OPEN ACCESS

Saudi Journal of Biomedical Research
Abbreviated Key Title: Saudi J Biomed Res
ISSN 2518-3214 (Print) | ISSN 2518-3222 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

Comparative Analysis for Bioremediation of Plastic and Dye Degradation

Ms. Manognya Pradeep[1], Ms. Pujaa B[1*]

[1]Assistant Professor, Stella Maris College, India

DOI: https://doi.org/10.36348/sjbr.2024.v09i08.001 | **Received:** 22.08.2024 | **Accepted:** 26.09.2024 | **Published:** 01.10.2024

***Corresponding author:** Ms. Pujaa B
Assistant Professor, Stella Maris College, India

Abstract

The leading environmental risk factor for disease and premature death is pollution, with plastic and dye pollutants being the most common in developing countries. The global textile industry contributes to pollution by releasing contaminated wastewater into water bodies, leading to a decline in water quality. Plastic pollution is a widespread issue affecting various environments, emphasizing the urgent need for a global response to combat the adverse effects of pollution on human health and the environment. *Anoxybacillus sp. PDR2* is a soil bacterium possessing natural competence. By nature, it is a thermophile, capable of biodegrading industrial dyes. *Pseudomonas sp. B10*, Gram-negative bacteria, is a strain capable of degrading Polyethylene Terephthalate (PET) plastic. To detect the genome-level mutations, comparative genomic analysis was performed using a free and open source software, Galaxy. Using five different variant callers (Samtools, Varscan, Freebayes, Sniffles, Ivar), mutations were detected at various loci resulting in the modifications of the genes. The primary goal of this investigation was to perform a comparative analysis of the whole genome sequencing of two bacterial species, along with their reference strains. The purpose was to identify potential solutions for the degradation of plastics and industrial dyes. By examining the genetic composition of these bacteria, this analysis had provided valuable insights into the genetic makeup of these bacteria and their ability to break down PETs and dyes.

Keywords: *Plastic Pollution, Anoxybacillus, Pseudomonas, Polyethylene Terephthalate, Bioremediation, Comparative Genomics.*

INTRODUCTION

Pollution was identified as the leading environmental risk factor for disease and premature death in 2015, resulting in 9 million premature fatalities. Subsequently, in 2019, findings confirmed that pollution continues to be responsible for around 9 million deaths annually, accounting for one in every six deaths globally (Fuller *et al.,* 2022).

The two most prevalent pollutants in developing countries are plastic and dye pollutants. The global textile industry is a source of pollution due to the accidental discharge or dumping of polluted waste into water bodies, leading to a notable deterioration in water quality. As outlined in a report by the World Bank, around 17–20% of industrial water pollution can be attributed to the dyeing and treatment activities within the textile sector. This poses a substantial environmental challenge for textile manufacturers. Given the rising environmental awareness, there is a pressing need for sustainable technology to eradicate dyes from industrial and local wastewater. The extensive presence of organic dyes in industrial wastewater from the paper, textile, and clothing industries leads to significant environmental contamination. Research studies have indicated that approximately 10-12% of dyes such as Rhodamine B, Victoria blue, Rose Bengal, Indigo Red, Carmine, Red 120, Eriochrome, Methylene Blue (MB), Black-T (EBT), and Thymol blue are annually utilized in textile industries, with a substantial portion (around 20%) being lost during synthesis and processing activities and ending up in wastewater. These dye-polluted effluents contain non-biodegradable, highly toxic, and colored pigments that pose a threat to living organisms (Rafiq *et al.,* 2021).

Plastic pollution is a widespread problem that affects diverse environments, ranging from deserts and farms to mountaintops and the deep sea, as well as tropical landfills and Arctic snow. Reports of plastic waste in marine ecosystems have been documented for the past fifty years, with a continuous buildup on the ocean surface over the past six decades (MacLeod *et al.,* 2021).

