

2010

# Uncultivated Environmental TM7 Model To Study Human Disease-Associated TM7 Bacteria

David Barton  
*San Jose State University*

Follow this and additional works at: [http://scholarworks.sjsu.edu/etd\\_theses](http://scholarworks.sjsu.edu/etd_theses)

---

## Recommended Citation

Barton, David, "Uncultivated Environmental TM7 Model To Study Human Disease-Associated TM7 Bacteria" (2010). *Master's Theses*. Paper 3845.

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact [scholarworks@sjsu.edu](mailto:scholarworks@sjsu.edu).

UNCULTIVATED ENVIRONMENTAL TM7 MODEL TO STUDY HUMAN  
DISEASE-ASSOCIATED TM7 BACTERIA

A Thesis  
Presented to  
The Faculty of the Department of Biological Sciences  
San Jose State University

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Science

by  
David Barton  
December 2010

© 2010  
David Barton  
ALL RIGHTS RESERVED



## ABSTRACT

### UNCULTIVATED ENVIRONMENTAL TM7 MODEL TO STUDY HUMAN DISEASE-ASSOCIATED TM7 BACTERIA

By David Barton

The TM7 bacterial phylum has no cultivated species and includes members that span a broad range of environmental and human habitats, some of which are associated with human periodontitis. In this project, activated wastewater TM7 bacteria were analyzed and their relatedness compared to human-associated TM7 bacteria for the long-term goal of using an environmental TM7 to better understand TM7 pathogenesis in humans. DNA was extracted from activated wastewater and PCR amplified using TM7 16S rRNA gene- specific primers. The ~1,170 base pair PCR products were then cloned and sequenced. DNA sequencing and phylogenetic analysis identified environmental TM7 clones with high 16S rRNA gene similarity (99%) and mask coverage (98%) with the human oral TM7, sub-gingival clone 3 (SBG3). Total bacterial, total TM7, and environmental SBG3 populations were quantified over a two-year period using real-time quantitative PCR. Results indicated TM7 bacteria were present year-round in the six sample sites studied, while environmental SBG3 levels varied between sample sites and sample dates. The relative abundance of TM7 ranged from 2.72% to 3.75% of the total sludge prokaryotic community. Based on qPCR, environmental SBG3 ranged from 0.001% to 0.100% of the total sludge prokaryotic community. TM7 cell morphology determined via fluorescent *in situ* hybridization was consistent with previously described morphologies. Given these results, this wastewater can serve as a reservoir to study TM7 and to better understand TM7 pathogenesis in human diseases, including periodontitis.

## ACKNOWLEDGEMENTS

I would like to thank Dr. Cleber Ouverney for introducing me to TM7 bacteria and giving me the opportunity to work on this project. Thanks are also due for his guidance and the chance to travel to present my research at national conferences.

I thank Dr. John Boothby and Dr. Sabine Rech for taking time out of their busy schedules to give me advice and review this thesis. I also would like to thank all my committee members for giving me the chance to instruct microbiology laboratories under their lectures. I get a great sense of accomplishment every time I come across former students who express their gratitude and say that their lab with me was the best class they have had.

A huge debt is owed to the National Institute of Health (grant number 1SC3GM0822 91-01) and the SJSU Department of Biological Sciences (Rocci and Marianna Pisano Research Fellowship). These fundings were necessary for the completion of this thesis.

I would like to express my gratitude to Mike Ortega and Tony Finch at the San Jose/Santa Clara Water Pollution Control Plant in San Jose, California for assisting in collecting samples for this study.

I would like to express my thanks to all my lab mates for the fun work atmosphere they provided. Special thanks are due to Deepa Surendar, Kavitha Boddugari, Jamsheed Ghadiri, and especially to Jorge Dinis for assisting with the work involved in this project. Their input was most valuable in completing this study.

Last, but certainly not least, I would like to thank my family. I thank my parents for their support and the belief that hard work will always pay off. I thank my daughter Katie whose smile and laugh turned the roughest days into happier ones. Finally, I thank my wife, Stephanie, for her unconditional devotion, encouragement, and belief in my abilities for which this thesis could not be completed without.

## Table of Contents

List of Figures .....	x
List of Tables .....	xii
Introduction.....	1
Literature Review.....	5
TM7 Bacteria .....	5
TM7 Pathogenesis.....	6
Molecular Methods for TM7 Detection, Quantification, and Visualization.....	8
DNA extraction.....	8
PCR amplification and clone generation .....	11
Phylogenetic analysis.....	11
Study of bacterial populations in wastewater .....	14
Fluorescent <i>in situ</i> hybridization.....	16
Quantitative PCR .....	18
TM7 Genome .....	20
Oral TM7 metagenome .....	20
Soil TM7 metagenome.....	22
Materials and Methods.....	24
Sample Collection and Processing.....	24
DNA Extraction: Mo Bio PowerSoil Kit .....	26
DNA Extraction: Bead-Beater Half-Lysis .....	27



