Investigating Suburban Locations in California for the Presence and Amount of Tetracycline Resistance Genes

Sean Yu

Received June 16, 2024
Accepted September 09, 2024
Electronic access October 15, 2024

Antibiotic resistance is a global public health crisis caused by the spread of resistant bacteria. Prior studies have identified environmental soil as a potential hotspot for antibiotic resistance, Thus, researchers have investigated agricultural, rural, and wilderness areas for resistant bacteria or antibiotic-resistance genes. However, only some studies have investigated suburban locations for the presence and amount of resistance genes. Investigating suburban locations is important because they have higher population densities than previously explored areas. This study collected soil samples from various suburban locations and extracted DNA from bacteria in the samples. The extracted DNA was run with gel electrophoresis. Afterward, linear calibration curves and densitometric analyses were used to verify and quantify any resistance presence. This study found most samples (62.5%: 6/16 samples) contained no resistance. Of the resistant locations, lawn soil (BAR) contained the most resistance based on normalized values [BAR – 1.068 | | COU – 0.604 | | PEU – 0.253 | | PEU – 0.145]. The other resistant locations had the greatest frequency of human activity. Lawn soil could contain more resistance because its unique chemical composition increases the selection of resistant bacteria. Locations with increased human activity often experience greater transfers in antibiotic-resistant bacteria and genes. This contributes to increased resistance in soil. These results indicate resistance in suburban locations is relatively limited. These results also show trends in resistance in suburban locations that could match trends in other locations. This knowledge could be used to pinpoint and mitigate significant hotspots of resistance in suburban areas. Future studies should continue affirming these trends and collect more data on resistance in suburban locations.

Keywords: antibiotic resistance, DNA extraction, environmental soil, lawn soil, human activity

Introduction

Antibiotic resistance is a global public health crisis. Increasing numbers of antibiotics are becoming ineffective as resistant bacteria spread[1]. The exact hotspots of resistant bacteria are unknown. However, studies have suggested these bacteria originate in environmental soil samples[2–4]. In this soil, antibiotic-resistant bacteria (ARBs) are created through new mutations or by receiving resistance genes (ARGs) through horizontal gene transfer. They theninfect humans and animals[3,4]. These bacteria multiply and spread from their host to other organisms, for example when humans eat contaminated poultry from infected animals[4]. For instance, infected humans may transfer resistant bacteria to the environment around hospitals or excrete resistant bacteria in waste. Although this waste is processed through wastewater treatment plants (WWTPs), ARBs and ARGs may survive and spread to the soil around WWTPs[2].

The spread of ARBs and ARGs to soil is augmented by the overuse of antibiotics in healthcare and agriculture[2,5–7]. Antibiotics provided to humans and animals are often not completely metabolized[7]. Animals spread their antibiotics into soil through waste. Although human waste is processed, WWTPs are not completely effective at removing antibiotics. Thus, antibiotics may leach into soil where WWTPs discharge treated water and select for resistant bacteria in this soil[7]. Antibiotic overuse increases the rate at which antibiotics eliminate nonresistant bacteria inside animals, humans, and environmental soil. This allows resistant bacteria to proliferate until only resistant bacteria are in a population[8].

Mitigating the causes of antibiotic resistance is even more important now. This is because current trends project infections from resistant bacteria will kill 10.0 million people annually by 2050[5]. Developing extensive knowledge of antibiotic resistance, such as studying its presence and amount in likely hotspots is crucial to mitigating this crisis[2,3,5,6,9]. Previous studies have sampled the soil of large areas (i.e. entire countries). These studies have primarily explored soil sampled from general agricultural, rural, or wilderness areas displaced from civilization. These studies have tested for the presence and amount of either resistant bacteria or a wide variety of resistance genes in these samples, as well as the spread and interrelationships of these bacteria and genes. The studies have also primarily analyzed the effects of present or added external environmental factors, such as manure, moisture, and temperature, on the amount of

resistance detected in soil samples[9–15]. However, few studies have analyzed the presence and amount of resistance genes in soil from suburban environments. Additionally, few studies have analyzed the effects of varying human interaction and internal chemical composition of soil on the presence and amount of resistance.

This study aims to address these gaps in research by providing information on the prevalence of resistance genes in suburban environments. This study will do so by extracting microbial genomes from soil collected in suburban locations throughout the Bay Area. The Bay Area was selected due to its proximity to the facilities used during the study. The Bay Area has many vast suburban landscapes that provide many sampling sites. Additionally, the Bay Area has a large population. There are significant human interactions with the environment, increasing the likelihood of the spread of antibiotic resistance. This study used high-school laboratory equipment to analyze the presence and amount of the tetracycline resistance genes TETB and TETM; it does not culture bacteria or alter locations sampled. Initially, the plan was to collect 18 distinct samples to cover a large area; however, it was later decided to collect 2 separate soil samples spaced an appropriate distance apart from each of 8 locations. This ensures repeatability in data, as soil from each location would be run twice (one run per separate sample). A potential bias includes how sampled locations may be skewed closer to the laboratory's location because of the ease of travel to closer locations. Two samples were also used to test the study's workflow to ensure expected data was produced.

The ultimate goal of this study is to identify the levels of human interaction and internal chemical composition for locations sampled. This is because these factors can affect the presence and amount of resistance detected, and which location has the most resistance. The collected data can be used to understand the distribution of resistant bacteria, which could aid in eliminating them and mitigating antibiotic resistance. Based on our review of the literature, there have not been similar studies reported previously. Thus, this study is significant because it provides preliminary information about the distribution of antibiotic resistance in a suburban environment. This information can be useful to future studies attempting to locate and explain the exact origins of antibiotic resistance in suburban environments. Additionally, this information could be used to develop ways to mitigate the levels of resistance genes in suburban environments. This would help combat antibiotic resistance.