In recent times, it has become increasingly evident that pollution is a grave concern for the entire

planet. The factors that contribute to pollution, its widespread dispersion, and its adverse effects on human health transcend local boundaries, demanding a global response. It is imperative to undertake global action to address all significant contemporary pollutants. By harnessing the synergy between global efforts and other environmental policy programs, particularly in swiftly transitioning from fossil fuels to clean and renewable energy sources, we can effectively combat pollution while also mitigating the pace of climate change.

Bioremediation

Bioremediation emerges as an efficient and economical solution for the eradication of environmental pollutants. Soil microorganisms perform essential functions such as promoting plant growth, controlling insects, conserving soil integrity, recycling nutrients, and reducing pollutants. The progress made in bioremediation has resulted in improved efficiency, cost-effectiveness, and societal approval. Research in bioremediation has predominantly focused on bacterial processes, which have proven to be versatile in numerous applications. Archaea also contribute to bioremediation efforts in collaboration with bacteria in various environmental remediation contexts. Bioremediation involves the utilization of microorganisms, including bacteria, algae, fungi, and plants, to degrade, alter, eliminate, immobilize, or detoxify diverse physical and chemical contaminants present in the surrounding environment (Bala *et al*., 2022).

Comparative Genomics

Comparative genomics serves as a valuable instrument for comprehending the genetic alterations occurring in microbial adaptation and investigating the evolutionary processes experienced by microorganisms. The identification of genetic changes and their subsequent effects presents a formidable challenge that demands considerable time and effort. Several comparative genomic tools have been developed to facilitate the comparison between laboratory-evolved strains and their wild counterparts, enabling the detection of mutations and subsequent analysis of the acquired data to derive meaningful interpretations.

METHODOLOGY

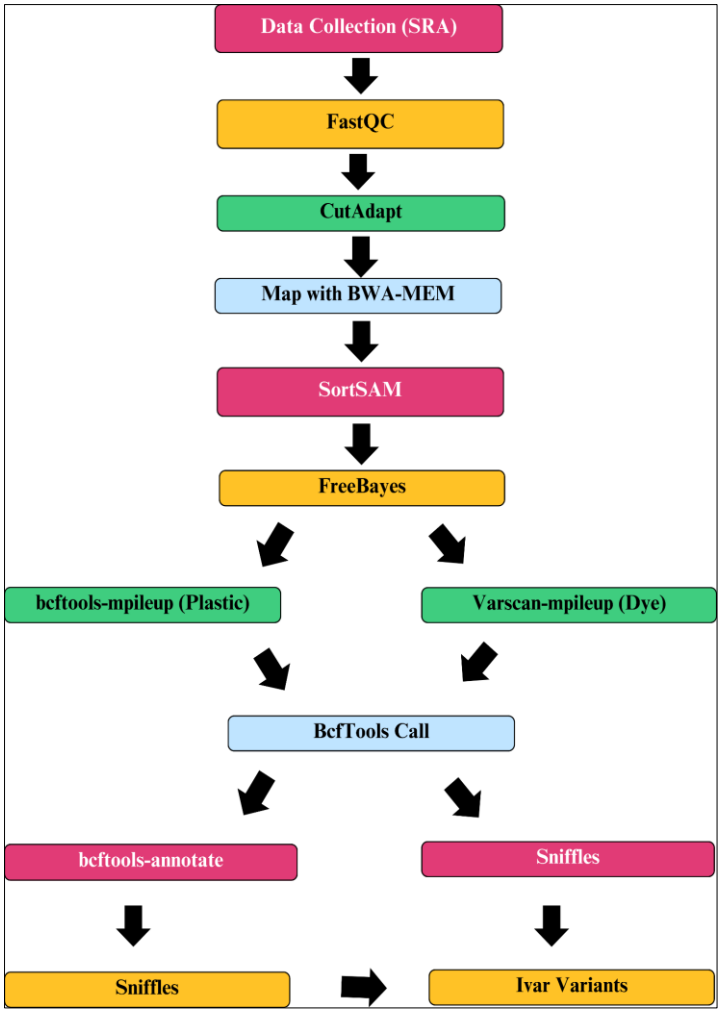

Fig. 1: Workflow of the analysis.

