses, the MBL also acts to stabilize the transition states of the intermediate compounds formed

V. CONCLUSION

It was found that the unamended soil sample contained an abundant phylum of proteobacteria (45%) which are known to degrade methyl parathion. However, upon amending the soil with MP a net reduction of microbial species was observed, including in species harboring degradation biomarkers. Based on bioinformatics data collected from MG-RAST, the MP degradation pathway is likely incomplete, effectively blocked with the formation of PNP. PNP is also poisonous and significantly reduces bacterial metabolism and growth. PNP genes are primarily found on mobile elements that were not present within these soil samples. It is therefore likely that upon inoculation into MP, the underlying bacterial population degraded MP to PNP, but were unable to proceed further and slowly died off.

The comparative metagenomic data for the unamended and amended farmland soil revealed three possible MPD genes in species *Burkholderia sp. AK-5, Rhizobium sp. MTP*-10005, and *Rhodanobacter sp. SCN.* A study on *Burkholderia sp. AK-5* revealed that it can degrade aromatic compounds like 4-aminophenol. Since the MPD pathway has no 4-aminophenol intermediate pathway, *Burkholderia sp. AK-5* will not be considered an methyl parathion gene for this study. The *Rhodanobacter sp. SCN* has little known degradation capability so it will also not be considered as an MPD gene for this study. The putative MPD species found for this study is the ***Rhizobium sp. MTP-10005***, an alphaproteobacterium, since it is the one that has the biodegradation capability via the maleylacetate and hydroxyquinol of the methyl parathion degradation pathway.

REFERENCES

1. EMNZ. (2015). *How pesticides affect soil microbes*. https://www.emnz.com/article/how-pesticides-affect-soil-microbes.
2. Adeyinka, A., Muco, E., & Pierre, L. (2021). Organophosphates. *StatPearls.* StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK499860/.
3. Tabassum, N., Rafique, U., Balkhair, K. S., & Ashraf, M. A. (2014). Chemodynamics of Methyl Parathion and Ethyl Parathion: Adsorption models for sustainable agriculture. *Biomed Research International,* Volume 2014. http://dx.doi.org/10.1155/2014/831989.
4. Garcia, S., Abu-Qare, A., Meeker-O'Connell, W., Borton, A., & Abou-Donia, M. (2003). Methyl Parathion: A review of health effects. *Journal of Toxicology and Environmental Health, Part B: Critical Reviews,* 6(2), 185–210.
5. Edwards, F. L., & Tchounwou, P. B. (2005). Environmental toxicology and health effects associated with Methyl Parathion exposure - A scientific review. *International Journal of Environmental Research and Public Health,* 2005, 2(3), 430-441.
6. Purg, M., Pabis, A., Baier, F., Tokoriki, N., Jackson, C., & Kamerlin, S. C. L. (2016). Probing the mechanisms for selectivity and promiscuity of methyl parathion hydrolase. *Philosophical Transactions of the Royal Society A.* 13 November 2016, Volume 374, Issue 2080. DOI: 10.1098/rsta.2016.0150.
7. Dong, Y. Bartlam, M., Sun, L., Zhou, Y., Zhang, Z., Zhang, C., Rao, Z., & Zhang, X. (2005). Crystal structure of methyl parathion hydrolase from Pseudomonas Sp. WBC-3. *Journal of Molecular Biology,* 2005, 353, 655-663. DOI: 10.1016/j.jmb.2005.08.057.
8. Thomas, T., Gilbert, J., & Meyer, F. (2012). Metagenomics - a guide from sampling to analysis. *Microbial Informatics and Experimentation, 2012; 2*(3). doi:10.1186/2042-5783-2-3