Since there is currently great interest in analyzing soil for resistance, the efficacy of this study's workflow was also be analyzed to determine if it could be adopted by similar studies. This could facilitate future studies that involve DNA extraction, as researchers could use the workflow developed in this study.

Results

Samples and Gels

Two soil samples were collected from each of the 8 locations and the experiment was run 16 times. Of the 8 locations, 50% were private and 50% were public. The locations were broadly distributed and had varying characteristics (TABLE 1). Additionally, 2 runs were performed before the experimental runs. This was done to ensure all equipment and processes worked properly, and to adjust the workflow to produce optimal results. 18 runs were conducted in total.

TETB was observed in no runs. TETM was observed in 37.5% of 16 runs [6 runs] and absent in 62.5% of 16 runs [10 runs], so all positive samples contained TETM. These were both BAR, COU, and PEU runs (TABLE 2). The DNA ladder was present in all samples and used to determine the relative sizes of the other bands. 16S was present in 94% of the 16 runs [15 runs] and absent from only the 1st MAU run. For negative and positive controls, negative controls were absent in 94% of runs [15 runs] while present in only the 1st CIR run, where the bands were very faint but still noticeable. Positive controls were present in all samples.

Presence of Resistance

All TETM bands were verified to be present using the linear calibration curves (Figure 1). A percent error calculation found the difference between the observed base pairs of all TETM bands and the actual base pairs of a TETM band were within a magnitude of 10% and therefore insignificant. The actual bp of a TETM band was 406, and the observed bps were: [BAR – 407.314 || COU – 435.400 || PEU – 400.213]. The absolute values of percent errors were: [BAR – 0.324% || COU – 7.241% || PEU – 1.425%] (TABLE 3). A chi-squared test was also used to verify this finding. The calculated chi-squared value was 0.00546, which was less than the critical value of 5.991 ($p > 0.05$). This indicates there were no significant differences between the measured and actual base pairs of the TETM bands. This supported the presence of the TETM band in the BAR, COU, and PEU locations. The confidence interval is (29.425, 46.043), which also indicates there is no significant difference.

Amount of Resistance

The TETM bands, in order of most to least resistance, were the BAR, COU, PEU, and MOU bands. The normalized values for the bands in this order were: [BAR – 1.068 || COU – 0.604 || PEU – 0.253 || MOU – 0.145]. A higher normalized value represents a band of higher density, thus containing greater amounts of TETM (FIGURE 2). The trend in the normalized values is supported by the descending percent area intensities

TABLE 1	Description
BAR	Soil from the backyard lawn of a **private** residence.
CIR	Soil from a parking lot divider of a **private** company.
IAR	Soil from a planter box in the garden of a **private** school.
IBR	Soil from the ground in the garden of a **private** school.
COU	Soil from a parking lot divider of a **public** warehouse.
MAU	Soil from a grassy area of a **public** park and pool.
MOU	Soil from a sidewalk-adjacent soil plot of a **public** road.
PEU	Soil from a sidewalk-adjacent soil plot of a **public** school.

*The grassy area is not occupied by lawn grass and is not the same as a lawn.

TABLE 2	BAR	CIR		IAR	IBR	COU	MAU	MOU	PEU
Run/ Sample	1 + 2	1	2	1 + 2	1 + 2	1 + 2	1	1 + 2	1 + 2
Ladder	+	+	+	+	+	+	+	+	+
16S	+	+	+	+	+	+	–	+	+
TETB	–	–	–	–	–	–	–	–	–
TETM	+	–	–	–	–	+	–	–	+
- Control	–	+	–	–	–	–	–	–	–
+ Control	+	+	+	+	+	+	+	+	+

*+ - Visible in Gel || – - Not Visible in Gel

and areas of the bands [BAR – 31.769%, 48581.350 || COU – 24.957%, 30269.229 || PEU – 4.309%, 9352.442]. To verify these findings, a chi-squared test was used. The calculated chi-squared value was 1.450, which was less than the critical value of 7.815 ($p > 0.05$). This indicates there were no significant differences between the amounts of resistance of the TETM bands. We also calculated the confidence interval to be (0.0507, 0.7673), indicating there is no significant difference. Nevertheless, the BAR band still has the greatest amount of resistance, followed by the COU and PEU bands.

Efficacy of Workflow

The SD and CV for the areas of 16S, TETM, + control TETB/TETM were: [16S – SD: 15798.265, CV: 0.145 —— TETM – SD: 15743.218, CV: 0.627 —— + control TETB – SD: 7876.276, CV: 0.144 —— + control TETM – SD: 8229.150, CV: 0.159] (TABLE 4). The larger deviation in TETM is further supported by the relatively larger SD, and CV of the TETM percent errors, which were: [TETM % Error – SD: 5.696, CV: 8.659].

Discussion

The presence of TETM was verified in both runs of samples from the BAR, COU, and PEU. BAR had the greatest amount of resistance, followed by COU and PEU. Considering locations with TETM, samples from BAR were collected from backyard lawn soil from a private residence and had the most resistance. Lawn soil has limited human interaction and exposure to the surrounding environment. However, lawn soil could contain the greatest amount of resistance because it accumulates more fertilizer, heavy metals, and plastics such as polycyclic aromatic hydrocarbons[16]. These substances are toxic to bacteria. Thus, these bacteria benefit from gaining metal and plastic-resistance genes