Data Collection

The raw read sequences of both the sample sequences and the reference genomes for both sample sequences were obtained from SRA and NCBI Genome datasets. The information on the samples retrieved are mentioned in tables 1 and 2.

The specific samples were collected from designated sources. The analysis was carried out utilizing UseGalaxy, a web-based analysis platform that is open source and provides access to a range of tools. The tools utilized within the platform included FastQC for evaluating sequence quality, GC percentage, and adaptor content. Furthermore, the Cutadapt tool was used to eliminate any inaccurate or highly expressed sequences from the sample data. The reference genome was aligned with the sample dataset using Map with BWA-MEM with default parameters. The aligned datasets were then sorted using the SortSam tool with default settings. Variation analysis was conducted using FreeBayes, bcftools-mpileup, bcftools-call, bcftools-annotate, and Varscan-mpileup tools to pinpoint specific positions where variants such as SNPs or indels were identified. Structural variant prediction was performed using the Sniffles tool on the data obtained from MarkDuplicates with default parameters. Finally, Ivar variants were identified using the data from MarkDuplicates in BAM file format with default parameters.

Table 1: Raw read sequences retrieved from SRA

	SRA Accession ID	SRX ID	Study
1.	SRR8835125	**SRX5623360**: 1 ILLUMINA (Illumina MiSeq) run: 1.1M spots, 324.3M bases, 198.3Mb downloads	Draft Genome Sequences of Five Environmental Bacterial Isolates That Degrade PET and LDPE Plastic
2.	SRR10820142	**SRX7493616**: 1 ILLUMINA (Illumina HiSeq 2500) run: 5.3M spots, 1.6G bases, 470.2Mb downloads	Molecular mechanism of degradation of azo dyes by thermophilic anaerobic bacillus

Table 2: Reference genomes downloaded from NCBI Genome datasets.

	GenBank ID	Genome assembly	Description
1.	GCF_001187595.1	ASM118759v1	Microbe sample from *Anoxybacillus gonensis*
2.	GCF_000009225.2	ASM922v1	BioSample entry for genome collection GCA_000009225

RESULTS

Prepossessing of Reads

The QC report generated by FASTQC provided an overview of the basic statistics for the raw reads data. It indicated that the quality of the data was satisfactory as most of the tests were successfully passed. The plastic data set comprised a total of 21,47,878 sequences with a GC content of 59%, while the dye data set consisted of 1,06,89,622 sequences with a GC content of 48%. Table 3, presents detailed information regarding the quality control measures. To ensure accuracy, the duplicates were eliminated using MarkDuplicates after the mapping process. Additionally, the reads underwent trimming using CutAdapt, and the resulting output files were aligned with the reference genome.

Table 3: QC report of raw read data.

Test name	SRR8835125	SRR10820142
Per base sequence quality	passed	passed
Per sequence quality scores	passed	passed
Per base sequence content	failed	failed
Per base GC content	passed	passed
Per sequence GC content	warning	failed
Per base N content	passed	passed
Sequence Length Distribution	passed	passed
Sequence Duplication Levels	passed	failed
Overrepresented sequences	passed	failed
Adapter Content	passed	passed

Mapping and Variant Calling

Mapping was done by BWA-MEM. BWA is a software package that is specifically designed for mapping low-divergent sequences against a large reference genome, such as the human genome. The sample dataset, consisting of *Pseudomonas sp.B10* and *Anoxybacillus sp.PDR2*, was mapped to the reference genomes of *Pseudomonas fluorescens* and *Anoxybacillus gonensis* using the default parameters. The input data was in fastq format, and the output obtained was a BAM file. Duplicate reads were removed by MarkDuplicates. Variants were called using four variant analysis tools

(Varscan, Sniffles, Bcftools, Ivar variants). These programs analyzed the alignments and created a vertical slice for each genomic location, compiling all reads covering that specific position in a pileup. They take a BAM file as input and generate a VCF file as the result.

