Title

Characterization of Anthropogenic Influence and Levels of Antibiotic Resistance Gene, erm(F), in Soil at 26 Public Parks in Four Cities and Two Pristine Sites in California

Author

Sun, Fangfang

Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Characterization of Anthropogenic Influence and Levels of Antibiotic Resistance Gene, *erm*(F), in Soil at 26 Public Parks in Four Cities and Two Pristine Sites in California

A thesis submitted in partial satisfaction

of the requirements for the degree Master of Science

in Civil Engineering

by

Fangfang Sun

2016

ABSTRACT OF THE THESIS

Characterization of Anthropogenic Influence and Levels of Antibiotic Resistance Gene, *erm*(F), in Soil at 26 Public Parks in Four Cities and Two Pristine Sites in California

by

Fangfang Sun

Master of Science in Civil Engineering

University of California, Los Angeles, 2016

Professor Jennifer Ayala Jay, Chair

The goal of this study was to characterize the soil constitution and the presence of antibiotic resistance gene (ARG) *erm*(F) in soils across California to investigate the relationship between the levels of anthropogenic influence and ARG *erm*(F) levels locally and regionally. Soil samples were collected from 26 parks in four cities, San Diego, Los Angeles, Bakersfield, Fresno, where six publicly accessible parks sites were selected in each, and two pristine sites, Yosemite National Park and Mount Baldy. DNA was extracted, normalized, and tested for the presence and quantity of genes 16S rRNA and *erm*(F) using quantitative polymerase chain reaction (qPCR). *Erm*(F) quantities were determined via two analysis

approaches—on a per gram of soil basis, and relative to 16S gene abundance. Locally among 26 parks, the five quantifiably detected *erm*(F) sites all occurred in high-level anthropogenic influenced sites. Regionally among four cities, urban and agricultural cities shared the similar soil constitution respectively and urban cities had higher *erm*(F) gene levels than agricultural ones. Population and population density in each city is positively correlated with erm(F) gene levels whichever analysis approach was applied.

The thesis of Fangfang Sun is approved.

Michael K. Stenstrom

Shaily Mahendra

Jennifer Ayala Jay, Committee Chair

University of California, Los Angeles

2016

Contents

Introduction

One of the first used antibiotics was penicillin, which was discovered by Alexander Fleming in 1928 and entered mass production in the early 1940s (Bergstrom & Feldgarden 2008). Antibiotics are thought to be the most successful family of drugs developed for human disease treatment thus far and are widely used to prevent or treat animal diseases in livestock production (Wang et al., 2014).

However, the overuse, or misuse, of antibiotics in medical and agricultural operations can play a role in increased antibiotic resistance (Knapp et al., 2011). In manure-amended soils or agricultural soils with a history of antibiotic use, bacteria were detected with a higher Minimal Inhibitory Concentration (MIC), while non-manure-amended soils yielded larger proportions of bacteria that had lower MICs, carried fewer antibiotic resistance genes, and did not display multidrug resistance (Popowska et al., 2012). Moreover, extensive horizontal gene transfer of ARGs occurred in the colon in human body and resulted in an increase in transfer of ARGs from farm to the human intestinal tract (Salyers et al., 2006). Furthermore ARGs are persistent in the environment. *Tet* and *erm* genes in the waste-impacted farm soil were still detected two years after waste was removed from the farm (You et al. 2012).

The World Health Organization has identified multidrug resistance around the world as a public health priority since the year of 2011 (Nardelli et al., 2012). The spread of antibiotic-resistant bacteria is a growing public health threat in the U. S.

and around the world. Antibiotic resistance genes (ARGs) are now considered emerging environmental contaminants and have been measured in several environmental compartments (Pruden et al., 2006).

Further, ARGs in soil have been shown to be related not only to the direct application of antibiotics, but also to anthropogenic influences in the environment. ARG *sul*1 occurrence frequency was significantly and positively related to the level of anthropogenic influence, although the other genes (*qacE1*/*qacEΔ1*, *intI*1, *ISCR1* and *tniC* genes) were not in a model of a culture-based method in a gradient of anthropogenic disturbances in a Patagonian island in South America (Nardelli et al., 2012).

To date there is limited research to characterize the relationship between the ARGs in soil and the anthropogenic influence levels in California. The hypothesis in this paper is that parks locally or cities regionally in California with a higher anthropogenic influence level are more likely to have higher *erm*(F) gene levels.

In this study *erm*(F) and 16S rRNA genes were selected for gene copies detection through qPCR. The *erm* genes are among the most common antibiotic resistance genes of macrolides-lincosamides-streptogramin B (MLSB) and are the most commonly acquired resistance genes in bacteria (Roberts, 2004; Chen et al., 2007). 16S rRNA is a specific signature sequence that is present in all bacteria, serving as a useful marker in the identification of bacterial presence (Woo et al., 2008).

Material and Methods

This study was conducted at 26 public parks in four cities and two pristine areas of California, USA. Cities sampled consisted of Los Angeles, San Diego, Fresno, and Bakersfield while pristine sites included Yosemite National Park and Mount Baldy. Six public parks were selected within each city, with each park classified to three levels of anthropogenic influence (Table 1, and see Appendix B for park name and geographical coordinates for each site).

The level of anthropogenic influence was quantitatively estimated by counting the number of buildings and roads that were contained in a 500-meter-radius area around each site (Nardelli et al., 2012). These estimates were conducted using satellite images provided by Google Earth (See Appendix A for aerial photos in each site). This paper defined high-level anthropogenic influence as ≥ 75 buildings and ≥15 roads, and low-level as <50 buildings and <10 roads. Anthropogenic influence levels that fell between these criteria were classified as mid-level. When considering buildings that fell on the circumference line, if over half the building fell within the radius of analysis, the building was included for anthropogenic influence estimation; otherwise, it was omitted from the analysis.

Table 1 Name, latitude, longitude, roads, buildings and urbanization degree of different soil sites. See Appendix B for numbers of roads and buildings at each site and geographical coordinates.

City	Site ID	Anthropogenic Influence Category	City	Site ID	Anthropogenic Influence Category
Los Angeles	1	low	Fresno	1	low
	2	low		2	high
	3	high		3	low
	4	high		4	high
	5	high		5	high
	6	mid		6	high
San Diego	1	high	Bakersfield	1	high
	2	mid		2	high
	3	high		3	high
	4	high		4	high
	5	low		5	low
	6	low		6	low
Yosemite		mid	Mount Baldy		low

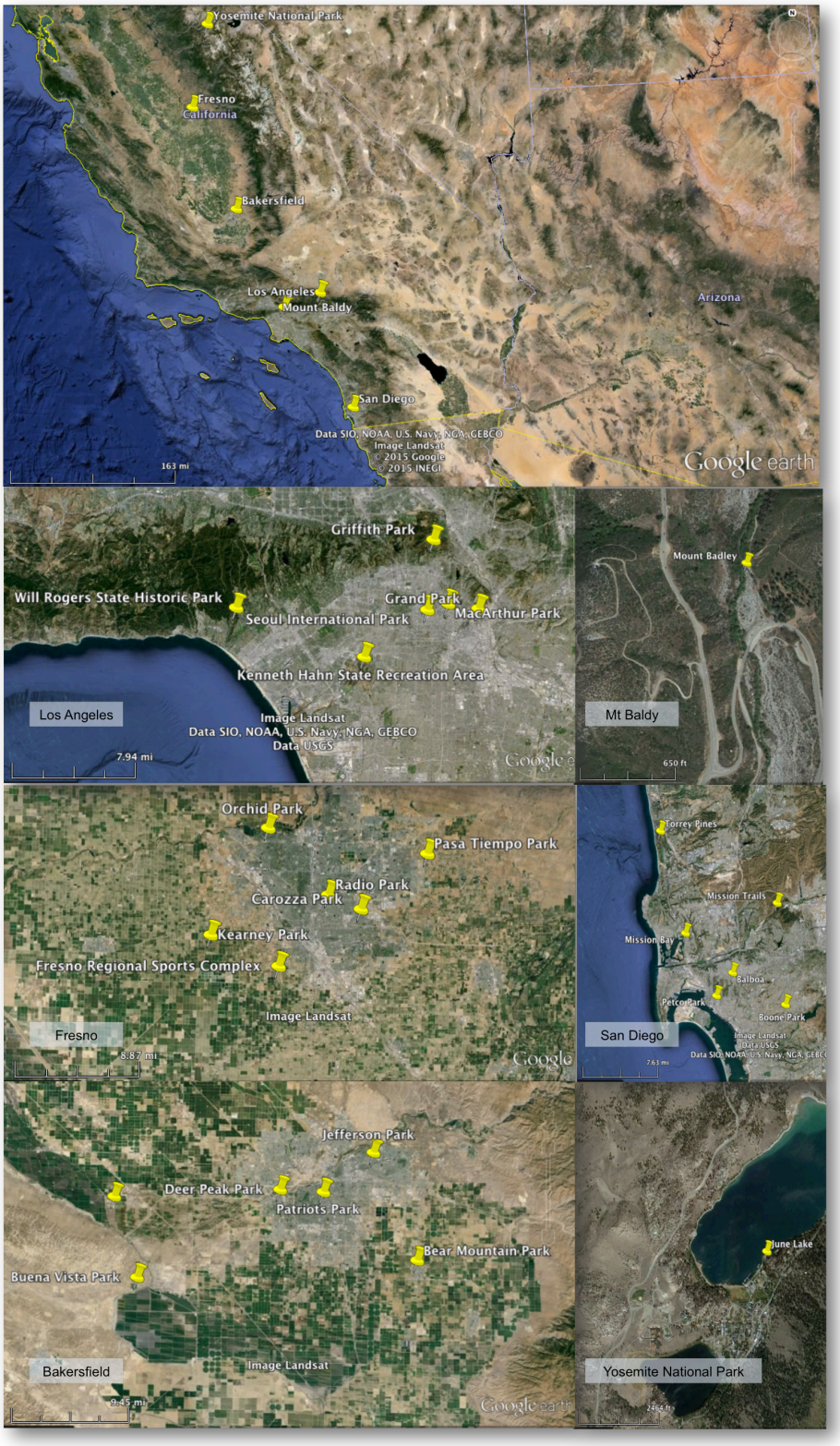


Figure 1 Sampling locations

Sampling Techniques

Sample Collection

When collecting soil samples, each site had three separate 50 mL samples gathered from three one-meter square areas within the selected sample site. Ten points were then chosen at random within the one-meter square areas and 5mL was collected from each of those ten locations. This sampling protocol ensured that the samples were representative of the site and also supplied the site triplicates with an accurate standard deviation. Samples were preserved with ice bags once they were collected and were stored in cold room at 4°C when arriving at the lab.