TM7 16S rRNA Gene Amplification.....	28
Cloning of PCR Products.....	30
Phylogenetic Analysis.....	31
Quantitative PCR.....	33
Total bacterial quantification .....	36
Total TM7 quantification .....	36
Environmental SBG3 quantification.....	37
Cloning quantitative PCR products .....	37
Preliminary Microscopy of Bacteria in Wastewater.....	39
Sonication/Emulsifier Optimization .....	41
Fluorescent <i>in situ</i> Hybridization.....	41
Results.....	44
Sludge Processing .....	44
DNA Extraction .....	50
TM7 16S rRNA Gene Amplification Results.....	55
Cloning TM7 PCR Products .....	56
Phylogenetic Analysis.....	58
Quantitative PCR .....	60
Cloning of qPCR Products.....	69
Fluorescent <i>in situ</i> Hybridization.....	69
Discussion.....	73
Identification of Environmental, Human-Like TM7 Bacteria .....	73

Evaluation of Wastewater Sample Sites for Further Study .....	75
Possible Further Studies.....	77
Conclusions.....	78
References.....	81

## List of Figures

Figure 1. Secondary aerator treatment tanks at the San Jose/Santa Clara Water Pollution Control Plant (Google Maps, 2010).....	26
Figure 2. Method for converting raw qPCR rRNA gene copy number/ $\mu$ l DNA solution to cells/g activated wastewater.....	35
Figure 3. Sonication results on wastewater samples with the addition of PBS only.....	48
Figure 4. Sonication results on wastewater samples with the addition of PBS and Tween 80 emulsifier. ....	49
Figure 5. Raw wastewater from secondary aerated treatment tank.....	50
Figure 6. Comparison between modified bead-beater half-lysis extracted nucleic acid (BB) and Mo-Bio Powersoil extracted nucleic acid (PS)..	55
Figure 7. PCR results from BAC 8F and TM7 1177R screening of nucleic acid extracted with modified bead-beater protocol....	57
Figure 8. Evolutionary distance of human-associated and environmental TM7 phylotypes based on 16S rRNA gene phylogenetic analysis.....	59
Figure 9. Temporal study of Q1 East total bacterial population and TM7 population from both “A” and “B” aerator tank groups..	63
Figure 10. Temporal study of Q1 Middle total bacterial population and TM7 population from both “A” and “B” aerator tank groups..	64
Figure 11. Temporal study of Q3 East total bacterial population and TM7 population from both “A” and “B” aerator tank groups. ....	65

Figure 12. Average number of environmental SBG3 in activated wastewater from the  
three “A” aerator tank group sample sites from two year period. .... 67

Figure 13. FISH results on wastewater samples using TM7 specific probe TM7 905  
CY3 (red) and YO-PRO-1 general DNA stain (green)..... 71

Figure 14. FISH results on wastewater samples using TM7 specific probe TM7 905  
CY5 (green), SBG3 Subgroup probe SBG3\_1 FISH CY3 (red), and  
DAPI general DNA stain (Blue). .... 72

## List of Tables

Table 1. Dates and sample sites where raw wastewater samples were collected and DNA was extracted using a modified bead-beater half-lysis protocol .....	25
Table 2. Primers and probes.....	29
Table 3. Raw wastewater and wastewater pellet weights for A1Q1 East, A1Q1 Middle, and A1Q3 East.....	45
Table 4. Raw wastewater and wastewater pellet weights for B1Q1 East, B1Q1 Middle, and B1Q3 East.....	46
Table 5. Average cell count per ml of raw activated wastewater..	51
Table 6. DNA extracted from “A” aerator tank group using modified bead-beater half-lysis protocol and Mo-Bio Powersoil kit.....	52
Table 7. DNA extracted from “A” aerator tank group samples from 3/19/2008 to 12/22/2009 using a modified bead-beater half-lysis protocol.....	53
Table 8. DNA extracted from “B” aerator tank group samples using a modified bead- beater half-lysis protocol.....	54
Table 9. Combined “A” and “B” aerator tank group results for DNA extraction, PCR amplification, and FISH.....	57
Table 10. Similarity of human-associated and environmental TM7 phylotypes to environmental SBG3 HM208132 (this work) based on 16S rRNA gene sequence phylogenetic analysis .....	60

Table 11. Experimental values included in qPCR analysis from temporal study of total bacterial, total TM7, and TM7 subgroup, environmental SBG populations in activated wastewater.. .....62

Table 12. Average TM7 percentage of the total bacterial population at the three sample sites.. .....62

Table 13. Quantification of environmental SBG3 and percent environmental SBG3 population of the total TM7 and total bacterial populations for samples that were positive for environmental SBG3 16S rRNA gene presence.. ....68

## Introduction

Until recently, much of the bacterial diversity of the human body was unknown due to limitations of cultivation-based techniques. These techniques first attempt to isolate a single type of bacterium in pure culture before it is characterized and identified. An estimated 50% of the human oral bacterial community has been cultivated (Fredricks, 2001; Loesche & Syed, 1973). In contrast, cultivation-independent studies have greatly increased the knowledge of the total bacterial community that resides on or in the human body. An estimated 85% of the oral microbial community has been identified using culture-independent techniques (Kazor *et al.*, 2003). Culture-independent studies use molecular methods, such as the polymerase chain reaction (PCR), quantitative PCR (qPCR), and fluorescent *in situ* hybridization (FISH) to identify and characterize uncultivated bacteria based on the sequence of the 16S ribosomal RNA (rRNA) gene.

Known uncultivated or difficult-to-culture pathogenic bacteria have been studied by culture-independent studies. The spirochete *Treponema pallidum* (Heymas *et al.*, 2010), the causative agent of syphilis and *Mycoplasma pneumoniae* (Waites & Atkinson, 2009), one of many microbes known to cause pneumonia, have been studied using culture-independent techniques. It is likely that additional uncultivated bacteria cause disease.