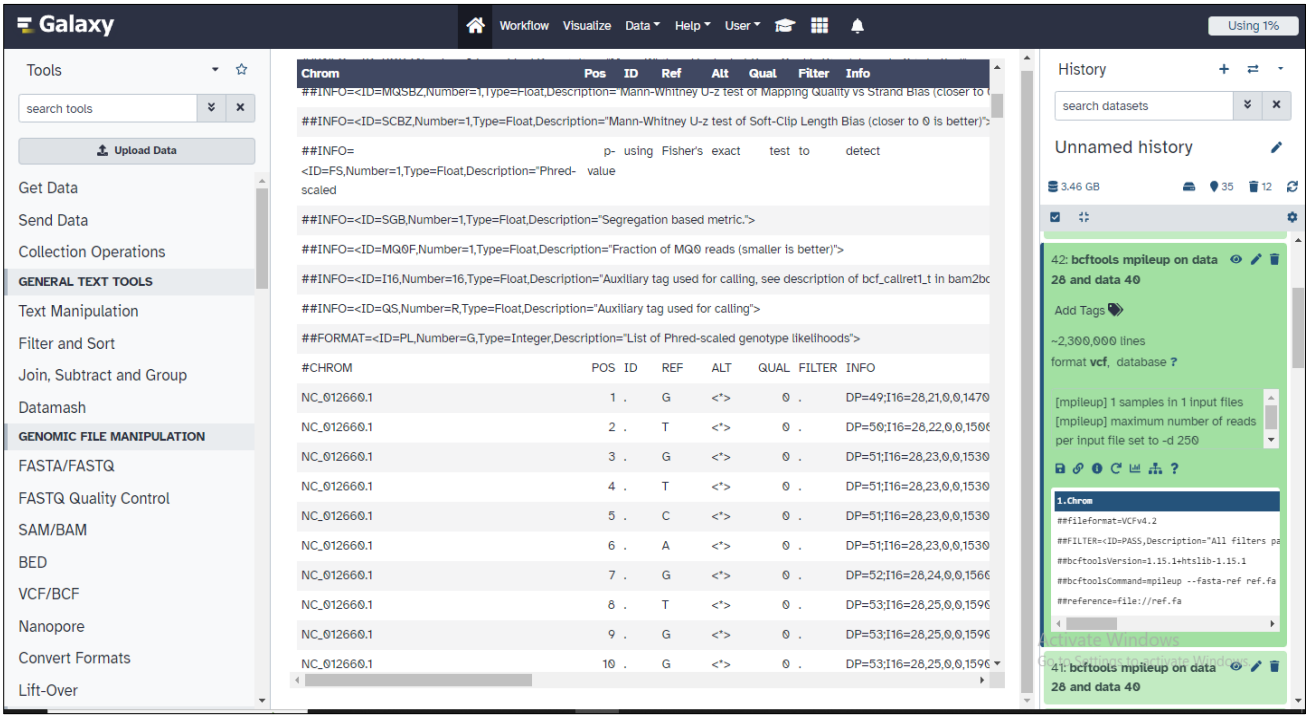

Fig. 2: First ten variants observed from bcftools-mpileup.