9. Segre, J. A. (n.d.). Metagenomics. *National Human Genome Research Institute.* https://www.genome.gov/genetics-glossary/Metagenomics .
10. Kalipci, E., Ozdemis, C., Oztas, F., & Schinkaya, S. (2010). Ecotoxicological effects of Methyl Parathion on living things and environment. *African Journal of Agricultural Research,* Volume 5(8), pp 712-718.
11. Imfeld, G., & Vuilleumier, S. (2012). Measuring the effects of pesticides on bacterial communities in soil: A critical review. *European Journal of Soil Biology,* 49 (2012; 22-30.
12. Bindhya, R., Sunny, S. A., & Thanga, V. S. G. (2009). In vitro study of the influence of methyl parathion on soil bacterial activity. *Journal of Environmental Biology*, May 2009, 30(3), 417-419.
13. Agency for Toxic Substances and Disease Registry (2001). Toxicological profile for methyl parathion. *U. S. Department of Health and Human Services.* https://www.atsdr.cdc.gov/toxprofiles/tp48.pdf.
14. Rayu, S. (2016). The interaction between xenobiotics and soil microbial communities. Ph.D. Thesis. *Western Sydney University, Australia.* Available at https://researchdirect.westernsydney.edu.au/islandora/object/uws:34253. (Accessed 13 October 2021).
15. Popoca-Ursino, E. C., Martinez-Ocampo, F., Dantan-Gonzales, E., Sanchez-Salinas, E., & Ortiz-Hernandez, M. L. (2017). Characterization of methyl parathion degradation by a *Burkholderia zhejangensis* strain, CEIB S4-3, isolated from agricultural soils. *Springer.* DOI: 10.1007/s10532-017-9801-1.
16. Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., & Edwards, R. A. (2008). The Metagenomics RAST server - a public resource for the automatic phylogenetic and functional

analysis of metagenomes. *BMC Bioinformatics,* 386 (2008). DOI:10.1186/1471-2105-9-386

17. National Center for Biotechnology Information (NCBI) (2021). BLAST. Available at https://blast.ncbi.nlm.nih.gov/Blast.cgi. (Accessed 17 October 2021).
18. Galperin, M. Y., Wolf, Y., Makarova, K. S., Alvarez, R. V., Landsman, D., & Koonin, E. V. (2021). COG database update: Focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Research*, Volume 49, Issue D1, 8 January 2021, pp. 274-281.
19. Huerta-Cepas, J., Szklarczyk, D., Heller. D., Hernandez-Plaza, A., Forslund, S. K., Cook, H., Mende, D. R., Letunic, I., Rattei, T., Jensen, L. J., Von Mering, C., & Bork, P. (2019). eggNOG 5.0; a hierarchical, functionally, and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*, Volume 47, 8 January 2019, pp. 309-314
20. Kanehisa Laboratories (2021). KEGG overview. Available from https://www.genome.jp/kegg/kegg1a.html. (Accessed 17 October 2021).
21. Microbe Wiki (2015). Conexibacter. Available from https://microbewiki.kenyon.edu/index.php/Conexibacter. (Accessed 17 October 2021).
22. Olanrewaju, O. S., & Babaluna, O. O. (2019). Streptomyces: implications and interactions in plant growth promotion. *Applied Microbiology and Biotechnology,* 2019, 103(3); 1179-1188.
23. De Souza, J. A. M., Alves, L. M. C., Varani, A. M., & Lemos, E. G. M. (2014). The family Bradyrhizobiaceae. *The Prokaryotes.* 2014, pp. 134-154.
24. Fuerst, J. A., & Sagulenko, E. (2011). Beyond the bacterium: Planctomycetes challenge our concepts of microbial structure and function. *Nature Reviews Microbiology,* 9, 2011, 403–413.
25. Bhatti, A. A., Haq, S., & Bhat, R. A. (2017). Actinomycetes' benefaction role in soil and plant health. *Microbial Pathogenesis*, Volume 111, October 2017, pp. 458–467.

26. Teng, Y., Wang, X., Li, L., Li, Z., & Luo, Y. (2015). Rhizobia and their bio-partners as novel drivers for functional remediation in contaminated soils. *Frontiers in Plant Science,* 05 February 2015. DOI:10.3389/fpls.2015.00032.

27. Lumen Learning (n.d). Proteobacteria. Available at https://courses.lumenlearning.com/boundless-microbiology/chapter/proteobacteria/.(Accessed 17 October 2021).

28. Zhang, B., Wu, X., Tai, X., Sun, L., Wu, M., Zhang, W., Chen, X., Zhang, G., Chen, T., Liu, G., & Dyson, P. (2019). Variation in actinobacterial community composition and potential function in different soil ecosystems belonging to the arid Heihe river basin of Northwest China. *Frontiers in Microbiology,* 24 September 2019. DOI:10.3389/fmicb.2019.02209.

29. Berkeley University of California (2021). NCBI bioinformatics resources: An introduction: BLAST: compare & identify sequences. Available at https://guides.lib.berkeley.edu/ncbi/blast. (Accessed 18 October 2021).