Molecular Analysis

DNA Extraction

Each 50mL Falcon tube containing soil samples was mixed for five minutes. 0.25 grams of soil were added into prepared beaded tubes, which contained one gram of 0.7mm garnet beads. Then MoBio PowerFecal DNA extraction kit was used for DNA extraction.

After DNA extraction, 100μL solution containing all DNA in each sample was preserved in freezer for following analysis.

DNA Normalization

To make sure that a consistent amount of template was used for the qPCR

reactions, DNA concentrations of each sample were made uniform to a working concentration 0.25ng/μL.

NanoDrop 2000 was used to determine initial DNA concentrations in each sample. The difference between two trials was no bigger than 0.4 ng/μL to ensure the representativeness of the readings. If the difference between two trials was bigger than 0.4ng/μL, third and fourth trials were applied and two most close readings were recorded for future analysis.

Normalization was performed on a sterile workbench with UV disinfection to keep the minimal contamination in dilution process. After dilution, 0.25ng/μL working concentration solution was allocated into different 1.7mL labeled microcentrifuge tubes for future use.

Quantitative Polymerase Chain Reaction

After DNA extraction and normalization, samples were analyzed by quantitative Polymerase Chain Reaction (qPCR). A 96-well plate was prepared. A 25μL volume was added into each well, including 12.5μL SYBR Green I (Thermo Fisher Scientific Inc., MA), 1.25μL for forward and reverse primers, and 10μL for samples on working concentration. Forward and reverse primers were used to detect the target genes.

Table 2 Primers used for qPCR on 16S, *erm*(F) and the concentration of *erm*(F) primers

Gene	Forward Primer	Reverse Primer
16S	GCGGACGGGTGAGTAATGT	TCATCCTCTCAGACCAGCTA
erm(F)	TCGTTTTACGGGTCAGCACTT	CAACCAAAGCTGTGTCGTTT
Primer concentration of *erm*(F)	500nM	500nM

For the detection of 16S gene and *erm(*F) gene, the StepOnePlus Real Time PCR System (Thermo Fisher Scientific Inc., MA) was applied. The heat cycle to detect 16S gene was: 10 minutes of 95°C and 40 cycles of 95°C for 15 seconds and 60°C for 1 minutes.

The heat cycles for *erm(*F) gene detection: 95°C for 10 minutes and 45 cycles at 94°C for 20 seconds and 60°C for 30 seconds. And the melt curve for reaction quality verification was 50-95°C, ΔT = 0.1 °C/second (Knapp et al., 2010).

Soil Characterization

Total Solids

Preparation of evaporating dishes: clean evaporating dishes and watch glasses were ignited at 550°C for one hour in a muffle furnace. The dried equipment was cooled and stored in a desiccator. Every dish and watch glass was weighed before use.

A 5g allocation of the pulverized sample was placed on a prepared evaporating dish. Each sample was covered with a watch glass, and was weighed. The samples were dried at 103°C to 105°C for at least twelve hours, and were cooled to balance temperature in an individual desiccator containing fresh desiccant. And the sample was weighed. The residue was heated for one hour, and then was cooled to balance temperature in a desiccator, and was weighed again. This heating, cooling, desiccating, and weighing procedure was repeated until the weight change was less than 4% or 50 mg, whichever was less. The final weight was recorded (Environmental Protection Agency [EPA], 2001).

Fixed and volatile solids

The evaporating dishes containing the dried residues were transferred to a cool muffle furnace. The furnace was heated to 550°C and was ignited for 2 hours. The residue in a desiccator was cooled to balance the temperature. The residues were weighed. Igniting, cooling, desiccating, and weighing steps were repeated until

the weight change was less than 4% or 50 mg, whichever was less. The final weight was recorded (EPA, 2001).

$$\% \text{ total solids} = \frac{(Wtotal - Wdish) * 100}{Wsample - Wdish}$$

$$\% \text{ fixed solids} = \frac{(Wvolatile - Wdish) * 100}{Wtotal - Wdish}$$

Where, W_{dish}=Weight of dish (mg); W_{sample}=Weight of wet sample and dish (mg); W_{total}=Weight of dried residue and dish (mg)

Sediment Grain Analysis

Samples must be frozen since collection, but they were dried for Sediment Grain Analysis. Enough wet sample was weighed, in a previously weighed Petri dish, to obtain at least 50g dry weight of sample. Wet samples were placed on Petri dishes in the drying oven at 70°C and were recorded as 'initial weight'. Samples were considered dry when weighing on at least two different occasions yields a maximum difference of less than 0.2g. Weight of dried sample was determined by subtracting the dry weight of the Petri dish from the weight of the sample and Petri dish, and was recorded as 'final weight'.

A mortar and pestle was applied and the dried sample was carefully crushed until sample bits were approximately 3 mm in diameter. Things that were not soil, such as shells and sticks were picked out with forceps. 50 grams of the moderately crushed sample were placed into a 600mL beaker, and the beaker was labeled. This step was repeated for all samples.

Preparation of sodium metaphosphate solution: 5g of sodium metaphosphate were added per 100mL of deionized water (DI) water. 1L of solution was used for 10 samples. When adding sodium metaphosphate, the water was mixed continuously and vigorously. Only small incremental amounts of chemical were added, allowing each to fully dissolve before adding the next increment. After the sodium metaphosphate solution was prepared, 100mL was added to each sample beaker. Approximately 200mL DI was added to each beaker. The beakers were placed on the shaker table at 125 rpm for at least 24 hours.

Contents of beaker were transferred to a 1L cylinder (marked at the 1L level), and DI water was used in a wash bottle to break up and wash all the particles out of the beaker into the cylinder. The label was taken from each beaker and was placed on its respective cylinder. DI water was added from the room-temperature jugs into each cylinder until the total volume in each cylinder was 1L.

A blank was made by adding 5g sodium metaphosphate to 1L of DI water, mixing vigorously until all the clumps have dissolved and was poured into a 1L cylinder.

The water temperature was taken in the first cylinder (temperatures must be between 16.5-24.4 °C for results to be accurate). The thermometer was rinsed with DI water after each reading.

Parafilm was placed over the top of the first cylinder. The contents were mixed thoroughly by inverting the cylinder several times until all of the soil was suspended in solution.

Immediately after placing the cylinder back down on the countertop, timing was started at an interval of 40 seconds. The Parafilm was removed and the

hydrometer was gently lowered into the cylinder. The hydrometer reading was recorded 40 seconds after setting back down on the countertop.

The hydrometer was rinsed with DI water between each cylinder reading. Once a sample reading has been taken, the cylinder could not be shaken again before the 2 hours reading. If the sample was shaken again, then the procedure had to be repeated.

After letting the cylinders sit undisturbed for 2 hours, the hydrometer reading was taken and the hydrometer was rinsed off between samples. The temperatures were recorded again for each sample.

When calculating grain size, these equations were needed:

- Soil moisture percent (%)=((initial weight – final weight)*100)/(initial weight)
- Corrected hydrometer reading=hydrometer reading - hydrometer reading of the blank
- TAHR=[(temperature-20°C)*0.35]+corrected reading

  Where, TAHR means Temperature Adjusted Hydrometer Reading.
- TAHR at 40 seconds *Volume (L)/grams of dry soil= % silt and clay
- 100 - (%silt and clay) =% sand
- (TAHR at 2 hours *volume (L)/grams of dry soil = % clay
- (% silt and clay) - % clay = % silt

Data Analysis

The qPCR plates were matched to the samples and a graph was created for visual illustration of the amplification of each well in the plate. A threshold line was set at 0.03, high enough to distinguish background noise from the qPCR test result and low enough for the intersection of DNA amplification curves and the threshold line.

The CT values of 16S rRNA were taken for relative data analysis. 16S CT values were normalized to 100/2n, in which n was equal to the average of 16S CT values subtracted by the lowest CT for that gene for all samples in all cities.

CT values of *erm*(F) were converted through standard curve into quantities for averages and standard deviations calculations. Since each site was tested in triplicate, presence was only counted if 2/3 of the samples had ARG *erm*(F) quantifiably detected. The presence of *erm*(F) in one site was tested first and if existing, the average of triplicates in the site was calculated and was converted to *erm*(F) gene copies per gram soil. For 26 sites, bar graphs with error bars were created to show the quantity of genes per gram soil present from each site.

The average of triplicates in four city sites from qPCR assays was also normalized to the relative amount of 16S rRNA gene copies, thus to be normalized to the total bacterial community. This provided an approach to correcting for potential variations in DNA extraction efficiencies (Pruden et al.,2006). For 24 sites in four cities, bar graphs with error bars were created to show the quantity of genes normalized to relative amount of 16S rRNA.

Afterwards, the city average *erm*(F) copies of six sites in two analysis approaches were applied for a regional anthropogenic influence comparison among four cities on a regional scale. Soil constitution in grain size and fixed and volatile solids was presented with pie graphs. The population and population density in each city were cited from U.S. Census Bureau and dot graphs between *erm*(F) gene levels in two analysis approaches and population or population density were plotted, and simple linear regression method was performed for analysis.

Results

Four cities, Los Angeles, San Diego, Fresno, Bakersfield and two pristine sites, Yosemite National Park and Mount Baldy were included in this study. Table 3 describes the results on grain size analysis, total solids and fixed and volatile factions. For soil conditions, Los Angeles and San Diego, the two urban cities had similar soil contents in the percentage of sand, clay, silt, soil moisture and total, fixed, volatile solids while Bakersfield and Fresno, the two agricultural cities were similar. Yosemite National Park and Mount Baldy, which were assumed as pristine sites shared the similar background information on the percentage of sand, clay, silt, and total solids but differed in soil moisture and fixed and volatile solids.

Table 3 Parameters on soil characterization

	Grain Size Analysis				Total, Fixed and Volatile Solids		
	Sand (%)	**Clay (%)**	**Silt (%)**	**Soil Moisture (%)**	**Total solids (%)**	**Fixed faction (%)**	**Volatile faction (%)**
Los Angeles	74.50	11.60	14.00	10.00	93.33	81.57	18.43
San Diego	74.50	11.20	14.30	11.50	89.96	86.26	13.74
Fresno	64.30	11.70	24.00	14.30	87.85	92.55	7.45
Bakersfield	68.00	9.90	22.20	17.60	85.14	93.43	6.57
Mt Baldy	81.50	8.60	10.00	1.90	98.18	96.40	3.60
Yosemite	85.50	8.60	7.00	4.60	95.65	86.37	13.63

Table 4 showed *erm*(F) detectable and quantifiable readings in soils. Each site was tested in triplicate, presence was only counted if 2/3 of the samples had the gene of interest was detected. In Los Angeles, there were three sites detected and quantified on ARG *erm*(F); in San Diego, there was two sites detected but only one site was quantifiable on *erm*(F); in Fresno, there was one detectable and quantifiable site; in Bakersfield, one site was detected but no site was quantifiable; in pristine sties, Yosemite National Park and Mount Baldy were not detected on *erm*(F).