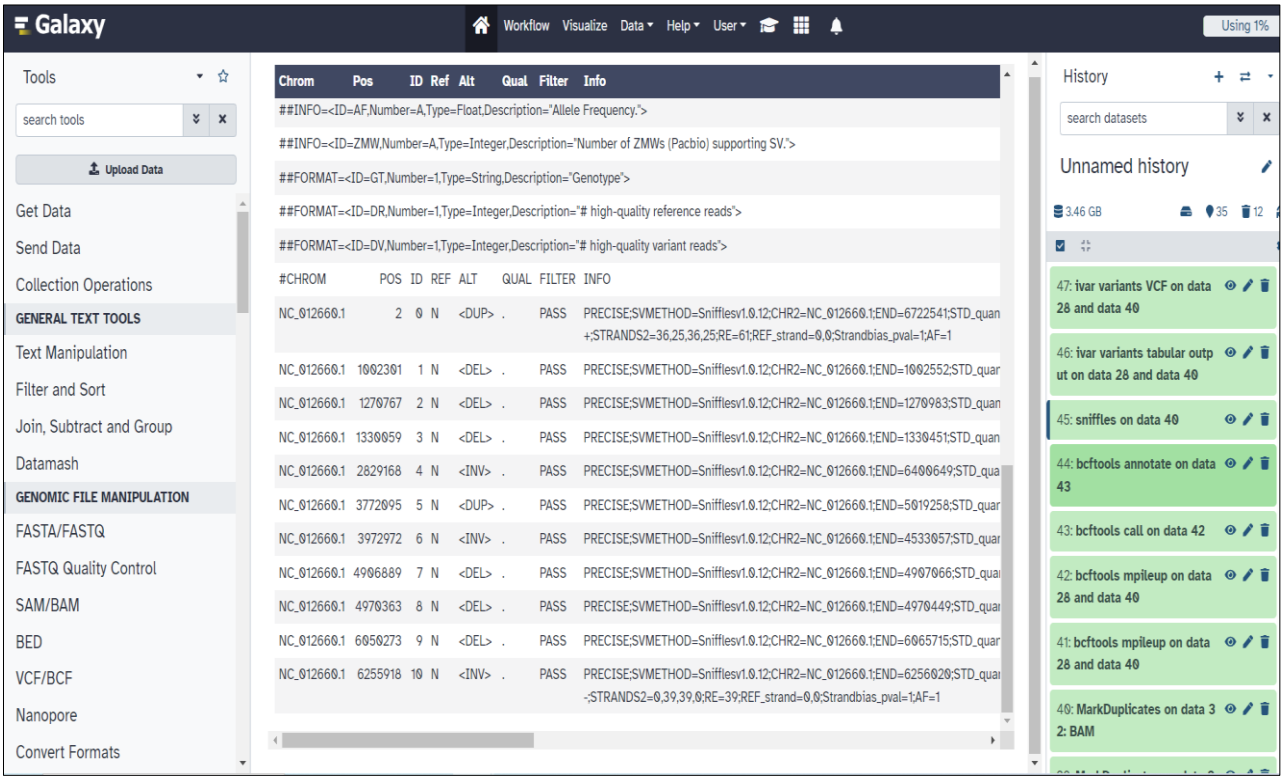

Fig. 4: Structural variants produced by sniffles for the plastic dataset.

Table 4: All structural variants with mutation types for the plastic dataset

Chromosome	Position	Mutation	Quality
NC_012660.1	2	Duplication	Pass
NC_012660.1	1002301	Deletion	Pass
NC_012660.1	1270767	Deletion	Pass
NC_012660.1	1330059	Deletion	Pass
NC_012660.1	2829168	Inversion	Pass
NC_012660.1	3772095	Duplication	Pass
NC_012660.1	3972972	Inversion	Pass
NC_012660.1	4906889	Deletion	Pass
NC_012660.1	4970363	Deletion	Pass
NC_012660.1	6050273	Deletion	Pass
NC_012660.1	6255918	Inversion	Pass