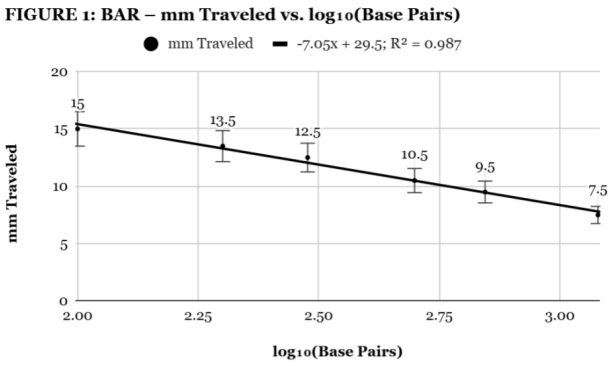


FIGURE 1: BAR – mm Traveled vs. log10(Base Pairs)

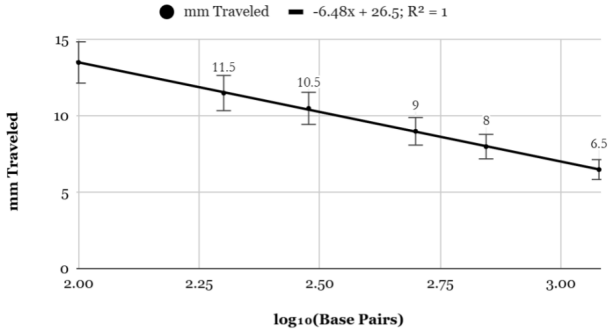


FIGURE 1: COU – mm Traveled vs. log10(Base Pairs)

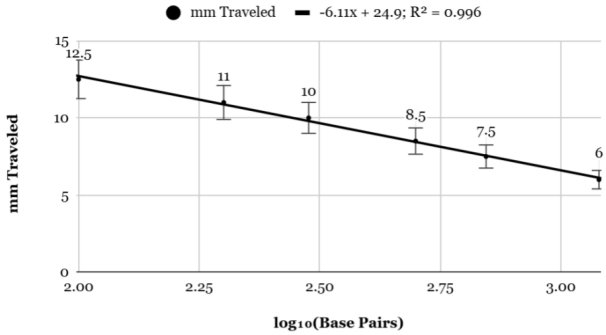


FIGURE 1: PEU – mm Traveled vs. log10(Base Pairs)

through increased horizontal gene transfer (HGF). A byproduct of increased HGF is increased chances bacteria obtain antibiotic resistance genes from the soil. Thus, this data supports that chemical differences in suburban lawn soils increase the likelihood this soil a hotspot for resistant bacteria. However, chemical analysis of suburban lawn soil to verify this relationship was not feasible in this study. Thus, this is a hypothesis that must undergo further testing.

The other resistant samples were collected from public locations with a high frequency of human activity. The frequency of human activity in a location matched the amount of resistance in the location. COU had a higher frequency and greater amount of resistance. PEU had a lower frequency than COU and thus a lower amount of resistance. The other 4 nonresistant samples (CIR, IAR, IBR, MAU, and MOU) were locations with low frequency of human activity and/or private locations. The absence of resistance in these locations suggests the soil there is less likely to have resistance. The overall data shows public suburban areas with a higher frequency of human activity could have greater amounts of resistance. It supports the general finding that areas with more human activity have more resistance [10].

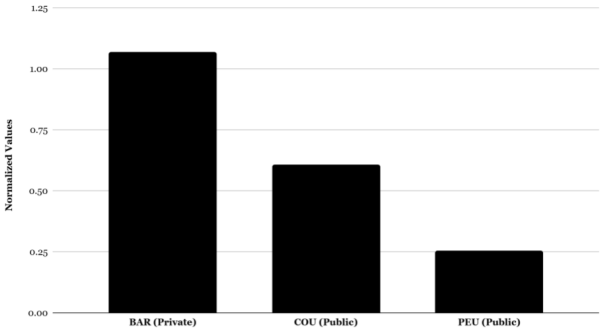
Considering efficacy, all 16S bands except one were present. This indicated reliable DNA extraction and processing and most soil samples contained DNA [17]. The absence of 16S in the first MGP run may indicate no DNA was present in the soil processed or poor DNA extraction. Locations with resistance had TETM for both runs, indicating the study has good repeatability. Negative control bands were absent from all runs except the first CIR run, which may indicate contamination during the study workflow. The presence of all positive control bands indicates the PCR process was reliable. Additionally, the satisfactory presence/absence of appropriate bands indicates the overall workflow is reliable. The relatively low SD and CV $<$ 1.000 between 16S, TETM, and positive control bands indicate minor deviations in areas between bands of each type. This supports the workflow's precision and the study's consistency [18]. As expected, there are greater SD and CV for 16S and TETM bands. This reflects the differing amounts of DNA between samples from different locations and within the same location. This expectation is supported by the relatively larger values from the percent errors calculated while verifying each TETM band. In general, these relatively low values indicate the workflow was effective at eluting DNA within a precise range. However, the workflow still permitted differences between the bands of each type to be detected.

Although this study was successful, it has a few limitations. First, the processes to standardize the amount of DNA eluted were relatively effective. However, they could be overcome if a sample processed contained too much or too little DNA [19]. In future studies, a method should be implemented to quantify the DNA. This would ensure the DNA in each soil sample is constant. This would also improve the workflow's precision