One uncultivated, possible disease-causing bacterial strain belongs to the TM7 phylum. Members can be found in a broad range of environmental habitats, such as peat bog, fresh water, saltwater, hot springs, mouse feces, termite guts, and activated

wastewater (Hugenholtz *et al.*, 2001). TM7 bacteria are also found on human sites, such as the skin (Gao *et al.*, 2007), vaginal lining (Fredricks *et al.*, 2007), esophageal lining (Pei *et al.*, 2004), intestines (Kuehbacher *et al.*, 2008), tongue (Kazor *et al.*, 2003), and sub-gingival crevices (Paster *et al.*, 2001). TM7 bacteria have been associated with two human diseases: periodontitis (Brinig *et al.*, 2003), an oral disease affecting 15% of the adult population in the United States (Oliver *et al.*, 1998) and inflammatory bowel disease (Kuehbacher *et al.*, 2008), which affects the colon and small intestine. TM7 clone IO25 has been linked to periodontitis. Based on qPCR, TM7 IO25 numbers increased significantly during periodontitis compared to healthy human mouth controls (Brinig *et al.*, 2003). FISH imaging showed that TM7 IO25 cells were larger in latter periodontitic stages, further suggesting that they may be growing more actively under disease conditions. Other TM7 bacteria, such as those in the TM7 sub-group SBG (for subgingival), have been identified in the sub-gingival crevice. In Brinig *et al.* (2003), seven SBG sub-groups were proposed. The pathogenicity of TM7 bacteria has yet to be determined.

Because TM7 cannot be cultured, research on this bacterium requires the identification of a reliable, continuous source. Human samples impose numerous restrictions, they are heavily regulated (NIH requires Institutional Review Board approvals), are known to be highly diverse from one individual to another (Brinig *et al.*, 2003; Ouverney *et al.*, 2004), lack consistency over time compromising follow up sampling to replicate results (tooth surfaces are brushed, skin is washed daily, and dietary or medicinal changes), or present challenges to access (intestinal or vaginal samples).



Once the TM7 bacteria associated with disease are detected in a human sample, IRB regulations require the patient be notified and harm to the patient be minimized. Individuals would be treated, which could change the oral microbial community and possibly remove TM7 bacteria (Oliver *et al.*, 1991). TM7 bacteria have been grouped primarily into two groups: a strict environmental group and a strict human-associated group (Hugenholtz *et al.*, 2001). However, this study suggests that there are no distinct lines between environmental and human-associated TM7 bacteria. An accessible environmental reservoir with TM7 bacteria, which are closely related to human oral disease-associated TM7, will provide an abundant source of TM7 to further investigate TM7 human pathogenesis.

The purpose of this study was to identify TM7 bacteria from environmental sources that are closely related to human-associated TM7 based on 16S rRNA gene sequence similarity. The long-term goal is to use the environmental TM7 as a model to further understand TM7 bacteria and their role in human diseases. Molecular methods used included PCR with TM7-specific primers and cloning combined with 16S rRNA sequence analysis using ARB software to identify human-like TM7 in activated wastewater. An environmental TM7 that was 98.60% similar to the human-associated TM7 sub-group SBG3 was identified by ARB phylogenetic analysis. This environmental TM7 bacterium closely related to human-associated SBG3 was labeled “environmental SBG3.” Quantitative PCR was used to quantify TM7 and environmental SBG3 abundance over a two-year period to confirm activated wastewater served as a reservoir for human-like TM7 for additional studies. Fluorescent *in situ* hybridization was used to

characterize TM7 and environmental SBG3 cell morphology. Quantitative PCR and FISH provided evidence that TM7 was present year-round in the activated wastewater samples, but environmental SBG3 was detectable only at select locations and times during the years.

## Literature Review

### TM7 Bacteria

TM7 bacteria are uncultivated bacteria. TM7 cannot be grown using traditional microbial techniques and there are no representatives that have been isolated in pure culture (Zengler *et al.*, 2002). The exact reasons for TM7 resistance to cultivation are unknown. It is possible that presence of high nutrient levels in traditional microbial growth media or the lack of a specific substrate found in its original environment inhibit isolation of most uncultivated bacteria in the laboratory (Zengler *et al.*, 2002). Since these organisms cannot be cultured using traditional microbiological methods, molecular methods that target the 16S ribosomal RNA (rRNA) gene are used to characterize TM7. These molecular methods include polymerase chain reaction (PCR) to identify TM7 presence in a sample, cloning the amplified PCR product, sequencing the cloned product, comparing the cloned sequences to a 16S rRNA gene database, quantitative PCR (qPCR) to determine TM7 abundance of the total mixed bacterial population, and fluorescent *in situ* hybridization (FISH) using TM7 16S rRNA molecule specific probes to microscopically visualize TM7 cells in a mixed bacterial community.

TM7 bacteria were first identified in a German peat bog using molecular methods listed above. TM7 have been found in other environmental habitats, including soil from old growth forests, fresh and marine water, and activated wastewater from wastewater treatment plants (Hugenholtz *et al.*, 2001). Environmental TM7 cell morphologies have been characterized by FISH and epifluorescence microscopy. Those morphologies range

from cocci, to bacilli, to filaments (Hugenholtz *et al.*, 2001). Cocci were 0.5  $\mu\text{m}$  in diameter, while bacilli and individual cells in filaments ranged from 1  $\mu\text{m}$  to 5  $\mu\text{m}$  in length (Hugenholtz *et al.*, 2001). TM7 bacteria are thought to be related to the Gram-positive bacteria (Hugenholtz *et al.*, 2001). Larger filaments had a cell wall that was similar to Gram-positive type cell walls when visualized using transmission electron microscopy (Hugenholtz *et al.*, 2001).