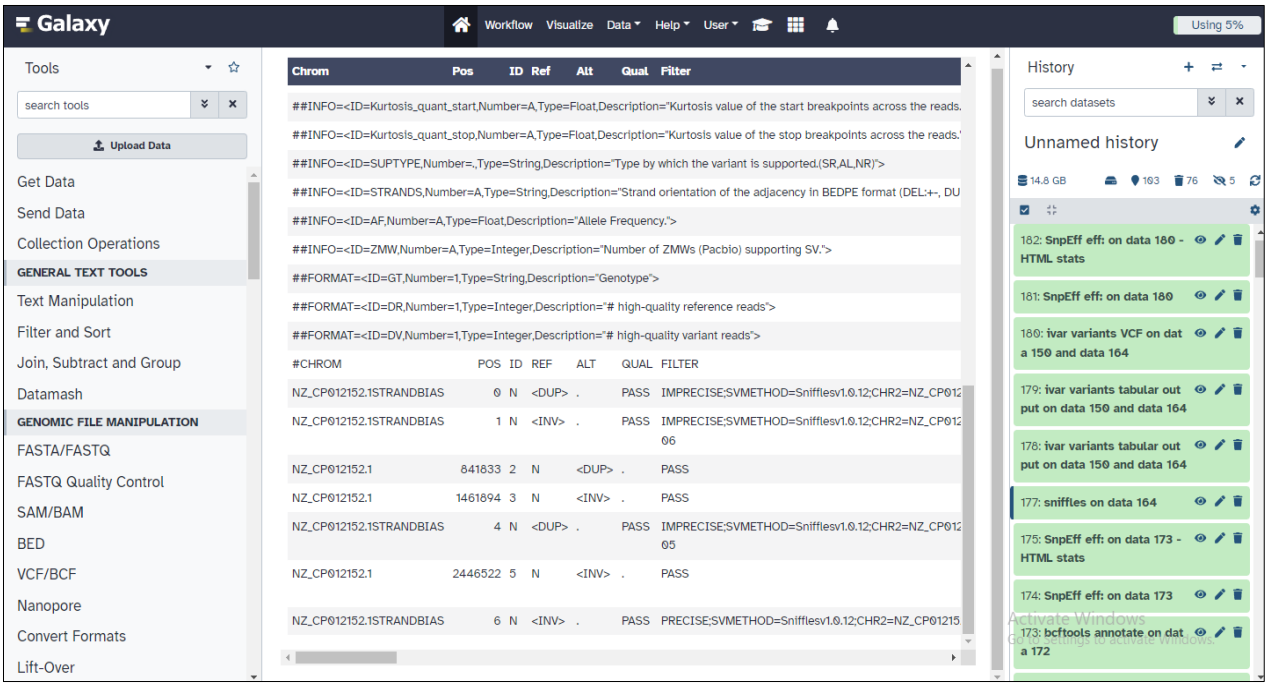

Fig. 5: Structural variants produced by sniffles for the dye dataset.

Table 5: All structural variants with mutation types for the dye dataset

Chromosome	Position	Mutation	Quality
NZ_CP012152.1	No positional information	Duplication	Pass
NZ_CP012152.1	No positional information	Inversion	Pass
NZ_CP012152.1	841833	Duplication	Pass
NZ_CP012152.1	1461894	Inversion	Pass
NZ_CP012152.1	No positional information	Duplication	Pass
NZ_CP012152.1	2446522	Inversion	Pass
NZ_CP012152.1	No positional information	Inversion	Pass

Annotation of Variants

The structural variants that were identified by sniffles were taken for further analysis. There were eleven SNPs identified in the plastic dataset. These eleven SNPs were annotated. Six variants were found to be deletions, while three were inversions and two were duplications. When the dye dataset was observed, seven SNPs were identified. There were four inversions and three duplications. Subsequently the ivar variants were identified and were taken for further analysis. The first ten variants identified for both the datasets (plastic and dye), are listed down as tables below. These variants were further annotated with the consequence of the mutation.

Table 6: Ivar variants with mutations for dye

Chromosome	Product	Location	Reference allele	Alternate allele	Quality
NZ_CP012152.1	DNA replication initiation ATPase	489	C	T	FAIL
NZ_CP012152.1	DNA replication initiation ATPase	492	A	G	FAIL
NZ_CP012152.1	DNA replication initiation ATPase	498	T	C	FAIL
NZ_CP012152.1	DNA replication initiation ATPase	504	C	T	FAIL
NZ_CP012152.1	DNA replication initiation ATPase	507	T	G	FAIL
NZ_CP012152.1	DNA replication initiation ATPase	508	C	A	FAIL
NZ_CP012152.1	DNA replication initiation ATPase	522	G	A	PASS
NZ_CP012152.1	DNA replication initiation ATPase	528	G	C	PASS
NZ_CP012152.1	DNA replication initiation ATPase	549	G	T	PASS
NZ_CP012152.1	DNA replication initiation ATPase	555	G	A	PASS