Table 4 *erm*(F) detectable and quantifiable readings in soils

Sites in Los Angeles	Detectable Readings	Quantifiable Readings
1	3/9	2/9
2	3/9	3/9
3	9/9	9/9
4	9/9	8/9
5	5/9	5/9
6	9/9	8/9
Total Sites	**3**	**3**
Number of Occurrences	**38/54**	**35/54**
Sites in San Diego	**Detectable Readings**	**Quantifiable Readings**
1	3/9	3/9
2	4/9	3/9
3	7/9	4/9
4	8/9	6/9
5	4/9	2/9
6	3/9	3/9
Total Sites	**2**	**1**
Number of Occurrences	**29/54**	**21/54**
Sites in Fresno	**Detectable Readings**	**Quantifiable Readings**
1	2/9	2/9
2	5/9	4/9
3	5/9	4/9
4	5/9	5/9
5	6/9	6/9
6	5/9	5/9

Total Sites	**1**	**1**
Number of Occurrences	**28/54**	**26/54**
Sites in Bakersfield	**Detectable Readings**	**Quantifiable Readings**
1	1/9	0/9
2	5/9	4/9
3	1/9	1/9
4	7/9	4/9
5	3/9	2/9
6	5/9	4/9
Total Sites	**1**	**0**
Number of Occurrences	**22/54**	**15/54**
Pristine sites	**Detectable Readings**	**Quantifiable Readings**
Yosemite National Park	4/9	3/9
Mt Baldy	3/9	2/9
Total Sites	**0**	**0**
Number of Occurrences	**7/18**	**5/18**

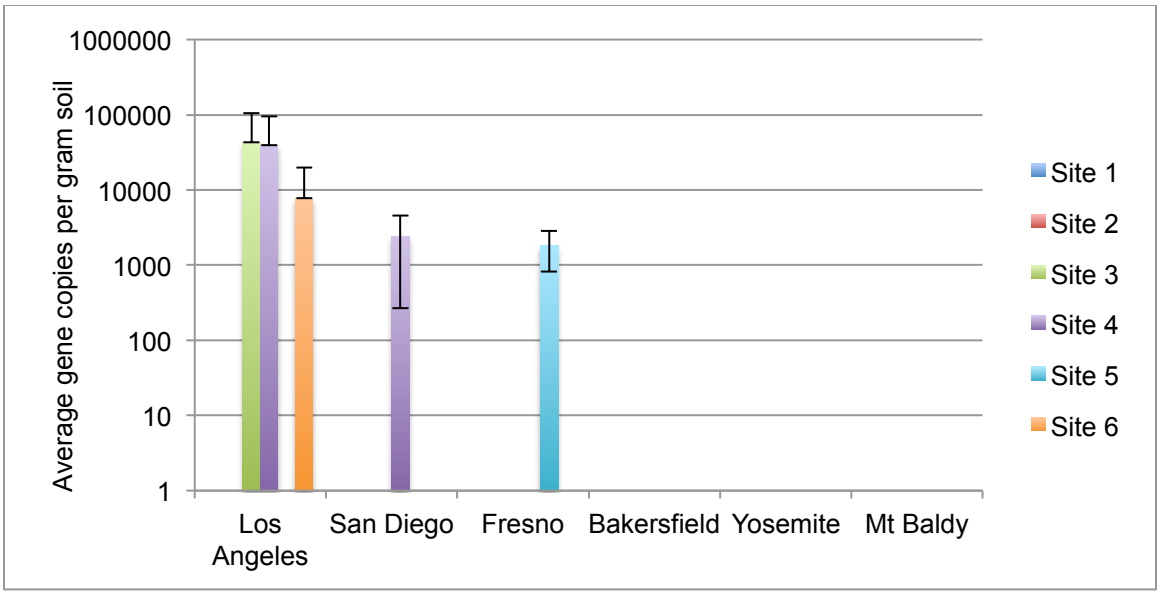


Figure 2 Detected levels of *erm*(F) gene copies per gram soil present in soil samples from 26 sites (Error bars indicate the standard deviation of three replicates. Error bars showing only the standard deviation in the plus direction are due to the log scale. Missing bars indicate non-detects.)

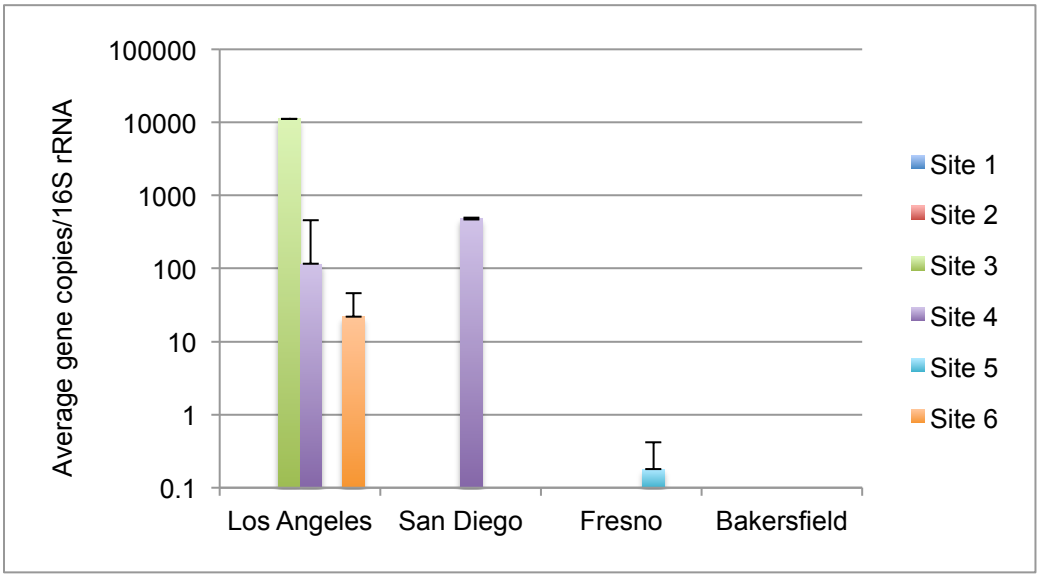


Figure 3 Detected levels of *erm*(F) gene copies normalized to relative amount of 16S rRNA gene present in soil samples from 24 sites in four cities (Error bars indicate the standard deviation of three replicates. Error bars showing only the

standard deviation in the plus direction are due to the log scale. Missing bars indicate non-detects.)

There were two analysis methods on *erm*(F) gene levels—on a per gram of soil basis, and relative to 16S gene abundance. From Figure 2 and Figure 3, site 3, 4 and 6 in Los Angeles, site 4 in San Diego, and site 5 in Fresno were quantifiable on both analysis methods while the other sites were not, including the two pristine sites. The highest *erm*(F) gene level was in site 3, Los Angeles, and the lowest was in site 5, Fresno, whichever analysis approach was utilized.

Table 5 showed the population, land area and population density in each city (U.S. Census Bureau, 2014 estimate). Population in each city was divided by land area in square miles to obtain population density. In Figure 4 and 5, average *erm*(F) gene copies per gram soil and normalized to relative16s rRNA in four cities were presented respectively.

Table 5 Population and population density in four cities

City	Population, 2014 estimate	Land area in square miles, 2010	Population density (Persons per square mile)
Los Angeles	3,928,864	468.67	8383
San Diego	1,381,069	325.19	4246
Fresno	515,986	111.96	4609
Bakersfield	368,759	142.16	2601

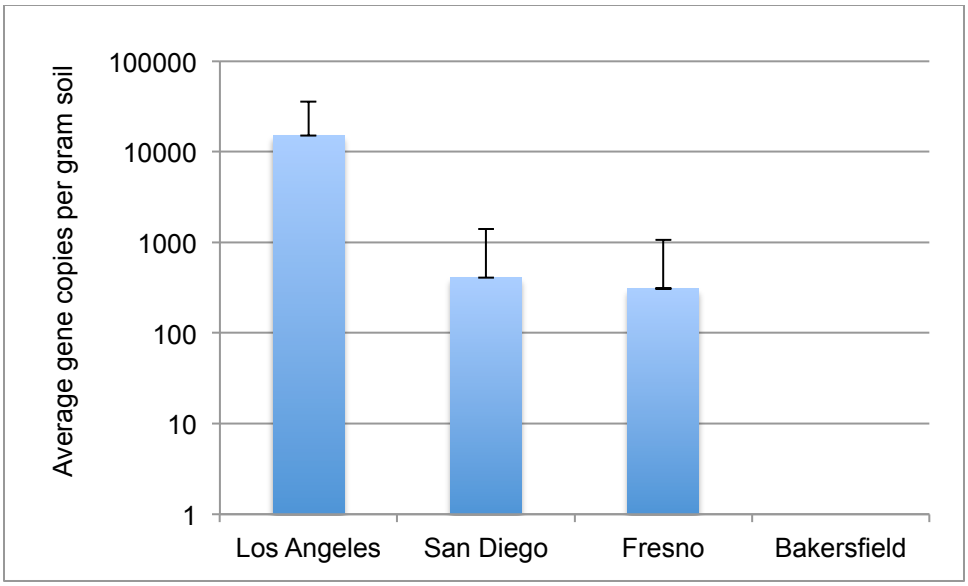


Figure 4 Average *erm*(F) gene copies per gram soil in four cities (Error bars indicate the standard deviation of six sites in each city. Error bars showing only the standard deviation in the plus direction are due to the log scale. Missing bars indicate non-detects.)