30. PerkinElmer (2019). Improving the efficiency of metagenomic analysis of soil samples. Available at https://perkinelmer-appliedgenomics.com/wp-content/uploads/2019/05/Improving-the-Efficiency-of-Metagenomic-Analysis-of-Soil-Samples.pdf.

31. Kumar, M. Yadav, A. N., Saxena, R., Paul, D., & Tomar, R. S. (2021). Biodiversity of pesticides degrading microbial communities and their environmental impact. *Biocatalysis and Agricultural Biotechnology, Volume 31, Jan 2021, 101883.* DOI:10.1016/j.bcab.2020.101883

32. Yue, W., Chen, M., Cheng, Z., Xie, L., & Li, M. (2018). Bioaugmentation of strain Methylobacterium sp. C1 towards p-nitrophenol removal with broad spectrum coaggregating bacteria in sequencing batch biofilm reactors. *Journal of Hazardous Materials*, 344 (2018), 431-440. DOI: 10.1016/j.jhazmat.2017.10.039

33. Liu, Y., Zhang, C., Liao, X., Luo, Y., Wu, S., & Wang, J. (2015). Hydrolysis mechanism of methyl parathion evidenced by Q-extractive mass spectrometry. *Environmental Science and Pollution Research*, 22(24). DOI. 10.1007/s11356-015-5169-0

34. Iyer, R., & Damania, A. (2020). Shotgun metagenomics of indigenous bacteria collected from the banks of the San Jacinto River for the biodegradation of aromatic waste. *FEMS Microbiology Letters*, 367, 2020, fnaa133. DOI: 10.1093/femsle/fnaa133

35. Caspi et al (2018). MetaCyc Organism: Rhizobium sp. MTP-10005. *Nucleic Acid Research,* 46(D1): D633-D639: 2018.

36. Hata, Y., Fujii, T., Yoshida, M., Oikawa, T. (2009). Structure of Maleylacetate reductase from Rhizobium sp. Strain MTP-10005. *AsCA '09 Beijing Joint Conference of the Asian Crystallographic Association and Chinese Crystallographic Society,* October 22-25, 2009; Beijing, China.

37. Caspi, et al (2014). MetaCyc Pathway: 4-aminophenol degradation. *Nucleic Acid Research,* 42: D459-D471. Available at **http://vm-trypanocyc.toulouse.inra.fr/META/NEW-IMAGE?type=PATHWAY&object=PWY-7081**

38. Caspi, et al (2018). MetaCyc Pathway: 4-nitrophenol degradation. *Nucleic Acid Research, Nucleic Acid Research,* 46(D1): D633-D639: 2018. Available at **https://biocyc.org/META/NEW-IMAGE?type=PATHWAY&object=PWY-5487**

39. Yang, W., Zhang, X. (2008). Application of methyl parathion hydrolase as a labeling enzyme. *Analytics and Bioanalytical Chemistry.* DOI:10.1007/s00216-008-1987-y

40. Kaehler, B. D., Bokulich, N. A., McDonald, D., Knight, R., Caporaso, J. G., & Huttley, G. A. (2019). Special abundance information improves sequence taxonomy classification accuracy. *Nature Communications*, 10, 4643 (2019).DOI:/10.1038/s41467-019-12669-6

41. Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., Harend, H., Huerta-Cepas J., Medema, M.H., Maltz, M.R., Mundra, S., Olsson, P.A., Pent, M,. Polme, S., Sunagawa S., Ryberg, M., Tedersoo, L. & Bork, P. (2018), Structure and function of the global topsoil microbiome, *Nature* 560: 233-237. DOI: 10.1038/s41586-018-0386-6

APPENDIX

List of Abbreviations

AChE	Acetylcholine Esterase
BLAST	Basic Local Alignment Search Tool
CEC	Cation Exchange Capacity
COG	Clusters of Orthologous Groups
CSM	Carbon Selective Media
eggNOG	Evolutionary Genealogy of Genes, Non-supervised orthologous Groups
KEGG	Kyoto Encyclopedia of Genes and Genomes
MBL	Metallo-β-lactamase
MEGA	Molecular Evolutionary Genetics Analysis
MG-RAST	Metagenomic Rapid Annotations using Subsystems Technology
MP	Methyl Parathion
MPH	Methyl Parathion Hydrolase
MSM	Mineral Salts Medium
OP	Organophosphate
ORF	Open-reading Frame
PCR	Polymerase Chain Reaction
PFGE	Pulsed-field and Electrophoresis Technique
PICT	Pollutant Induced Community Tolerance