Table 7: Ivar variants with mutations for plastic

Chromosome	Product	Location	Reference allele	Alternate allele	Quality
NC_012660.1	Chromosomal replication initiator protein DnaA	45	T	A	PASS
NC_012660.1	Chromosomal replication initiator protein DnaA	60	G	A	PASS
NC_012660.1	Chromosomal replication initiator protein DnaA	117	T	C	PASS
NC_012660.1	Chromosomal replication initiator protein DnaA	120	T	C	PASS
NC_012660.1	Chromosomal replication initiator protein DnaA	126	G	A	PASS
NC_012660.1	Chromosomal replication initiator protein DnaA	129	C	G	PASS
NC_012660.1	Chromosomal replication initiator protein DnaA	132	T	C	PASS
NC_012660.1	Chromosomal replication initiator protein DnaA	144	C	G	PASS
NC_012660.1	Chromosomal replication initiator protein DnaA	159	G	A	PASS
NC_012660.1	Chromosomal replication initiator protein DnaA	169	A	G	PASS

In NZ_CP012152.1 from the position 1 to 1744, the product was DNA replication initiation ATPase. The presence of ATP in the protein is crucial for the continuation of replication initiation, leading to a phase where ATP hydrolysis is activated, resulting in a change in the protein's activity (Lee *et al*., 2000). The first 10 variants of both samples were found to be point mutations. From the 189th position in the plastic dataset, there was a change in the mutation. There were some insertions, deletions and frameshift mutations. In NC_012660.1n from the position 1 to 1506, the product was chromosomal replication initiator protein DnaA. The initiation and regulation of chromosomal replication in bacteria heavily rely on the central domain of DnaA proteins. DnaA, a protein that binds both ATP and DNA, plays a crucial role in this process. Specifically, it binds to 9 bp nucleotide repeats called dnaA boxes, which are present in the origin of replication (oriC) on the chromosome. DnaA is approximately 50kDa in size and consists of two conserved regions. The first region, located in the N-terminal half, is responsible for ATP binding. The second region, found in the C-terminal half, is potentially involved in DNA binding. Additionally, DnaA has the ability to interact with other proteins such as the RNA polymerase beta subunit, dnaB and dnaZ proteins, as well as the groE gene products (chaperonins). Overall, the central domain of bacterial DnaA proteins plays a critical role in the initiation and regulation of chromosomal replication, with DnaA itself acting as an ATP- and DNA-binding protein. Its specific binding to dnaA boxes in the oriC region further contributes to the precise control of replication in bacteria (Skarstad and Boye, 1994). From the 30th position in the dye dataset, there was a change in the mutation, there are insertions and frameshift mutations.

DISCUSSION

Anoxybacillus sp. PDR2 exhibited a response to different carbon sources, wherein the addition of any external carbon source resulted in the promotion of methyl red (MR) decolorization. The optimization of the carbon source, nitrogen source, and temperature in the culture medium is a common practice among researchers to enhance the decolorization efficacy of bacteria towards azo dyes. However, the initial pH, initial inoculation amount, and metal ion concentration during the decolorization process are equally crucial factors influencing the bacteria's decolorization of azo dyes. Therefore, this study concentrated on optimizing the initial pH, initial inoculation amount, and metal ion concentration as variables that impact the rate of MR decolorization (Zhang *et al*., 2023).

To identify the metabolic pathways and genes connected to PET degradation, the genome sequences of the five strains were procured. In brief, the bacteria were cultured in lysogeny broth at a temperature of 26°C overnight. Preliminary metabolic comparisons were conducted using the KEGG database via BlastKOALA, which revealed that the genomes possess common central carbohydrate metabolism and biosynthetic capabilities, encompassing the synthesis of nucleotides

and amino acids. Furthermore, the genomes share genes associated with transport systems for both simple and complex biomolecules. It is worth noting that each genome contains a substantial number of genes that are predicted to encode lipase-like enzymes (Leon-Zayas *et al*., 2019).