### **TM7 Pathogenesis**

Using 16S rRNA gene PCR amplification, TM7 have also been identified in human body samples from skin (Gao, *et al.*, 2007), vaginal secretions (Fredricks *et al.*, 2007), esophagus (Pei *et al.*, 2004), intestines (Kuehbachner *et al.*, 2008), and mouth (Kazor *et al.*, 2003; Paster *et al.*, 2001). TM7 bacteria have been associated with chronic periodontitis (Kumar *et al.*, 2003). Periodontitis is an oral disease caused by bacteria (Oliver *et al.*, 1991). In the United States alone, 50% of individuals above the age of 30 have some form of periodontitis (Albandar & Kingman, 1999; Oliver *et al.*, 1998), with 15% of Americans suffering severe periodontitis (Oliver *et al.*, 1991). Advance stages of periodontitis cause tooth loss by destruction of ligaments in the mouth, jaw bone, and dental nerves (Oliver *et al.*, 1991). Oral bacteria have also been connected to bacterial endocarditis (Agarwal *et al.*, 1997), aspiration pneumonia (Smith & Thornton, 1993), and coronary heart disease (Renvert *et al.*, 2006). Previous research has shown that 40% of bacteria that form plaque are novel species with no known culture representatives (Paster

*et al.*, 2001). It is possible that TM7 or other uncultivated bacteria in the mouth may be linked to other diseases.

Amplification of 16S rRNA by PCR identified TM7 bacteria in 12 out of 12 individuals with signs of periodontitis (Kumar *et al.*, 2003). By comparison TM7 bacteria were found in four out of nine individuals showing no sign of periodontitis (Kumar *et al.*, 2003). Using FISH, TM7 numbers found in individuals with periodontitis was three times the number of TM7 in individuals with healthy mouths (Ouverney *et al.*, 2003). A qPCR study by Brining *et al.* (2003) showed that TM7 bacteria were 0.3-0.5% of the total bacterial population in healthy mouths. In the same study, TM7 bacteria were found in all individuals with periodontitis (Brinig *et al.*, 2003). Total bacterial counts from diseased individuals were about twice the bacterial counts from healthy individuals (Brinig *et al.*, 2003). TM7 bacterial counts from disease individuals were 10 times higher compared to healthy individuals (Brinig *et al.*, 2003). The highest ratio of TM7 bacteria versus the total bacterial population (0.5%) was observed in the early stages of periodontitis ( $P < 0.01$ ; [Student's *t* test]) (Brinig *et al.*, 2003).

TM7 IO25 was found to be significantly higher in individuals with periodontitis compared to healthy individuals ( $P < 0.001$ ; [Fisher's exact test]) (Brinig *et al.*, 2003). Another study by Ouverney *et al.* (2003) used FISH to show that IO25 cells were 2.5 times more abundant in disease sites compared to healthy sites; although, this was not significant ( $P = 0.778$ ; [unpaired *t* test]). In the same study by Ouverney *et al.* (2003), TM7 IO25 cells from diseased sites were longer than TM7 IO25 cells observed from healthy sites ( $P = 0.003$ ; unpaired *t* test). Visualization of diseased samples using FISH

suggests that TM7 filaments may be necessary for development of biofilms that leads to chronic periodontitis and that TM7 IO25 may be pathogenic and an active contributor to periodontitis formation (Ouverney *et al.*, 2003).

### **Molecular Methods for TM7 Detection, Quantification, and Visualization**

**DNA extraction.** Bacterial genomic DNA extraction from wastewater samples is complicated by several factors, including: **i)** the need to extract large size DNA molecules for PCR amplification, **ii)** the difficulty to extract DNA from the many bacterial representatives in the mixed community while maintaining their relative abundances in the final extract, and **iii)** the presence of humic acid contaminants which, due to its chemical properties, tend to be extracted with nucleic acid and later inhibit PCR (Yeates *et al.*, 1997). Considering that TM7 bacteria constitute 6% or less of the total bacterial population by qPCR (Brinig *et al.*, 2003) and FISH (Hugenholtz *et al.*, 2001), any loss of DNA material during the extraction steps could considerably affect the results. Hence, an optimal DNA extraction method is essential.

Many protocols for DNA extraction have been published. DNA extraction on soil samples using an ultrasonic processor, for instance, resulted in DNA fragments smaller than 500bp (Yeates *et al.*, 1998). These DNA fragments are too small to be useful for PCR due to chimera formation (Liesack *et al.*, 1991). Isolating bacterial cells from contaminating particulate matter, prior to DNA extraction, using a density gradient sets a bias for which organisms are extracted (Roose-Amsaleg, Garnier-Sillam, & Harry, 2001).

The Yeates study (1998) performed PCR amplification of DNA extracted from density gradient isolated microorganisms. Primers specific for the prokaryotic 16S rRNA gene were able to amplify bacterial DNA while primers specific for the eukaryotic 18S rRNA gene did not produce results (Yeates *et al.*, 1998). This suggests that smaller bacteria were isolated while larger eukaryotic cells were not. This method would work for smaller cocci and bacilli TM7 cells, but might exclude larger bacilli and filamentous TM7 members.