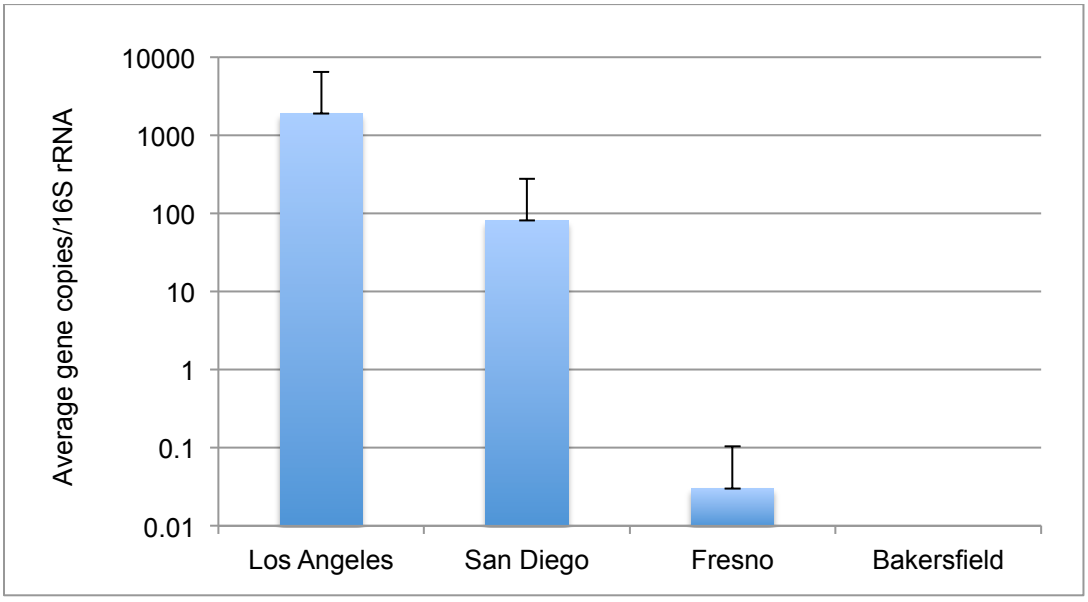


Figure 5 Average *erm*(F) gene copies divided by relative16s rRNA in four cities

(Error bars indicate the standard deviation of six sites in each city. Error bars showing only the standard deviation in the plus direction are due to the log scale. Missing bars indicate non-detects.)

Discussions

Discussions on local scale

From Figure 2, site 4 and 6 in Los Angeles on *erm*(F) gene copies per gram soil were higher than site 4 in San Diego, while they were lower in *erm*(F) gene copies normalized to 16S rRNA from Figure 3. It indicated that in site 4, San Diego, there was a higher proportion of bacteria present with the *erm*(F) gene than the two sites in Los Angeles.

In Table 1, 26 sites were classified to three groups with different levels of anthropogenic influence. And in Figure 6 and Figure 7, '1' represented low-level anthropogenic influence, '2' denoted mid-level and '3' indicated high-level. All sites detected with *erm*(F) gene fell into the high-level anthropogenic influence group. It indicated a positive correlationship between anthropogenic influence and ARG *erm*(F) levels. Although this correlation was not as strong as other study (Wang et al., 2014), the detections of erm(F) in five high-level anthropogenic influenced sites could not be treated as accidental events. Further researches are needed on more kinds of ARGs and more sites with different levels of anthropogenic influences.

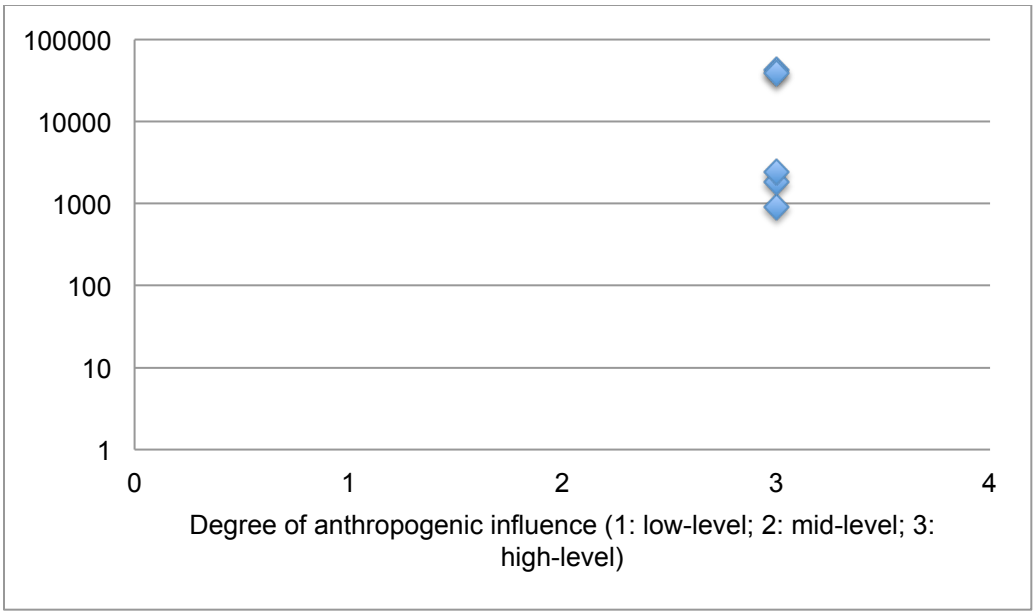


Figure 6 Relationships between *erm*(F) gene copies per gram soil and three different levels of anthropogenic influence

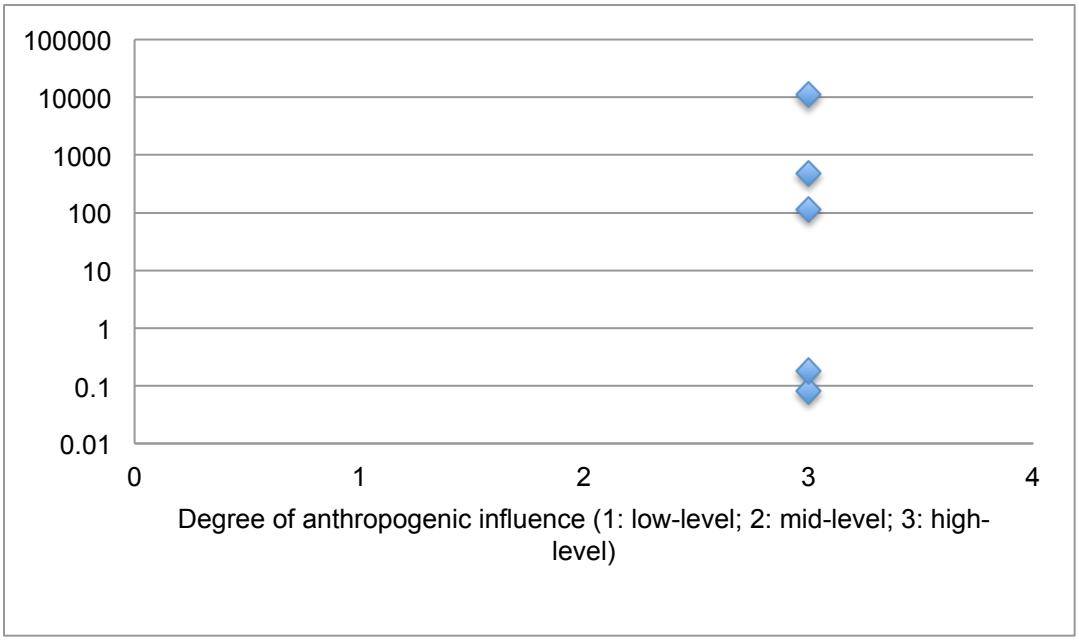
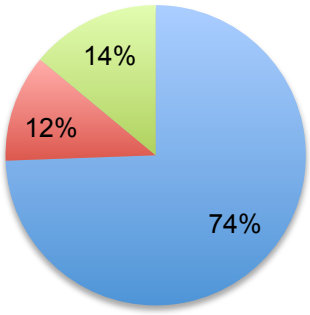
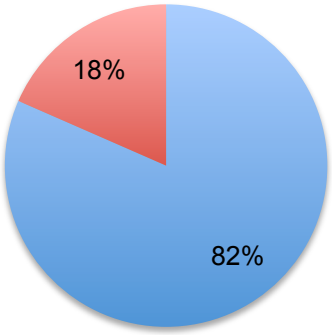


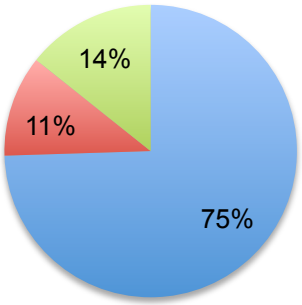
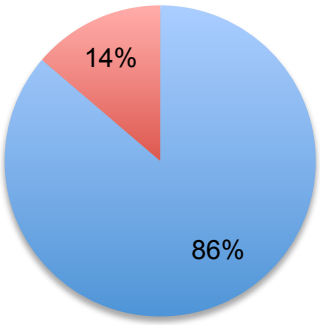
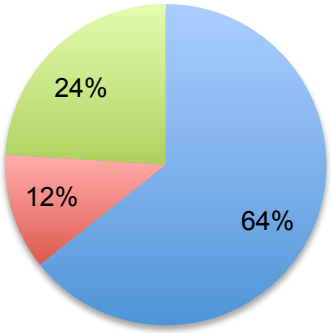
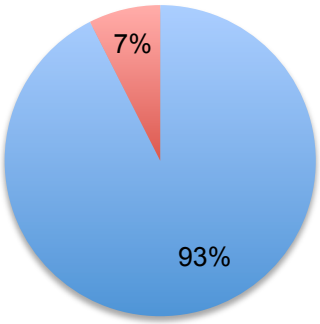
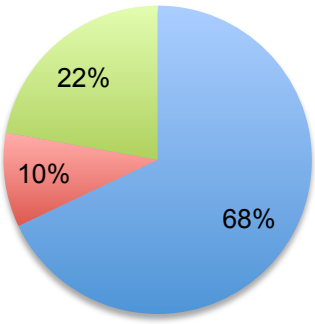
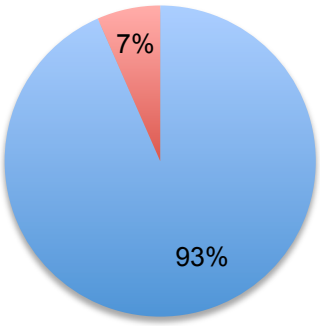
Figure 7 Relationships between *erm*(F) gene copies divided by relative 16s rRNA and three different levels of anthropogenic influence

Discussions on regional scale

Volatile fraction in total solids is an indicator of organic matter percentage in soil. In Table 6, urban cities and agricultural cities had the similar soil constitution respectively on grain sizes and volatile solids.