TABLE 3: BAR	Base Pairs	$\log_{10}$(Base Pairs)	mm Traveled	| Percent Error |
TETM (406 bp)	407.3137564	2.609929078	11.1	0.324
Linear Calibration Curve Values				
Base Pairs	**$\log_{10}$(Base Pairs)**	**mm Traveled**	-	
1200	3.079181246	7.5		
700	2.84509804	9.5		
500	2.698970004	10.5		
300	2.477121255	12.5		
200	2.301029996	13.5		
100	2	15		
TABLE 3: COU	**Base Pairs**	**$\log_{10}$(Base Pairs)**	**mm Traveled**	**| Percent Error |**
TETM (406 bp)	435.4004654	2.638888889	9.4	7.241
Linear Calibration Curve Values				
Base Pairs	**$\log_{10}$(Base Pairs)**	**mm Traveled**	-	
1200	3.079181246	6.5		
700	2.84509804	8		
500	2.698970004	9		
300	2.477121255	10.5		
200	2.301029996	11.5		
100	2	13.5		
TABLE 3: PEU	**Base Pairs**	**$\log_{10}$(Base Pairs)**	**mm Traveled**	**| Percent Error |**
TETM (406 bp)	400.2131236	2.602291326	9	1.425
Linear Calibration Curve Values				
Base Pairs	**$\log_{10}$(Base Pairs)**	**mm Traveled**	-	
1200	3.079181246	6		
700	2.84509804	7.5		
500	2.698970004	8.5		
300	2.477121255	10		
200	2.301029996	11		
100	2	12.5		

and decrease the SD. Another limitation includes how the study sampled 8 locations and collected 16 soil samples. While this is

FIGURE 2: Normalized Values vs. Resistant Locations



a large number across an extensive area, there is still a limited sample size. This restricts the finality of these findings and their generalizability to suburban Bay Area locations.

The limited number of samples per location also increases the chances of variability in the results. There could be a large difference in the amounts of resistance between the two samples from a location. This was primarily due to resource limitations, as the study's budget provided enough resources to analyze 18 samples. Furthermore, the study did not have a method of distinguishing between naturally occurring and human-introduced resistance genes in soil samples. Other limitations include how the workflow involved many steps. For instance, DNA extraction and processing relied on manual pipetting to transfer liquids. This could potentially accumulate minor errors throughout the workflow. Mitigating this could involve purchasing automatic pipettes (beyond this study's budget). In general, future studies should ensure a higher budget to secure chemicals and equipment to process more soil from more numerous and varied locations. This could help confirm the trends elucidated in this study.

Despite these limitations, this study met its objectives. The collected data helps better understand resistance in environmental soil from suburban environments. This study expanded the range of locations analyzed for resistance. Most prior studies have analyzed soil from agricultural, rural, or wilderness locations. The collected data shows that resistance in suburban environments is often in line with prior studies' findings. Specifically, the study's data supports a previous finding of how lawn soil tends to contain greater amounts of resistance because it accumulates more fertilizer and toxic materials [9,10,16]. However, this study's findings were different from another similar study [13]. This study found that differences in the sand content, moisture and nutrition in soils could result in specific soils having greater amounts of resistance [13].

This study also analyzed two specific resistance genes, TETB and TETM, while most prior studies analyzed multiple resistance genes. This narrower focus facilitates a comparison of resistance between locations with different characteristics. It establishes specific criteria for environment locations that are hotspots of resistance, especially significant hotspots. These findings can be applied to public health practice by ensuring soil around areas with a high population density doesn't accumulate fertilizers or other toxic substances that could potentially select for resistant bacteria. This could entail using microbial bioremediation to ensure the amount of fertilizer and toxic substances in such soils remains low. Additionally, governments could implement policies that restrict the extensive use of fertilizer in soil around large populations, along with industrial practices that cause an accumulation of fertilizers or other toxic substances in soil. To ensure such policies are followed, governments could regularly test soil near large populations to ensure the amount of fertilizers and toxic substances remains under a threshold.

This study's workflow could also be used in similar studies. The chemicals and equipment are relatively affordable and simple to use. Future studies should continue affirming lawn soil is a significant hotspot of resistance because of its different chemical composition. Affirming this explanation involves testing lawn soil from other suburban residences and locations. Studies could also compare lawn soils' chemical composition to nonlawn soils. They could also analyze the amount of resistance in nonlawn soil after changing its chemical composition to match that of lawn soils. In general, there continues to be limited data on resistance in suburban and urban locations. Specific locations include industrial and commercial zones such as drug factories, hospitals, and wastewater treatment plants [20,21]. This should be a focus of future studies collecting data on antibiotic resistance in the environment.