Humic acids have a similar charge and size to DNA making their separation difficult when purifying DNA (Holben, 1994). Humic acids have been able to block DNA polymerase enzyme activity used in PCR despite DNA quantity (Holben, 1994). Humic acids also interfere when quantifying DNA through a spectrophotometer since humic acids, like DNA, absorb at 260nm wavelength (Yeates *et al.*, 1997). DNA purity can be determined using a spectrophotometer in the presence of humic acids by looking at the 260/230 ratio. A 260/230 ratio above 2 indicates pure DNA, while lower values suggest that humic acid contamination is present (Yeates *et al.*, 1997). Mechanical bead beating, physically shaking sample with small glass beads at high speeds, on soil samples gave 260/230 ratios closer to 2 ( $1.69 \pm 0.02$ ) when compared to sonication ( $1.47 \pm 0.07$ ), density separation ( $1.10 \pm 0.003$ ), chemical lysis, and Proteinase K enzymatic lysis (chemical and enzymatic lysis both  $1.31 \pm 0.03$ ) protocols (Yeates *et al.*, 1998). In addition, bead beating protocols gave a 260/280 ratio closer to the ideal 1.8 value. A 1.8 260/280 ratio indicates no protein contamination in the extracted DNA (Yeates *et al.*, 1997). DNA extracted using a bead beating protocol showed inhibition of PCR when

DNA sample was undiluted (Yeates *et al.*, 1998). PCR products were obtained when the extracted DNA was diluted in a 1:50 ratio (Yeates *et al.*, 1998).

A mechanical bead beating DNA extraction performed by Roh *et al.* (2006) was used to extract DNA from activated wastewater samples. Samples were incubated for 30 min at 65°C with Triton-X, a cellular denaturant, and Proteinase K. After incubation, guanidine thiocyanate and EDTA were added to prevent enzyme activity. Half the sample was transferred to tubes containing 0.1 mm and 0.5 mm silica bead. Tubes with beads were vortexed at 5,000 rpm for 30 sec. The genomic DNA in the tube with glass beads was broken into smaller pieces (Roh *et al.*, 2006). Breaking up DNA reduced supercoiling which can produce PCR inhibition (Roh *et al.*, 2006). The tube without beads served as a source of intact genomic DNA in case the glass beads cut the genes of interest during the bead beating step (Roh *et al.*, 2006). This would also inhibit PCR amplification later (Roh *et al.*, 2006). Vortexed and non-agitated sample were combined. Benzyl alcohol was added to separate DNA from proteins and lipids (Roh *et al.*, 2006). DNA was pelleted with the addition of 3M sodium acetate and 100% ethanol (Roh *et al.*, 2006). The resultant DNA pellet was sequentially washed with 100% and 70% ethanol solutions (Roh *et al.*, 2006). The DNA pellet was resuspended in a low salt buffered solution (Roh *et al.*, 2006). Mechanical bead beating gave the optimal combination of high DNA yield, large DNA size, and minimal humic acid contamination for PCR amplification (Roh *et al.*, 2006; Yeates *et al.*, 1997).



**PCR amplification and clone generation.** Originally, TM7 was found by screening a peat bog prokaryotic population using 16S rRNA gene primers BAC 8F and BAC 1492R (Table 2 in Materials and Methods section) (Hugenholtz *et al.*, 2001). BAC 8F and BAC 1492R are broad range primers designed to target most 16S rRNA genes of the bacterial domain (DeLong, 1992). Later, a more specific TM7 primer, TM7 1177R (Table 2 in Materials and Methods section), was used with BAC 8F to screen human oral samples specifically for TM7 16S rRNA genes (Brinig *et al.*, 2003). In the latter approach, PCR was setup using *Taq* DNA polymerase (Brinig *et al.*, 2003; Hugenholtz *et al.*, 2001) with each primer at final concentration of 10 pmol and annealing temperature of 64°C (Brinig *et al.*, 2003). PCR products were cloned using TOPO-TA cloning kit and TOP10 chemically competent *Escherichia coli* cells (Invitrogen, Carlsbad, California) (Brinig *et al.*, 2003). Inserts were sequenced using BigDye sequencing chemistry (Applied Biosystems, Foster City, California). Sequencing with vector specific primers (e.g. M13 Forward and M13 Reverse) was used to sequence the insert ends (Paster *et al.*, 2001). Sequencing with BAC 8F was used to sequence the middle of the insert (Paster *et al.*, 2001).

**Phylogenetic analysis.** CodonCode Aligner (CodonCode Corporation, Dedham, Massachusetts) is a program used to clean DNA sequences by analyzing the chromatogram for each sequence (CodonCode Corporation, 2009). The chromatogram shows the relative signal strength of each base (adenine, thymine, cytosine, and guanine) for each position in the sequence (CodonCode Corporation, 2009). CodonCode Aligner

expedites nucleic acid sequence analysis when programmed to automatically delete parts of the sequence with poor or mixed signals, contaminating vector sequence, and primer sequence based on user specific thresholds. Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1997) in the GenBank database provided by the National Center for Biotechnology Information (NCBI) compares sequences to previously published 16S rRNA gene sequences. This is done to determine submitted sample sequences novelty or similarity to published 16S rRNA gene sequences. BLAST uses parts of the submitted DNA sequence to compare to the same segments of previously published, known clean DNA sequences in the NCBI GenBank DNA database (Altschul *et al.*, 1997). This greatly increases the efficiency of comparing query sequences to larger databases.