Table 6 Visualization on the results of grain size analysis and fixed and volatile solids in four cities

Site	Grain Size Analysis (Blue: sand; red: clay; yellow: silt)	Fixed and Volatile Solids (Blue: fixed solids; red: volatile solids)
Los Angeles		

San Diego		
Fresno		
Bakersfield		

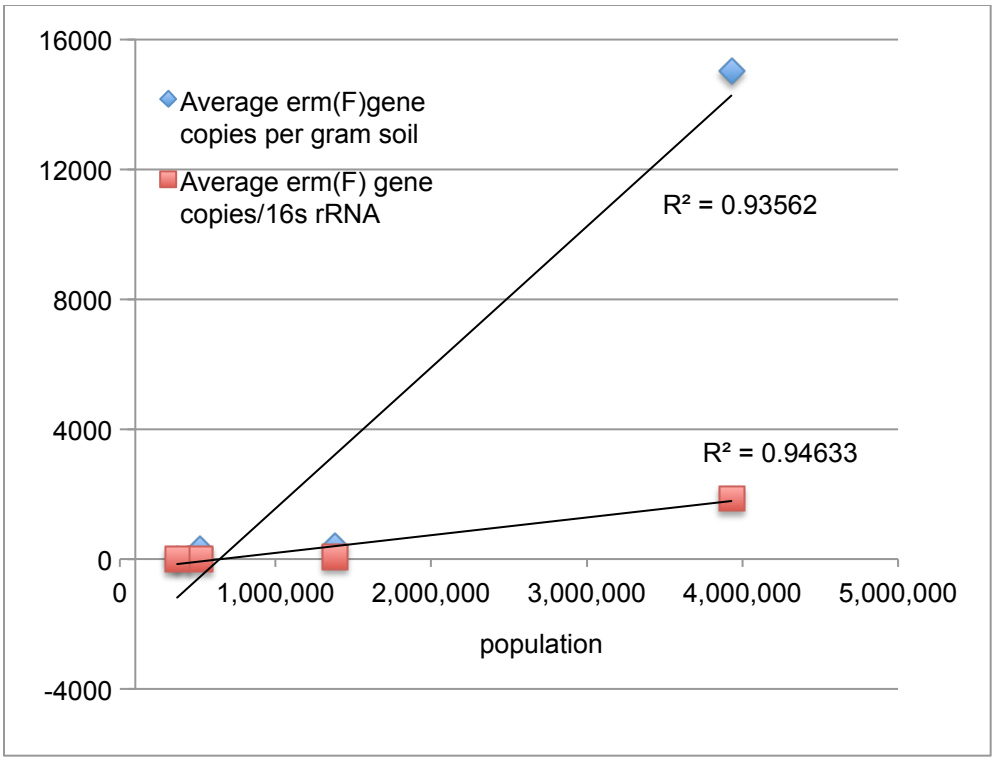


Figure 8 The relationship between population and average *erm*(F) gene copies per gram soil, average gene copies normalized to relative amount of 16S rRNA gene in four cities

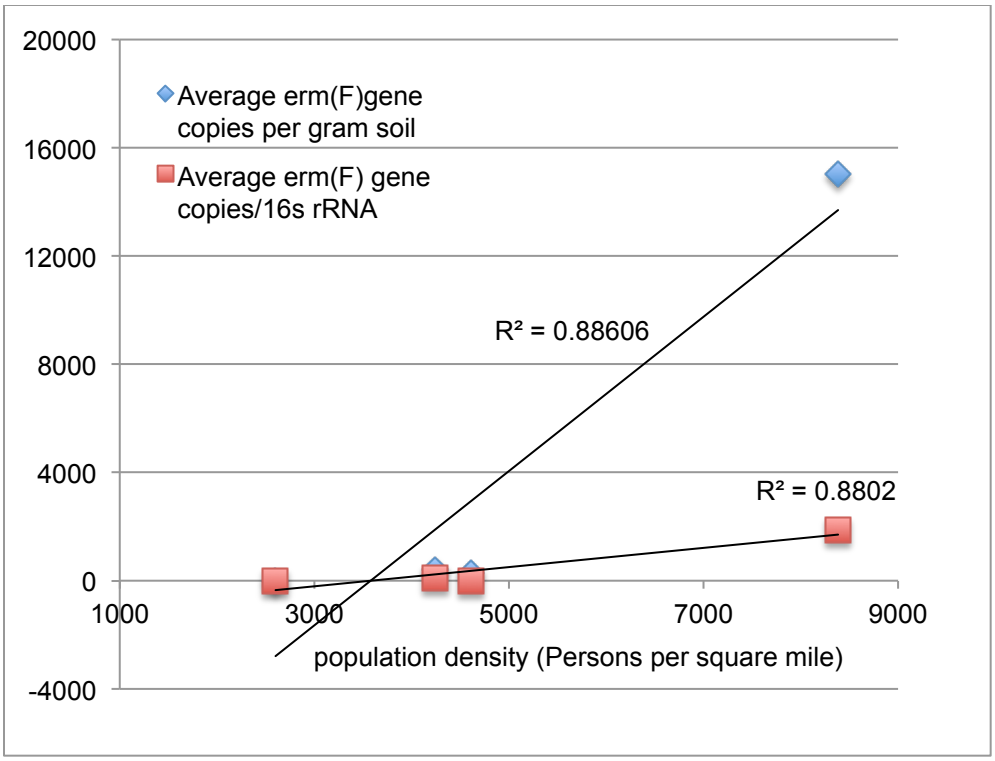


Figure 9 The relationship between population density and average *erm*(F) gene copies per gram soil, average gene copies normalized to relative amount of 16S rRNA gene in four cities

On regional scale, population and population density were indicators of anthropogenic influence. The higher the population or population density was, the higher the anthropogenic influence would be. From Figure 8, there was a positive correlationship (R^2 = 0.9356, 0.9463 respectively) between population in each city and *erm*(F) gene copies per gram soil or normalized to relative amount of 16S rRNA gene copies. From Figure 9, population density positively correlated with *erm*(F) gene in two analysis as well (R^2 = 0.8861, 0.8802).

From Table 6, Figure 4 and Figure 5, urban cities had a higher population and population density than agricultural cities and differed in soil constitution. The difference in soil constitution might result from different population or population density as well as different geochemical backgrounds.

Population and population density could contribute to human activities, such as construction of roads and buildings. Those human activities could be anthropogenic influences on the decrease of fertility and species variability in the soils and thus influenced the soil constitution. Because population and population density positively correlated with ARG *erm*(F) gene levels, soil constitution could reflect ARG *erm*(F) gene levels to some degree. Moreover, different geochemical soil background should be taken into account as well (Knapp, et al., 2011). In this part, more studies are also needed in the future.

The hypothesis in this paper is established to some extent, which is in parks in California with a stronger anthropogenic influence, the level of ARG, *erm*(F) is very likely to be higher. On the regional scale, for cities with a higher population and population density, they might have a higher level of ARG, *erm*(F).

Because of limited time, only *erm*(F) was assessed and only six sites in each city were included. Further research work is needed for the relationship between the levels of ARGs and the levels of anthropogenic influence on local and regional scale.

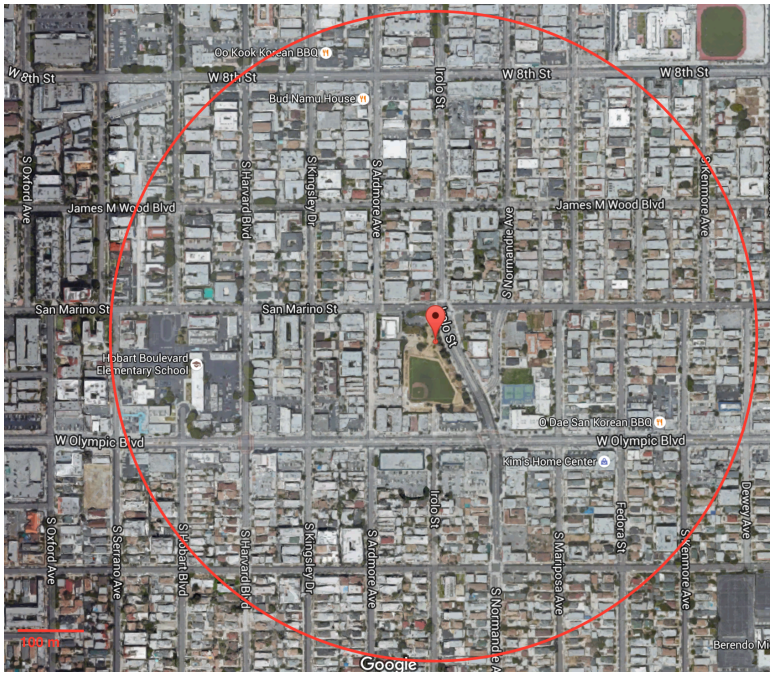
Appendix

A. Aerial photos of each site

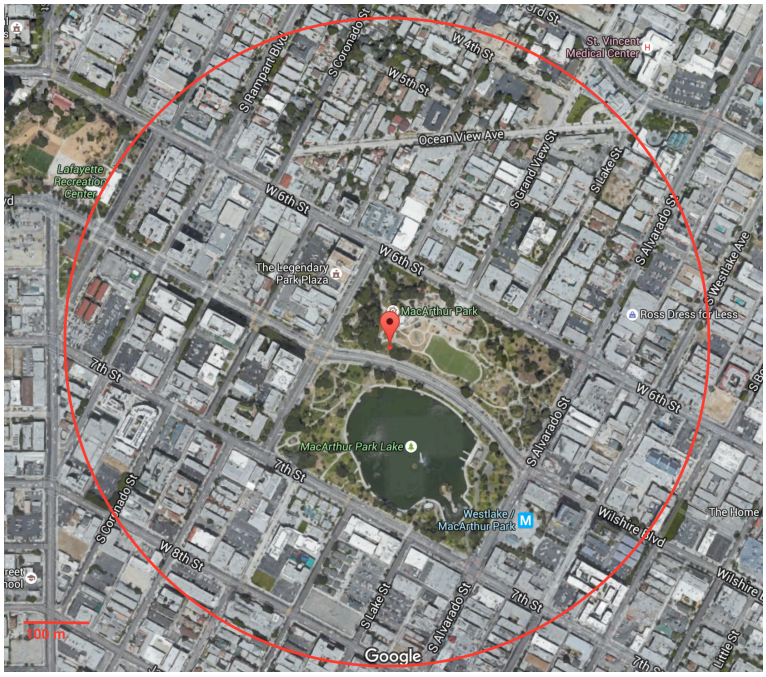
1. Los Angeles Site 1 Will Rogers State Historic Park