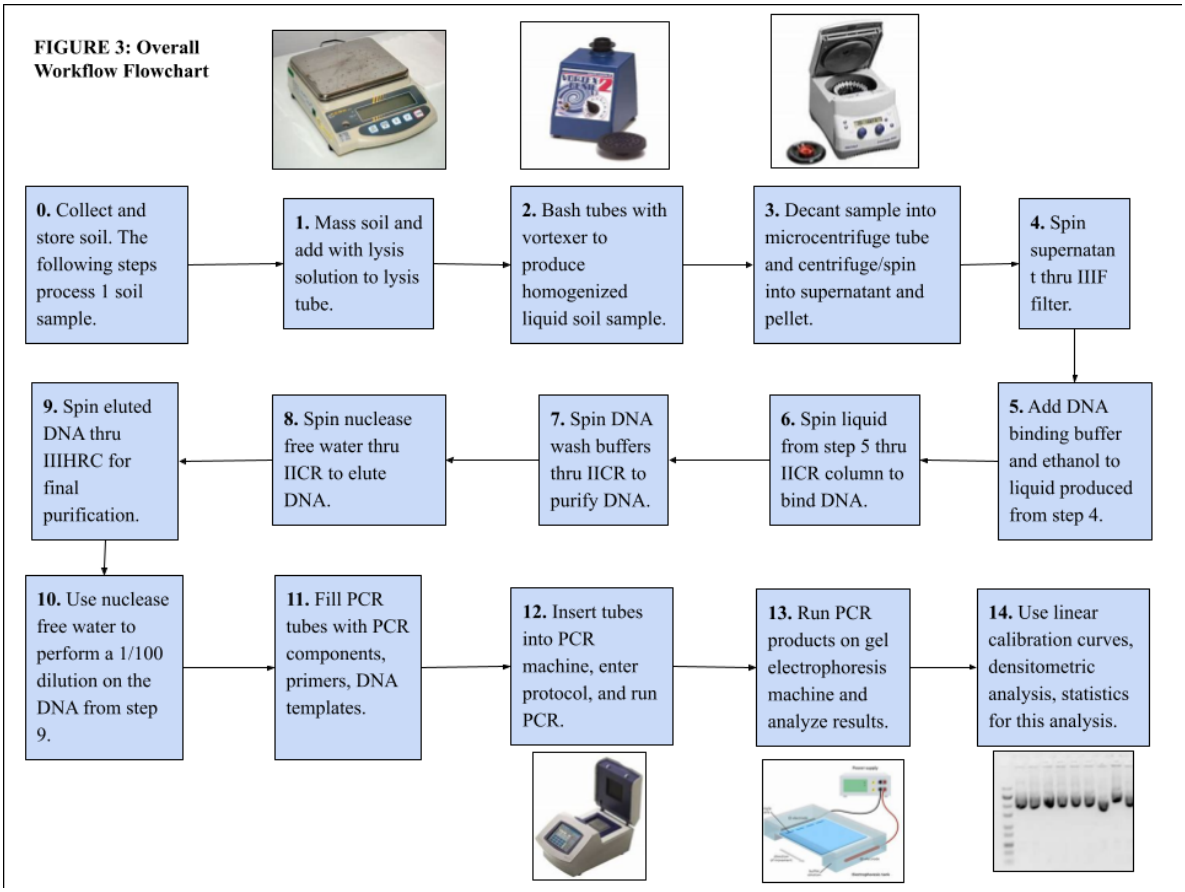
In conclusion, this study's data addresses antibiotic resistance, an important global medical challenge. We hope it helps future studies on the science of antibiotic resistance meet their objectives. To mitigate antibiotic resistance, society must also adapt global healthcare policies to prevent antibiotic overuse and introduction to the environment.

Methods

Samples and Objectives

This experimental study was conducted from 10/23 to 03/24. Soil samples were collected from private and public suburban locations in the Bay Area, California. Inclusion criteria were that sampled locations must be accessible for physically collecting soil and have decent to frequent human traffic. Additionally, soil in the location had a consistent microenvironment: a defined plot of small to moderate size (i.e. soil from a lawn) and exposed to the environment (i.e. soil relatively uncovered). Locations indoors or at extreme distances ($>$ 60 miles) from the laboratory were excluded.

TABLE 4	SD	| Mean |	| CV |
16S	15798.265	108736.711	0.145
TETM	15743.218	25111.375	0.627
TETM % Error	5.696	0.658	8.659
+ Control TETB	7896.277	54727.288	0.144
+ Control TETM	8229.150	51492.358	0.160



Soil Collection and DNA Extraction

A shovel was used to collect soil. We only collected soil from the topsoil (horizon A) to ensure consistency between soil samples. Additionally, for each location we visited, we only sampled planter soil that was strictly used to support plants or used simply for aesthetic purposes. We did not sample soil used for farming or frequent gardening. Additionally, we only collected soil that was not subject to frequent human interaction, besides regular maintenance. Thus we were able to sufficiently control for soil type and land use among our soil samples. After collection, this soil was lightly irrigated with deionized water and mixed in a freezer-grade Ziplock bag. Irrigating and mixing the stock soil ensured the microbes and DNA were spread evenly, so the DNA extracted for each processed soil sample was standardized to a specific range. This also preserved the study's accuracy. It ensured the soil bacteria remained alive, as DNA tended to degrade when bacteria die[22]. Before processing, each soil sample was securely stored in a freezer-grade Ziplock bag inside of another. These bags were located in a cabinet in a dark and cool room to simulate the soil environment.

The genomes of microbes collected in soil samples were extracted using bead-bash lysis with a vortexer. They were then processed using sequential centrifugation in an open-rotor centrifuge; ZymoBIOMICS DNA Miniprep was used in the above process[23]. Bead-bashing lysis is when soil samples are placed in DNA lysis tubes with beads. These tubes were bashed in a vortexer, allowing the beads to lyse microbial cell walls through

physical force[24]. This ensured equitable DNA extraction from samples, which contained microbes of many varieties. The lysis solution helped penetrate cell membranes and maximized the DNA extracted. It also removed contaminants and prevented DNA denaturation[25,26]. Maximizing DNA extraction increased the chances bands were visible after PCR and gel electrophoresis which improved the accuracy of results.

DNA Isolation, Purification, and Elution

Sequential centrifugation isolated extracted DNA and further purified it through filtration. The III-F spin filter (C1057) used first contained a mesh that removes impurities such as soil debris from the supernatant. Ethanol was then added to remove detergents and salts and helped precipitate DNA[27], while DNA binding buffer (D4300) recovered this DNA from the supernatant. The binding buffer contained the chaotropic salt guanidinium thiocyanate (a chemical that improves the migration time of DNA molecules) and glycerol (alcohol that facilitates DNA separation), which helped recover more DNA in less time. This facilitated high-quality DNA extraction[28,29].

The II-CR spin column (C1078) contained silica, a compound that bound DNA as the previous mixture was centrifuged through. After, sequential volumes of DNA wash buffers 1 (D43003 - dilute) and 2 (D43004 - concentrated) were centrifuged through the spin column to thoroughly remove any remaining contaminants. This ensured only DNA was eventually eluted[30]. Lastly, nuclease-free water (W1001) was centrifuged through the spin column to elute DNA[31]. This water had low salt concentrations, which facilitated the elution of DNA as it detached from silica in these conditions. In the overall workflow, standardized spin columns and chemicals of a fixed volume were used to ensure the amount of DNA eluted was within a precise range.