Sequences similar to previously published TM7 sequences need to be aligned to a 16S rRNA gene standard (process described below) prior to insertion into a sequence database. Near Alignment Space Termination (NAST) is a web based aligner program supported by Greengenes (DeSantis *et al.*, 2006). Greengenes is a DNA database similar to GenBank, but it focuses on prokaryotic 16S rRNA genes while GenBank contains additional non-prokaryotic DNA. Greengenes provides faster analysis turnaround when compared to GenBank due to its smaller size and extensive coverage of the 16S rRNA gene. NAST expands submitted (query) sequences to a 7,682 character standard (DeSantis *et al.*, 2006). This expansion is done to consistently match conserved regions shared by most 16S rRNA gene sequences at the same location (DeSantis *et al.*, 2006). NAST aligns queried sequences to 10,000 core representative genes in the Greengenes

database (DeSantis *et al.*, 2006). Conserved regions in the submitted sequence are aligned to conserved regions in the core representative database (DeSantis *et al.*, 2006). Aligned sequences are checked for chimeras. A chimera is a PCR artifact in which the sequence is amplified from more than one parent sequence (Liesack *et al.*, 1991). When compared to other 16S rRNA gene sequences, a chimera will appear as a novel sequence due to its low percent similarity to any known organism (Liesack *et al.*, 1991). Chimera Check with Bellerophon is a chimera screening program (Cole *et al.*, 2003). Segments of submitted sequences that match very well to segments of previously published sequences, but not the whole sequence, are thought to be chimeras and discarded from the analysis (Cole *et al.*, 2003).

Once cleaned, aligned, and checked for chimera formation, 16S rRNA gene sequences are ready to be analyzed. Phylogenetic relationships to previously well characterized sequences can be performed by software packages such as ARB (Ludwig *et al.*, 2004). ARB is a robust software for constructing phylogenetic trees of query 16S rRNA sequences of interest against its own extensive pre-aligned 16S rRNA gene sequences (Ludwig *et al.*, 2004). ARB uses both the primary and the secondary structure when comparing submitted 16S rRNA gene sequences to the currently aligned 236,469 sequences in the ARB database (Ludwig *et al.*, 2004). ARB uses a neighbor-joining algorithm which pairs submitted sequences to their closest match in the database (Ludwig *et al.*, 2004). Closely related sequences are depicted in a phylogenetic tree (Ludwig *et al.*, 2004). Due to the neighbor-joining algorithm, ARB pairs organisms (or groups of organisms) on two separate “branches,” which diverge from a common ancestor (Ludwig

*et al.*, 2004). The point at which the branches separate is called the “node” of the tree (Ludwig *et al.*, 2004). The length of each branch represents how distantly the organism (or group of organisms) is (are) from that common ancestor (Ludwig *et al.*, 2004). The longer the branch is, the more genetically distant that branch is from the common ancestor (Ludwig *et al.*, 2004).

ARB tests phylogenetic tree topography by repeating tree generation 100 separate times or greater (Bik *et al.*, 2006). Bootstrap values are assigned to each node in the tree. Bootstrap values represent the percentage of ARB iterations that generate the same node in the tree over the 100+ replicates (Ludwig *et al.*, 2004). The higher the bootstrap value the more confidence ARB assigns that node and the more likely the two groups are distinct but related to a common ancestor (Ludwig *et al.*, 2004). Phylogenetic trees are rooted with the addition of a sequence from an organism not related to any represented organisms. This distantly related sequence (called the out-group) roots all representative sequences to a common ancestor (Ludwig *et al.*, 2004).

**Study of bacterial populations in wastewater.** Particles are present in activated wastewater that large numbers of bacteria attach to (Hugenholtz *et al.*, 2001). It is thought that bacterial adhesion to a substrate is required for cellular growth (Amalfitano & Fazi, 2008). Thus, there may be strong adhesive forces to overcome in order to separate bacteria from particles for microscopic visualization. Chemical and mechanical means are used to separate soil bacteria from particles. Two common methods are the addition of Tween 80 (Barkovskii & Fukui, 2004) and sonication (Barkovskii & Fukui,

2004). Tween 80 (C<sub>64</sub>H<sub>124</sub>O<sub>26</sub>) is an emulsifier and is soluble in water and organic compounds (Merck, 2001). An ultrasonic sonicator creates alternating high and low pressures works by transmitting high frequency (40 kHz or greater) sound waves through samples (Branson, 2002). In each cycle, bubbles form during low pressure and implode during high pressure (Branson, 2002). This combination works to break up adhesive forces that hold particles together (Branson, 2002).

The combination of Tween emulsifier, shaking, and sonication, followed by density gradient separation increased the number of soil bacteria recovered when compared to untreated soil by microscopic visualization (Amalfitano & Fazi, 2008). A study on peat used the addition of PBS pH 6.2, 0.01% Tween 80 solution with shaking, or vortexing, or sonication to extract bacteria from peat particles (Barkovskii & Fukui, 2004). DNA was extracted from bacteria recovered by each procedure, PCR amplified, cloned, and sequenced. No single procedure was able to fully extract all bacterial groups present in peat (Barkovskii & Fukui, 2004). Combining the three procedures recovered all bacterial groups (Barkovskii & Fukui, 2004). A study by Janssen *et al.* (2002) showed that 91% of soil bacteria survived sonication; however, no filamentous bacteria or fungal hyphae were seen after sonication with a high powered probe ultrasonic sonicator.