This study showed that adaptive laboratory evolution leads to genetic modification in the *Anoxybacillus sp*. PDR2 and *Pseudomonas sp*. B10 strains and the comparative analysis of these bacteria showed that they can be used for the bioremediation of dyes and plastics using lab-evolved techniques. This study has aimed to provide the first-ever analysis of these bacteria for bioremediation, computationally and using bioinformatics tools.

CONCLUSION

The research identified various mutations in the analyzed samples, with variants present in multiple genes across different chromosomal locations. Notably, the study highlighted a high prevalence of variants in chromosomes NC_012660.1 and NZ_CP012152.1. Given that each tool possesses a distinct collection of inherent parameters and algorithms for variant detection, it is expected that there will be variations in the statistical characteristics of the raw variants that have been identified. However, when considering the filtered variants identified by all four variant calling pipelines, they exhibit a high degree of similarity.

Through further investigation and analysis, the mutational profiles of these samples can be utilized to explore the functionality of affected genes and develop potential bio-remediation strategies for plastic and dye pollution. The findings of this study demonstrate that adaptive laboratory evolution results in genetic modifications in *Anoxybacillus sp*. PDR2 and *Pseudomonas sp*. B10 strains, indicating their potential for bioremediation of dyes and plastics using lab-evolved techniques.

REFERENCES

- Bala, S., Garg, D., Thirumalesh, B. V., Sharma, M., Sridhar, K., Inbaraj, B. S., & Tripathi, M. (2022). Recent Strategies for Bioremediation of Emerging Pollutants: A Review for a Green and Sustainable Environment. *Toxics*, *10*(8), 484. https://doi.org/10.3390/toxics10080484.
- Fuller, R., Landrigan, P. J., Balakrishnan, K., Bathan, G., Blüml, S., Bräuer, M., Caravanos, J., Chiles, T., Cohen, A., Corra, L., Cropper, M., Ferraro, G., Hanna, J. L., Hanrahan, D., Hu, H., Hunter, D., Janata, G., Kupka, R., Lanphear, B. P., & Yan, C. (2022). Pollution and health: a progress update. *The Lancet Planetary Health*, *6*(6), e535–e547. https://doi.org/10.1016/s2542-5196(22)00090-0
- https://broadinstitute.github.io/picard/
- https://broadinstitute.github.io/picard/command-line-overview.html
- https://usegalaxy.org/
- https://varscan.sourceforge.net/germline-calling.html
- https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
- Lee, D. G., & Bell, S. P. (2000). ATPase switches controlling DNA replication initiation. *Current opinion in cell biology*, *12*(3), 280-285.
- León-Zayas, R., Roberts, C., Vague, M., & Mellies, J. L. (2019). Draft genome sequences of five environmental bacterial isolates that degrade polyethylene terephthalate plastic. *Microbiology Resource Announcements*, *8*(25), 10-1128.
- MacLeod, M., Arp, H. P. H., Tekman, M. B., & Jahnke, A. (2021). The global threat from plastic pollution. *Science*, *373*(6550), 61–65. https://doi.org/10.1126/science.abg5433
- Rafiq, A., Ikram, M., Ali, S., Niaz, F., Khan, M. I., Khan, Q., & Maqbool, M. (2021). Photocatalytic degradation of dyes using semiconductor photocatalysts to clean industrial water pollution. *Journal of Industrial and Engineering Chemistry*, *97*, 111–128. https://doi.org/10.1016/j.jiec.2021.02.017
- Skarstad, K., & Boye, E. (1994). The initiator protein DnaA: evolution, properties and function. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, *1217*(2), 111-130.
- Zhang, S., Feng, L., Han, Y., Xu, Z., Xu, L., An, X., & Zhang, Q. Bioremediation Possibility of Two Azoreductases of Anoxybacillus Sp. Pdr2 in Degradation of Azo Dye Methyl Redbioremediation Possibility of Two Azoreductases of Anoxybacillus Sp. Pdr2 in Degradation of Azo Dye Methyl Red. *Pdr2 in Degradation of Azo Dye Methyl Redbioremediation Possibility of Two Azoreductases of Anoxybacillus Sp. Pdr2 in Degradation of Azo Dye Methyl Red*.