Before PCR, the DNA was centrifuged through an III-HRC spin filter (C1058) prepared with an HRC-preparation solution. This filter has a histidine-rich calcium-binding protein (HRC) matrix to help remove potential PCR and gel electrophoresis inhibitors such as melanin and tannins. These are ubiquitous natural components found in animals and plants[6]. Before the III-HRC spin filter was used, the HRC preparation solution (D4300-7) was centrifuged to prepare the filter matrix. The previously eluted DNA was centrifuged through this spin filter for final DNA purification.

PCR and Gel Electrophoresis

The collected DNA underwent a 1/100 dilution before PCR to ensure PCR wasn't inhibited by excessive DNA[31]. During PCR, 16S, TETB, and TETM primers provided starting points for Taq polymerase to amplify these DNA strands if they were present. These primers were purchased from an outside source. The amplified 16S DNA was a positive control which ensured proper DNA extraction. Nuclease-free water was amplified as a negative control to ensure no contamination occurred and positive control DNA was amplified to ensure proper PCR. After electrophoresis, valid results had the 16S and positive control bands and no negative control bands. After PCR, we observed each gel for the presence/absence of appropriate bands and used the DNA ladder to verify the relative size of each band. This helped determine the reliability of the workflow and contributed data to a study outcome.

Study Outcomes and Measures

The primary outcomes were the presence and amount of resistance in soil samples. Presence was defined by whether TETB/TETM bands were observable in a gel, had the correct size relative to the DNA ladder, and if their size had an insignificant difference compared to the real size of the band. Presence was measured with a linear calibration curve that plotted the distance traveled (mm) vs. $\log_10$ of the base pairs (bp) of the DNA ladder bands for each gel[32]. A coefficient of determination R^2 value was used to determine the accuracy of each curve to ensure it was accurate enough for usage[33]. The curve was used to verify that the measured distance the TETM bands traveled matched the band's expected bp. If a deviation was present, a percent error calculation was used to determine whether the deviation was significant. A percent error within 5-10% indicated the difference between observed and expected bp was minimal enough to verify the presence of a TETB/TETM band.

Only TETB/TETM bands verified to be present were analyzed for the amount of resistance used, which was defined by a band's density. The denser a band was, the more resistance was present, and vice versa. The amount was measured with a densitometric/density analysis of each band in ImageJ software. Notable tools used were the subtraction, invert, plot lanes, and label peaks tools. This analysis examined the area (space a band takes up) and percent area intensity (how saturated a band is) for each TETB/TETM band[34]. The density of each band was normalized. It's percent area intensity value was divided by the percent area intensity value of either the positive control TETB or TETM bands, whichever was appropriate. The normalized values were graphed to identify which sampled location had the most resistance. This also helped determine if there were trends in the amount of resistance between locations based on previously recorded characteristics of the locations.

The secondary outcome was the relative efficacy of the workflow, defined by how precise and accurate bands produced by the workflow were. This indicated the consistency and quality of the workflow, respectively. The standard deviation (SD) and coefficient of variation (CV) were calculated for the areas for each type of band and for the percent errors recorded[18,35]. The data was used to determine the deviation in the amount of

DNA between bands of the same type and the extent of this deviation. A lower standard deviation indicates less variance and greater precision among the amounts of DNA between the same bands, supporting the workflow's consistency. A mean and coefficient of variation (CV) where was also calculated for each type of band, with CVs < 1.000 indicating low variance and CVs > 1.000 indicating high variance. The lower the CV, the less deviation there is in the SD compared to the mean of the data. (SUPPLEMENT 1: detailed procedure; SUPPLEMENT 2: detailed materials).

Acknowledgments

I appreciate Mrs. Nguyen's help with this project.

References

1 *The Pew Charitable Trusts, 'My Worst Fear': A Doctor Faces Antibiotic Resistance*.

2 S. Kraemer, A. Ramachandran and G. Perron, *Antibiotic Pollution in the Environment*, From Microbial Ecology to Public Policy.

3 S. Kunhikannan, C. Thomas, A. Franks, S. Mahadevaiah, S. Kumar and S. Petrovski, *Environmental hotspots for antibiotic resistance genes*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8123917/,.

4 C.D.C., *Antimicrobial Resistance in the Environment and the Food Supply: Causes and How It Spreads*, https://www.cdc.gov/antimicrobial-resistance/causes/environmental-food.html?CDC_AAref_Val=https://www.cdc.gov/drugresistance/environment.html,.

5 A. Jacobs and U.N., *Issues Urgent Warning on the Growing Peril of Drug-Resistant Infections*, https://www.nytimes.com/2019/04/29/health/un-drug-resistance-antibiotics.html,.

6 K. Wu, *Antibiotic-resistant bacteria in farm animals are rising in low- and middle-income countries*, https://www.pbs.org/wgbh/nova/article/animal-antimicrobial-low-middle-income-countries/,.

7 M. Cycon, A. Mrozik and Z. Piotrowska-Seget, *Antibiotics in the Soil Environment—Degradation and Their Impact on Microbial Activity and Diversity*.

8 E. Yee, S. Cheng, G. Knappe and C. A, *Moomau, Antibiotic resistance: How to prevent the next public health emergency*, https://biology.mit.edu/antibiotic-resistance-how-to-prevent-the-next-public-health-emergency/,.

9 D. Zhu, J. Ma, G. Li, M. Rillig and Y. Zhu, *Soil plastispheres as hotspots of antibiotic resistance genes and potential pathogens*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8776808/,.

10 K. Osbiton, A. Oxbrough and L. Fernandez-Martínez, *Antibiotic resistance levels in soils from urban and rural land uses in Great Britain*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8115975/.