These procedures also work for larger cells that might be similar in size to larger TM7 cells and filaments. The addition of Tween 80 increased *Cryptosporidium parvum* oocyst dispersion from soil and fecal particles (Kuczynska & Shelton, 1999). Tween 80 addition before soil/fecal sample storage improved oocyst separation for stored samples (Kuczynska & Shelton, 1999).

**Fluorescent *in situ* hybridization.** Fluorescent *in situ* hybridization (FISH) can be used to microscopically visualize prokaryotic cells using fluorescently labeled probes targeting the 16S rRNA molecule in the prokaryotic small ribosomal subunit (Amann *et al.*, 1995). FISH allows for the visualization of a specific group, for example TM7, in a mixed bacterial community when using an epifluorescence microscope. Whole cell visualization is possible due to the large number of ribosomes, each with a target rRNA molecule, within one cell. Therefore, hundreds of fluorescent probes can label each bacterial cell (Hugenholtz *et al.*, 2002).

Probes, much like primers, can be tailored to target a broad (domain level) group or a smaller (specific genus or species level) group depending on how conserved the target sequence is (Amann *et al.*, 1995). Specificity of probes in FISH is also affected by the temperature used during the hybridization step in which a probe solution is introduced to fixed, permeabilized cells. The higher the hybridization temperature the closer the complementation between the probe sequence and target rRNA sequence needs to be to remain hybridized. Hybridization temperature can be kept constant when hybridizing probes of different annealing temperatures by adjusting formamide concentration. Formamide weakens hydrogen bond formation (Amann *et al.*, 1995). An increase of 1% formamide solution in the hybridization buffer is equivalent to raising hybridization temperature by about 1.5°C (Amann *et al.*, 1995). For instance, both probe BAC 338 (Murray *et al.*, 1998), which targets most prokaryotes, and TM7 group specific probe TM7 905 (Hugenholtz *et al.*, 2001) hybridize at 43°C and 20% formamide concentration (Amann *et al.*, 1995; Ouverney *et al.*, 2003). Similarly probe SBG3\_1 FISH (Table 2 in

Materials and Methods section), which targets the SBG3 subgroup of TM7, was validated to hybridize at 43°C and 30% formamide (Gavrilova *et al.*, 2009).

The first step in the FISH method is to preserve and permeabilize the cell wall and cell membrane. This allows probes to enter the cytoplasm and bind to target rRNA molecules present in the ribosomes (Amann *et al.*, 1995). The fixed sample is added to Teflon etched slides with ADD-Cell coating (Fisher Scientific, Waltham, Massachusetts). The Teflon etching creates wells in which sample and probe solutions are added and prevents probe solution from mixing between well to well (Ouverney *et al.*, 2003). The ADD-Cell coating increases bacterial cell adhesion to the glass slides. A hybridization buffer is prepared with at final concentration: 0.9 M NaCl, 20 mM Tris-HCl, 0.01% SDS, formamide (concentration determined by probe validation), and 5 ng/μl probe (Amann *et al.*, 1995; Hugenholtz *et al.*, 2002). Multiple probes, each with a distinct fluorescent dye, can be combined in the same hybridization buffer for multiple labeling. Each probe final concentration should be at 5 ng/μl (Hugenholtz *et al.*, 2002). Hybridization occurs for 2-24 h. During this period probes enter the cells and bind to the target rRNA (Amann *et al.*, 1995). Probes that are not hybridized or probes that are non-specifically hybridized are removed during a washing step (Amann *et al.*, 1995). Salt replaces formamide in the wash buffer as low salt concentration also inhibits DNA hybridization (Amann *et al.*, 1995). The wash buffer is warmed to 48°C (Amann *et al.*, 1995) prior to use to ensure at least the same stringency level as the hybridization buffer. Samples are stained with a general DNA dye, such as DAPI or YO-PRO-1, to visualize non-hybridized cells (Amann *et al.*, 1995). Plant material, which is common in sludge samples, is a major cause of

autofluorescence when viewed with an epifluorescent microscope (Bond *et al.*, 1999). A negative control is used to check for autofluorescence in which the sample is exposed to buffer containing no probe (Bond *et al.*, 1999).

**Quantitative PCR.** Quantitative PCR (qPCR), or real-time PCR, can determine the absolute quantity of a single nucleic acid target sequence within an unknown sample in real time (Applied Biosystems, 2005). Real time data acquisition is possible in qPCR because amplification of the target DNA is detected by fluorescent molecules. DNA quantification is determined by the increase in fluorescence at the end of each qPCR cycle (Applied Biosystems, 2005). In contrast, regular PCR measures DNA amount upon the completion of all PCR amplification cycles (Applied Biosystems, 2005). There are two main methods to quantify DNA amplification in qPCR.

The first method uses SYBR green. SYBR green is a general stain for double stranded DNA (Applied Biosystems, 2005). SYBR green binds to any double stranded DNA molecules present after amplification (Applied Biosystems, 2005). SYBR green will not bind to single stranded DNA such as probes, primers, or single strand template (Applied Biosystems, 2005). However, this method is imprecise as SYBR green will also bind to non-specific PCR products.