2. Los Angeles Site 2 Kenneth Hahn State Recreation Area

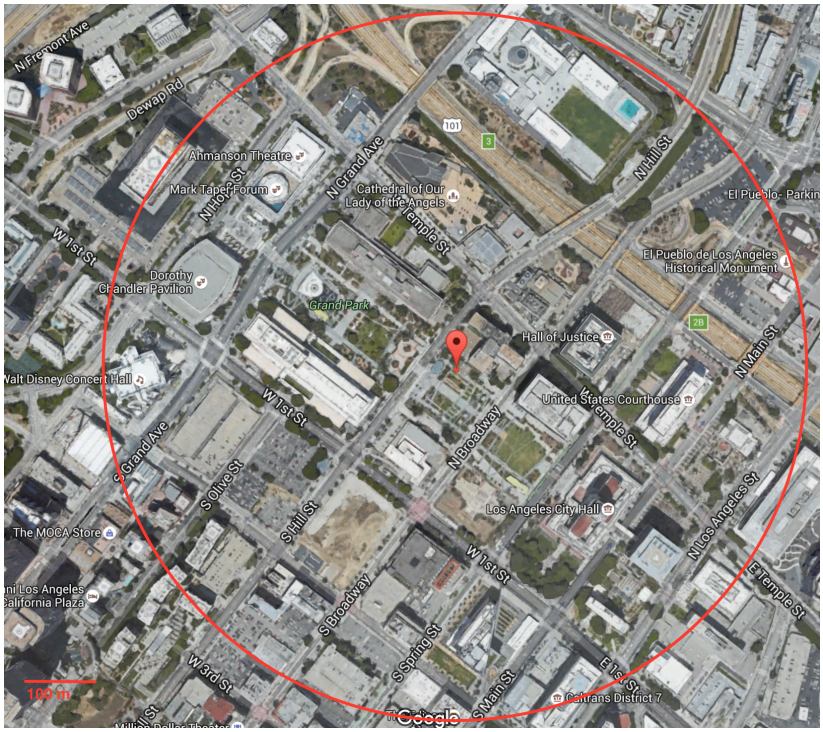
3. Los Angeles Site 3 Seoul International Park