11 H. Shi, X. Hu, W. Li, J. Zhang, B. Hu and L. Lou, *Soil Component: A Potential Factor Affecting the Occurrence and Spread of Antibiotic Resistance Genes*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9952537/,.

12 T. Wang, S. Sun, Y. Xu, M. Waigi, E. Odinga, G. Vasilyeva, Y. Gao and X. Hu, *Nitrogen Regulates the Distribution of Antibiotic Resistance Genes in the Soil–Vegetable System*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8964294/,.

13 S. Peng, D. Song, B. Zhou, Q. Hua, X. Lin and Y. Wang, *Persistence of Salmonella Typhimurium and antibiotic resistance genes in different types of soil influenced by flooding and soil properties*, https://pubmed.ncbi.nlm.nih.gov/36436254/,.

14 I. Kampouris, S. Agrawal, L. Orschler, D. Cacace, S. Kunze and T. U. Klumper, *Antibiotic resistance gene load and irrigation intensity determine the impact of wastewater irrigation on antimicrobial resistance in the soil microbiome*, https://pubmed.ncbi.nlm.nih.gov/33571903/,.

15 Q. Xiang, C. Fu, C. Lu, A. Sun, Q. Chen and M. Qiao, *Flooding drives the temporal turnover of antibiotic resistance gene in manure-amended soil–water continuum*, https://pubmed.ncbi.nlm.nih.gov/37647704/,.

16 C. Knapp, A. Callan, B. Aitken, R. Shearn, A. Koenders and A. Hinwood, *Relationship between antibiotic resistance genes and metals in residential soil samples from Western Australia*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5340841/,.

17 B. Hornung and R. E. Kuijper, *Issues and current standards of controls in microbiome research*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6469980/,.

18 Z. Bobbit, *What is Considered a Good Coefficient of Variation?*, https://www.statology.org/what-is-a-good-coefficient-of-variation/,.

19 N. Ali, R. Rampazzo, A. Costa and M. Kriger, *Current Nucleic Acid Extraction Methods and Their Implications to Point-of-Care Diagnostics*.

20 I. Samreen, H. Malak and H. Abulreesh, *Environmental antimicrobial resistance and its drivers: a potential threat to public health*, https://pubmed.ncbi.nlm.nih.gov/34454098/,.

21 Y. Zhu, Q. Zhang, J. Xu, Q. Qu, T. Lu, B. Du, M. Ke, M. Zhang and H. Qian, *Changes in bacterial community structure and antibiotic resistance genes in soil in the vicinity of a pharmaceutical factory*, https://pubmed.ncbi.nlm.nih.gov/29660617/,.

22 J. Hallsworth, *Water is a preservative of microbes*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8719826/,.

23 *Zymo Research, ZymoBIOMICS DNA Miniprep Kit*, https://www.zymoresearch.com/products/zymobiomics-dna-miniprep-kit,.

24 P. Vandeventer, K. Weigel, J. Salazar, B. Ewin, B. Irvine, R. Doebler, A. Nadim, G. Cangelosi and A. Niemz, *Disposable Device*.

25 *Thermo Fisher, Detergents for Cell Lysis and Protein Extraction*, https://www.thermofisher.com/us/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/detergents-cell-lysis-protein-extraction.html,.

26 S. El-Ashram, I. Nasr and X. Suo, *Nucleic acid protocols: Extraction and optimization*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5361071/.

27 J. Brennan, *What Does Ethanol Do in a DNA Extraction?*, https://sciencing.com/ethanol-do-dna-extraction-8336005.html,.

28 C. Shen, *Academic Press*, **1**, 143–166,.

29 A. Diaz, H. Jothiraman and V. Ramakrishnan, *Journal of Molecular Graphics and Modelling*, **144**, year.

30 Y. Tang and C. Stratton, *Nucleic Acid Extraction Techniques*.

31 *Thermo Fisher, Nuclease-Free Water, Nuclease-Free Water*, https://www.thermofisher.com/us/en/home/life-science/dna-rna-purification-analysis/rna-extraction/rna-extraction-products/nuclease-free-tubes-tips-and-buffers/nuclease-free-water.html,.

32 *Biomedical and Life Sciences Simplified, How to calculate the size of a DNA band on a gel?*, https://www.youtube.com/watch?v=z8Hz2WNnGY4&t=133s,.

33 S. Turney, *Coefficient of Determination (R2*, https://www.statology.org/what-is-a-good-coefficient-of-variation/,.

34 nrtTAYE, *Quantification or densitometric analysis of agarose gel electrophoresis bands using ImageJ*, https://www.youtube.com/watch?v=d1MZ08HWhlU,.

35 P. Bhandari, *How to Calculate Standard Deviation*, https://www.statology.org/what-is-a-good-coefficient-of-variation/,.

SUPPLEMENT 1: Study Detailed Procedure

Study Data Collection

Upon arrival at a location, indicate whether it's private (R-access for limited individuals) or public (U-access for all individuals). We recorded either U or R at the end of the three-letter code describing the location; the first two letters were from the location's name. For DNA extraction of one sample, we massed 25 mg soil with a sensitive balance and used a metal spatula to add the soil to a 2 mL lysis tube with silicon lysis beads. We then added 750 μl lysis solution and bashed the lysis tube in a vortexer with the proper attachment for 40 minutes. Afterwards, we decanted the homogenate into a microcentrifuge tube (all tubes used were 1.5 mL) and centrifuged this tube for 3 minutes (open rotor centrifuge at maximum speed) to separate homogenate into pellet and supernatant. We pipetted 400 μl supernatant into an III-F spin filter with a collection tube, and centrifuged for 1 minute. After, we added 800 μl DNA binding buffer and 400 μl 95% ethanol to the collection tube. We then pipetted up and down 5-10 times with a 1000 μl pipette to mix the liquid. Afterward, we transferred 800 μl of this liquid to an II-CR spin column with a collection tube and centrifuged for 1 minute, then discarded the liquid in the collection tube and repeated this process for the remaining mixture. This isolated the DNA on the spin column.

We pipetted 400 μl DNA wash buffer 1 (dilute) into the spin column and centrifuged for 1 minute, then discarded the collected liquid. We repeated the above process with 700 μl DNA wash buffer 1 then 200 μl DNA wash buffer 2 (concentrated). For the next step, we transferred the spin column into a microcentrifuge tube (labeled "F"), pipetted in 100 μl nuclease-free water, and centrifuged for 1 minute to elute DNA into the tube. To remove final contaminants, we prepared an III-HRC filter and added another microcentrifuge tube (labeled "FF") to the bottom of the HRC. We then pipetted the liquid from tube F into the HRC and centrifuged for 3 minutes to isolate DNA into tube FF. We prepared another microcentrifuge tube (labeled "1/100"), pipetted in 198 μl nuclease-free water, then pipetted in 2 μl DNA from tube FF to create a 1/100 DNA dilution in tube 1/100. Then, we prepared 20 μl PCR tubes:

Tubes	**1**	**2**	**3**	**4**	**5**	**6**	**7**
5x PCR mix	5 µl for each tube						
18 µl primer	*16S*	*TETB*	*TETM*	*TETB*	*TETM*	*TETB*	*TETM*
2 µl DNA	1/100 sample			Nuclease free water		Positive control	

We entered the following programming for 38 cycles of PCR:

Stage	**Initial denaturation**	**Denaturation**	**Annealing**	**Extension**	**Final extension**
Temperature (°C)	94	94	61	72	72
Time (s)	120	20	20	40	60

We then casted agarose gels by combining 1 agarose tablet with 25 mL 1x TBE buffer in an Erlenmeyer flask. We poured the heated mixture into a casting tray to create gels. Afterwards, we loaded samples for 25 minutes of gel electrophoresis:

Lane	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Sample (µl)	10 µl DNA ladder	15 µl tube 1	15 µl tube 2	15 µl tube 3	15 µl tube 4	15 µl tube 5	15 µl tube 6	15 µl tube 7

Study Data Analysis

We observed all gels for 16S and positive control DNA bands and ensured their base pairs (bps) were accurate relative to the present DNA ladder (FIGURE 1). The one band that appeared in TETB/TETM lanes must have had an accurate relative size to suggest the presence of resistance; additionally, negative control bands must be absent. Primer dimers and nonspecific bands from PCR may occur but didn't invalidate results.

A linear calibration curve [distance traveled (mm) vs the $\log_{10}$ of the base pairs (bp)] was plotted to verify the presence of each TETB/TETM band. To create the calibration curve, the bp and $\log_{10}$bp for each band in the DNA ladder were recorded in descending order. Then, the amount each DNA ladder band traveled was determined. We measured the distance between the well's center and the band's center. The distance each TETB/TETM band traveled was measured, converted into $\log_{10}$bp using the unique calibration curve created for each gel, then converted into bp using the antilog. Each chart received error bars for reference, a trendline, and an R^2 value. A percent error calculation was used to measure the deviation between the observed and expected bp of a band.

A densitometric analysis using ImageJ (Windows-64) software was used to determine each recorded band's density after verifying its presence. In TETB/TETM bands, the density of a band correlated with the amount of DNA/resistance present. Images of each agarose gel were converted into 8-bit. The subtraction tool was used to eliminate the background noise on the image of the gel. The invert tool was then used to change the color of the bands from white to black to make them more visible. Next, the "plot lanes" tool was used to create a wave-like plot that represented the selected bands. The line tool was used to connect the first bottom corners of each wave. Each completed wave represented a band on the gel. The wand tool was used to record the area of each complete wave. Then, the "label peaks" tool was used to record the percent area intensity of each complete wave. To generate normalized values, the percent area intensity of the TETB/TETM band was divided by the percent area intensity of the corresponding positive control band. The greater the normalized value, the denser the band and the greater the amount of resistance.

The normalized values for all TETB/TETM bands were graphed to identify trends in the amount of resistance across positive locations. To identify trends, we looked for commonalities in the recorded characteristics of positive locations and searched through any prior literature to determine if related studies could affirm or provide new explanations for trends we observed.

To determine the relative efficacy of the workflow, the above process was used for all non-TETB/TETM bands (except DNA ladder bands) to record their areas. Additionally, the standard deviation (SD) between the areas of each type of band was calculated. The SD and coefficient of variation (CV) of the percent errors recorded when verifying the presence of TETB/TETM bands were also calculated. This helped us determine the efficacy of the workflow in producing bands of correct bps.

SUPPLEMENT 2: Study Detailed Materials

General Laboratory Materials

- Beaker, Erlenmeyer flask, graduated cylinder
- Deionized water and bottle
- Gloves and safety goggles
- Lab tray with tray covers
- Metal spatula, sensitive balance, weight boats
- Micropipettes and tips
- Nuclease-free water
- Sharpie
- Shovel and Ziplock bags

DNA Extraction Materials

- Soil sample

- Lysis tube with lysis beads
- Lysis solution
- Vortexer

DNA Processing Materials

- Spin filter
- Spin column
- HRC filter
- DNA binding buffer
- 95% ethanol
- DNA wash buffers
- HRC preparation solution
- Centrifuge

PCR Operation Materials

- PCR tubes
- 5x PCR mix with Taq pol, nucleotides
- *16S*, *TETB*, *TETM* primers
- 1/100 DNA
- Positive control DNA
- PCR machine

Gel Electrophoresis Materials

- Agarose tables
- 1x TBE buffer
- DNA ladder
- Casting tray
- Electrophoresis machine