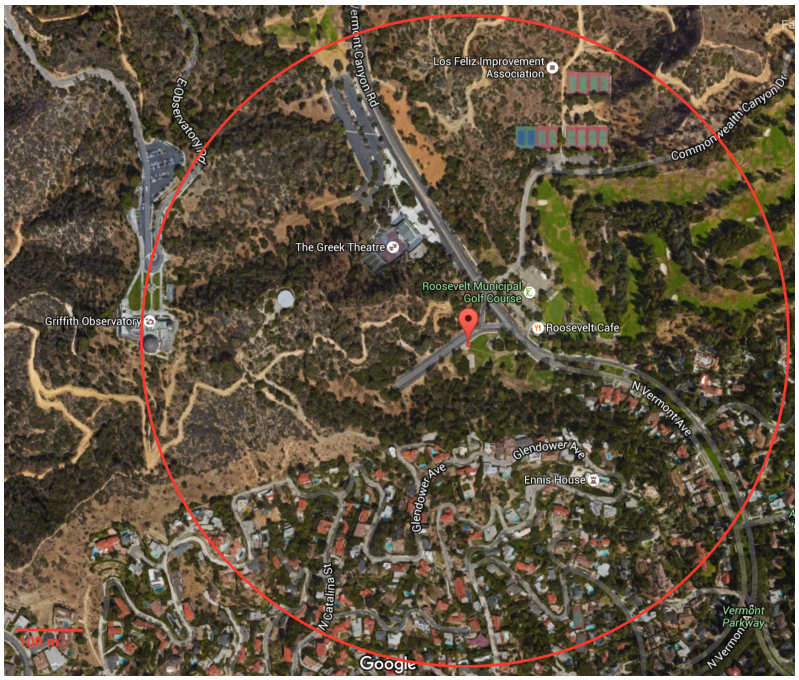
4. Los Angeles Site 4 MacArthur Park



5. Los Angeles Site 5 Grand Park



6. Los Angeles Site 6 Griffith Park

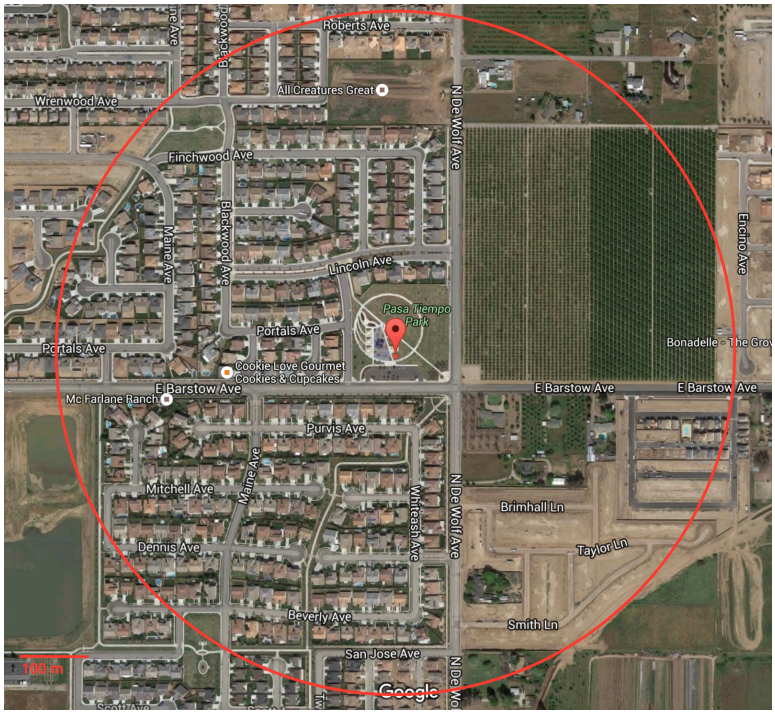


7. Fresno Site 1 Kearney Park

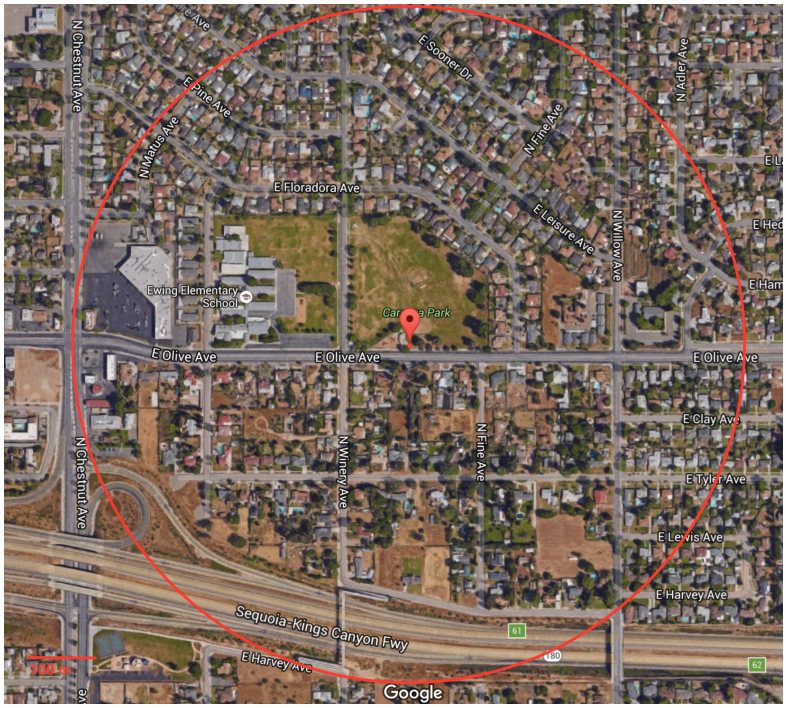
8. Fresno Site 2 Orchid Park

9. Fresno Site 3 Fresno Regional Sports Complex

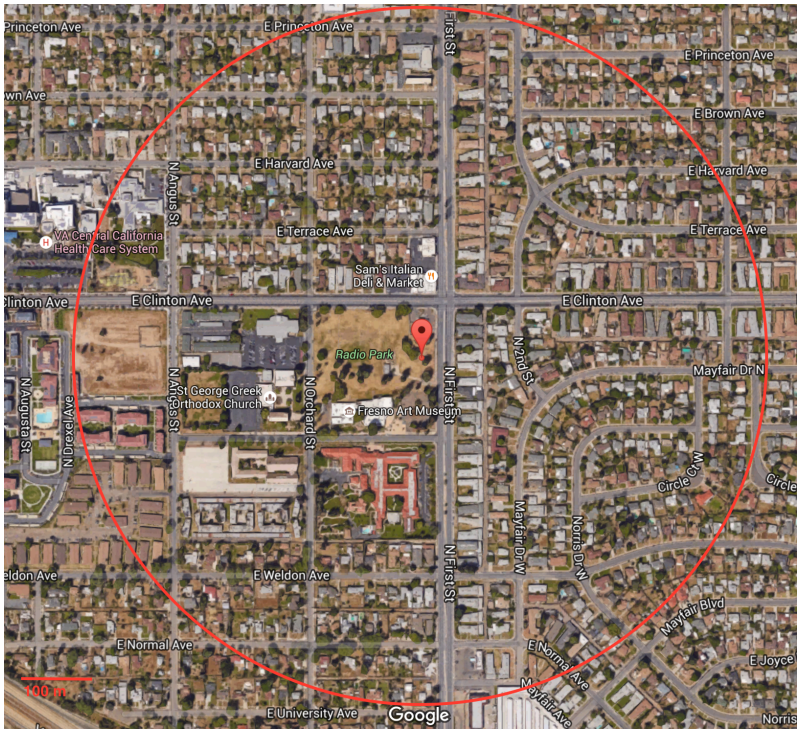
10. Fresno Site 4 Pasa Tiempo Park



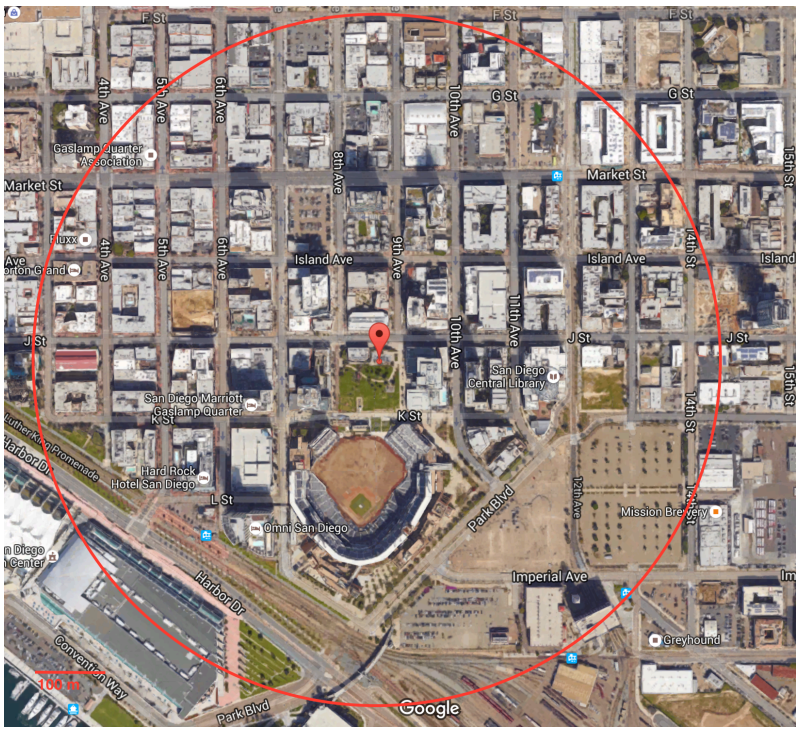
11. Fresno Site 5 Carozza Park



12. Fresno Site 6 Radio Park

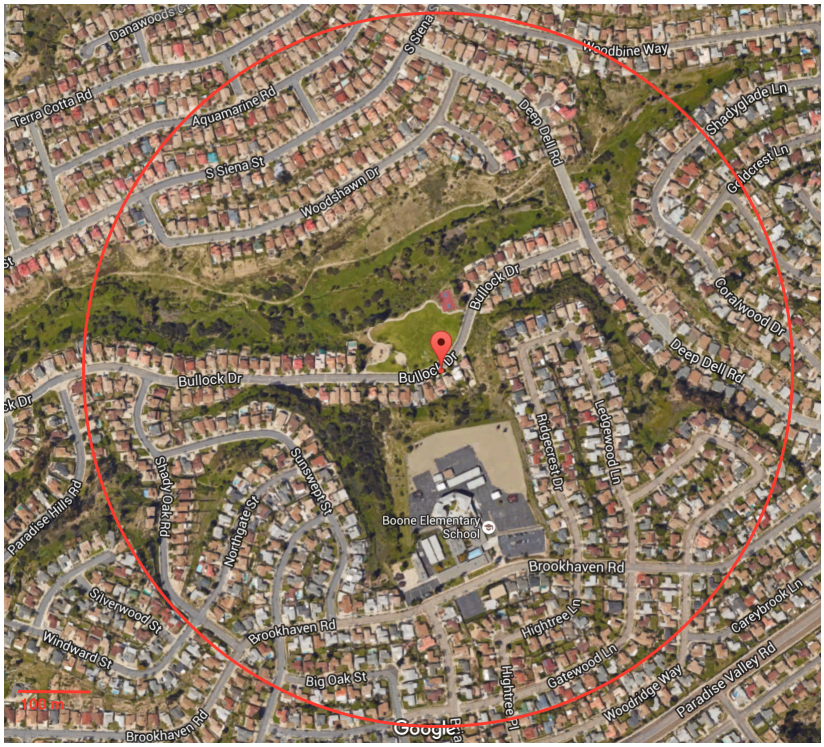


13. San Diego Site 1 Petco Park



14. San Diego Site 2 Mission Bay

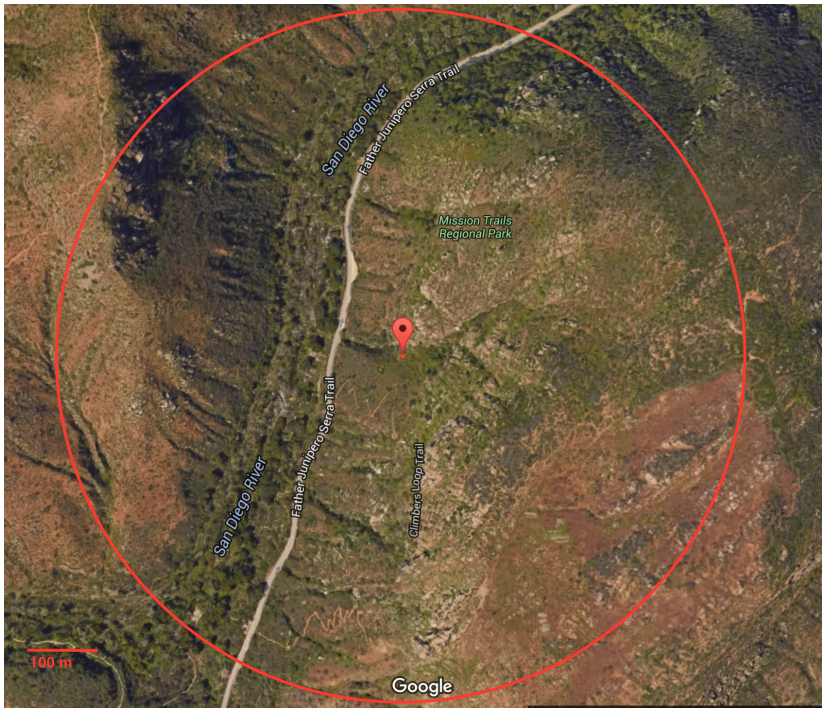
15. San Diego Site 3 Boone Park



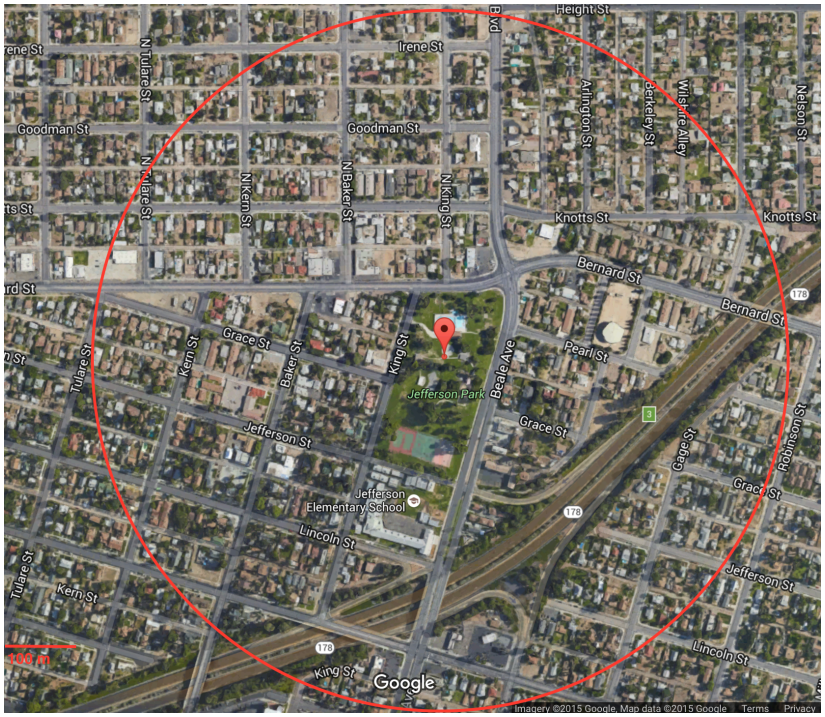
16. San Diego Site 4 Balboa

17. San Diego Site 5 Torrey Pines

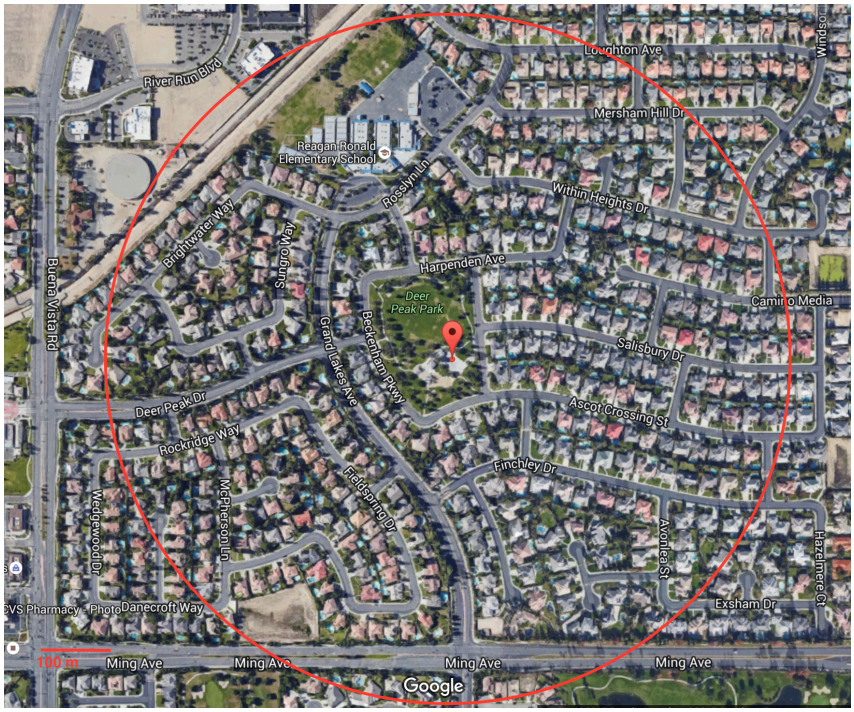
18. San Diego Site 6 Mission Trails



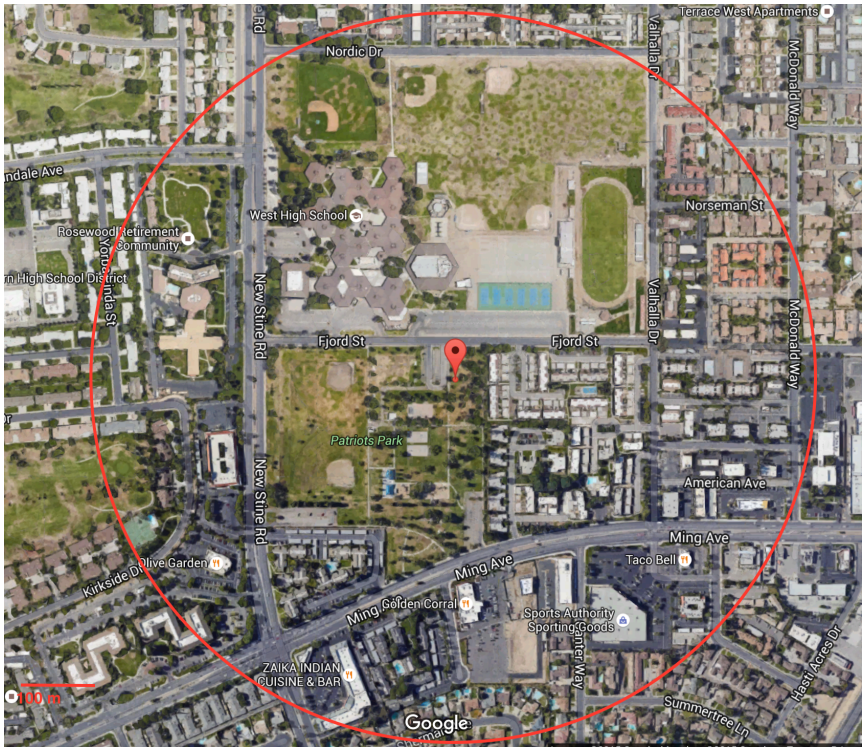
19. Bakersfield Site 1 Jefferson Park



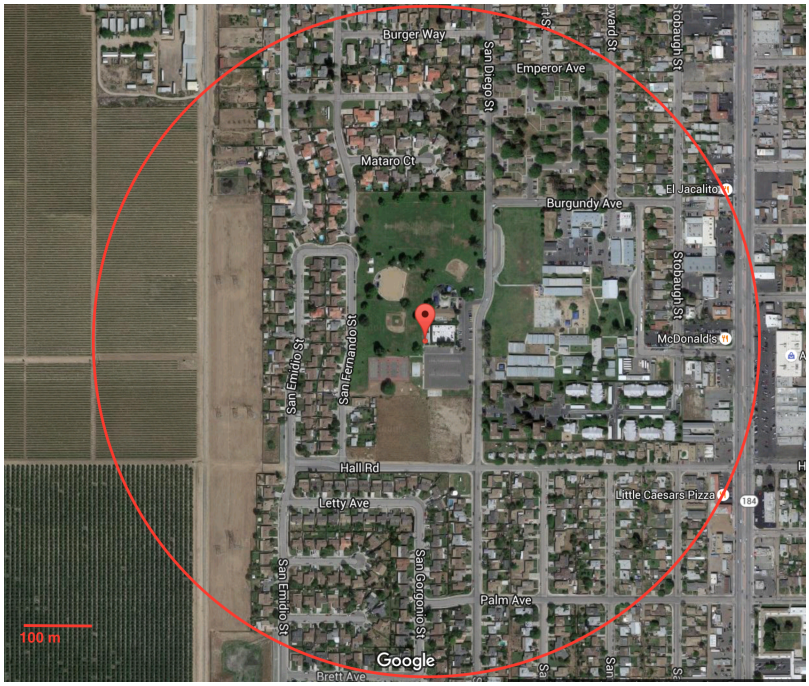
20. Bakersfield Site 2 Deer Peak Park



21. Bakersfield Site 3 Patriots Park



22. Bakersfield Site 4 Bear Mountain Park



23. Bakersfield Site 5 Tule Elk State Reserve Park

24. Bakersfield Site 6 Buena Vista Park

25. Yosemite National Park June Lake

26. Mount Baldy

B. Park name, latitude, longitude, roads, buildings and influence level

Table 7 Park name, latitude, longitude, roads, buildings and influence level

Name	site	Latitude	Longitude	Build-ings	Roads	Anthropogenic Influence level
Los Angeles Site 1	Will Rogers State Historic Park	N34.05489441010547	W118.51119685929129	42	6	low
Los Angeles Site 2	Kenneth Hahn State Recreation Area	N34.01068364273199	W118.37023613054153	42	4	low
Los Angeles Site 3	Seoul International Park	N34.054009	W118.301012	606	17	high
Los Angeles Site 4	MacArthur Park	N34.059710	W118.278300	302	20	high
Los Angeles Site 5	Grand Park	N34.055471	W118.244938	79	15	high
Los Angeles Site 6	Griffith Park	N34.118113	W118.295020	158	13	mid
Fresno Site 1	Kearney Park	N36.727584	W119.922808	3	8	low
Fresno Site 2	Orchid Park	N36.839263	W119.852161	449	21	high

Fresno Site 3	Fresno Regional Sports Complex	N36.697017	W119.834781	9	7	low
Fresno Site 4	Pasa Tiempo Park	N36.816133	W119.647037	377	24	high
Fresno Site 5	Carozza Park	N36.757641	W119.730777	453	18	high
Fresno Site 6	Radio Park	N36.771529	W119.772957	643	20	high
San Diego Site 1	Petco Park	N32.709112	W117.156767	197	21	high
San Diego Site 2	Mission Bay	N32.789578	W117.210361	29	14	mid
San Diego Site 3	Boone Park	N32.697275 1	W117.037251	696	22	high
San Diego Site 4	Balboa	N32.738572	W117.128649	1079	17	high
San Diego Site 5	Torrey Pines	N32.9210	W117.2532	0	2	low
San Diego Site 6	Mission Trails	N32.8278	W117.0511	0	1	low
Bakersfield Site 1	Jefferson Park	N35.389345	W118.986005	674	28	high
Bakersfield Site 2	Deer Peak Park	N35.343669	W119.121428	410	35	high

Bakersfield Site 3	Patriots Park	N35.341474	W119.058111	295	16	high
Bakersfield Site 4	Bear Mountain Park	N35.261390	W118.919701	441	24	high
Bakersfield Site 5	Tule Elk State Reserve Park	N35.332201	W119.363835	16	3	low
Bakersfield Site 6	Buena Vista Park	N35.23597	W119.328283	10	2	low
Yosemite	June Lake	N37.783975	W119.07494	64	13	mid
Mt Badley		N34.17945	W117.6757	0	3	low

References

Baquero, F., Martínez, J. L., & Cantón, R. (2008). Antibiotics and antibiotic resistance in water environments. *Current opinion in biotechnology*, *19*(3), 260-265.

Bergstrom, C. T., & Feldgarden, M. (2008). The ecology and evolution of antibiotic-resistant bacteria. Evolution in health and disease, 2.

Chen, J., Yu, Z., Michel, F. C., Wittum, T., & Morrison, M. (2007). Development and application of real-time PCR assays for quantification of *erm* genes conferring resistance to macrolides-lincosamides-streptogramin B in livestock manure and manure management systems. *Applied and environmental microbiology*, *73*(14), 4407-4416.

Environmental Protection Agency. (Jan. 2001). METHOD 1684, Total, Fixed, and Volatile Solids in Water, Solids, and Biosolids. EPA-821-R-01-015. Retrieved from http://water.epa.gov/scitech/methods/cwa/bioindicators/upload/2008_11_25_methods_method_biological_1684-bio.pdf

Knapp, C. W., McCluskey, S. M., Singh, B. K., Campbell, C. D., Hudson, G., and Graham, D. W. (2011). Antibiotic resistance gene abundances correlate with metal and geochemical conditions in archived Scottish soils. *PLoS One*, *6*(11), e27300.

Leclercq, R. (2002). Mechanisms of resistance to macrolides and lincosamides: nature of the resistance elements and their clinical implications. *Clinical Infectious Diseases*, *34*(4), 482-492.

Martínez, J. L. (2008). Antibiotics and antibiotic resistance genes in natural environments. *Science*, *321*(5887), 365-367.

Mellon, M., C. M. Benbrook, and K. L. Benbrook. 2001. Hogging it: estimates of antimicrobial abuse in livestock. UCS Publications, Cambridge, MA.

Nardelli, M., Scalzo, P. M., Ramírez, M. S., Quiroga, M. P., Cassini, M. H., & Centrón, D. (2012). Class 1 integrons in environments with different degrees of urbanization. *PloS one*, *7*(6), e39223-e39223.

Popowska, M., Rzeczycka, M., Miernik, A., Krawczyk-Balska, A., Walsh, F., & Duffy, B. (2012). Influence of soil use on prevalence of tetracycline, streptomycin, and erythromycin resistance and associated resistance genes. *Antimicrobial agents and chemotherapy*, *56*(3), 1434-1443.

Pruden, A., Pei, R., Storteboom, H., & Carlson, K. H. (2006). Antibiotic resistance genes as emerging contaminants: studies in northern Colorado. *Environmental Science & Technology*, *40*(23), 7445-7450.

Roberts, M. C. 2003. Acquired tetracycline and/or macrolide-lincosamides-streptogramin resistance in anaerobes. Anaerobe 9:63-69.

Roberts, M. C. 2004. Distribution of macrolide, lincosamide, streptogramin, ketolide and oxazolidinone (MLSKO) resistance genes in Gram-negative bacteria. Curr. Drug Targets Infect. Disord. 4:207-215.

Salyers, A., & Shoemaker, N. B. (2006). Reservoirs of antibiotic resistance genes. *Animal biotechnology*, *17*(2), 137-146.

Singer, R. S., & Hofacre, C. L. (2006). Potential impacts of antibiotic use in poultry production. *Avian diseases*, *50*(2), 161-172.

Stokes, H. W., Nesbø, C. L., Holley, M., Bahl, M. I., Gillings, M. R., & Boucher, Y. (2006). Class 1 integrons potentially predating the association with Tn402-like transposition genes are present in a sediment microbial community. *Journal of bacteriology*, *188*(16), 5722-5730.

Stokstad, E. L. R., and T. H. Jukes. "Growth-promoting effect of aureomycin on turkey poults." *Poultry Science* 29.4 (1950): 611-612.

U.S. Census Bureau (2014): State and County QuickFacts. Retrieved from http://www.census.gov/search-results.html?page=1&stateGeo=none&searchtype=web&cssp=SERP&q=San+Diego+city%2C+CA&search.x=0&search.y=0&search=submit

Wang, F. H., Qiao, M., Lv, Z. E., Guo, G. X., Jia, Y., Su, Y. H., & Zhu, Y. G. (2014). Impact of reclaimed water irrigation on antibiotic resistance in public parks, Beijing, China. *Environmental Pollution*, *184*, 247-253.

Woo, P. C. Y., Lau, S. K. P., Teng, J. L. L., Tse, H., & Yuen, K.-Y. (2008). Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clinical Microbiology and Infection*, *14* (10), 908–934. http://doi.org/10.1111/j.1469-0691.2008.02070.x

Woolcock, J. B. (1991). *Microbiology of animals and animal products. World Animal Science, A6* (No. 6).

You, Y., Hilpert, M., & Ward, M. J. (2012). Detection of a common and persistent tet (L)-carrying plasmid in chicken-waste-impacted farm soil. *Applied and environmental microbiology*, *78*(9), 3203-3213.