The second and more precise method is the Taqman probe assay (Applied Biosystems, 2005). A Taqman probe is an oligonucleotide DNA probe which is complimentary to the target DNA strand (Applied Biosystems, 2005). The Taqman probe is linked to a fluorescent reporter on the 5' end and a quencher on the 3' end



(Applied Biosystems, 2005). When excited, the reporter emits energy as light at a specific wavelength. As long as the reporter and quencher are both attached to the probe, the quencher prevents the reporter from fluorescing by absorbing the energy normally released by the reporter as light (Applied Biosystems, 2005). During qPCR, *Taq* DNA polymerase excises the probe (Applied Biosystems, 2005). This releases the reporter from the quencher, and the reporter is able to fluoresce. Similar to FISH probes, Taqman probes can be designed to encompass broad groups, such as most prokaryotes with probe BAC 338 (Murray *et al.*, 1998), or specific groups, such as TM7 with probe TM7 1093F (Brinig *et al.*, 2003). A fluorescent passive reference is provided in the ABI master mix. The machine reads the reference fluorescence and normalizes reporter fluorescent fluctuations between wells due to errors in volume or master mix concentration (Applied Biosystems, 2005).

A study by Brinig *et al.* (2003) looking at the relative prokaryotic populations in the human subgingival crevice developed a qPCR protocol to quantify total bacterial and TM7 specific bacterial populations. Broad range 16S rRNA gene primers, BAC 8F and BAC 515R (Lane *et al.*, 1985), were used for total bacterial quantification (Brinig *et al.*, 2003). TM7 16S rRNA gene specific primers TM7 1177R and TM7 910F were used for TM7 quantification reactions (Brinig *et al.*, 2003). Thermocycling conditions were performed according to manufacturer protocols. The bacterial quantification reaction had three annealing temperatures: 55°C for 30 sec, 60°C for 45 sec, and 65°C for 15 sec (Brinig *et al.*, 2003). The TM7 quantification reactions had a 61°C annealing temperature (Brinig *et al.*, 2003). A solution with a known copy number of a plasmid

with TM7 positive insert (amplified by BAC 8F and TM7 1177R) was serially diluted to generate a standard curve for each individual quantification run (Brinig *et al.*, 2003).

Some prokaryotic chromosomes have more than one copy of the same 16S rRNA gene (Crosby & Criddle, 2003). A study comparing National Center for Biotechnology Information (NCBI), TIGR (now The Gene Index, TGI), and Ribosomal RNA Operon Copy Number (rrndb) (Klappenbach *et al.*, 2001) databases showed that some prokaryotes have as many as 15 16S rRNA gene copies per cell with most having somewhere between one and seven copies (Acinas *et al.*, 2004). That same study recommended a 3.5 average when using 16S rRNA genes per cell for quantifying the total bacterial community (Acinas *et al.*, 2004). TM7 appear to have two rRNA gene copies based on two independent studies able to amplify the whole TM7 genome (Marcy *et al.*, 2007; Podar *et al.*, 2007). A two 16S rRNA gene copy number per cell will be used when quantifying TM7 in activated wastewater.

### **TM7 Genome**

Multiple displacement strand amplification sequencing has been used to characterize the complete genome for two TM7 members, one human oral TM7 (Marcy *et al.*, 2007) and a human oral related TM7 from soil (Podar *et al.*, 2007). A “metagenome” for each TM7 was created based on a composite of the sequences from each individual TM7. The word “metagenome” was used to describe the sequenced genome because the sequenced DNA is from a single organism isolated from a mixed community (Tringe & Rubin, 2005). However, there was the risk of sequence-

contaminated DNA being included into the genome. Also, the metagenome may not be the whole genome. The metagenome from the oral TM7 had an estimated 10% or less *Leptiricha* DNA contamination (Marcy *et al.*, 2007). The metagenome from the soil TM7 excluded DNA sequences that were possibly from *Pseudomonas* based on GC ratio, which possibly removed genuine TM7 genes (Podar *et al.*, 2007).

**Oral TM7 metagenome.** The oral TM7 metagenome contained 2.86 megabases (Marcy *et al.*, 2007). One gene number estimate placed 3,245 genes and gene fragments in the full genome, or 1,474 genes with a conservative estimate (Marcy *et al.*, 2007). When cultured bacterial genomes were compared to each other, 85% of protein genes were similar to each other, 15% of protein genes had <60% similarity, and only 3% had <30% similarity (Marcy *et al.*, 2007). When the oral TM7 metagenome was compared to previously published, cultured bacterial genomes 80% of protein genes had <60% similarity and 33% had <30% similarity (Marcy *et al.*, 2007). Other uncultivated organism metagenomes, such as one developed from *Rhodopirellula baltica*, were similar to the oral TM7 metagenome in that most protein genes were dissimilar to cultured bacterial genomes (Fuchsman & Roca, 2006).

Genes in the oral TM7 metagenome were most similar to *Bacilli*, *Clostridia* and *Fusobacteria* genes further suggesting TM7 is related to Gram-positive bacteria (Marcy *et al.*, 2007). Genes were found for common pathways such as: glycolysis, tricarboxylic acid cycle, nucleotide biosynthesis, and some amino acid biosynthesis (Marcy *et al.*, 2007). More specialized genes have also been characterized. Genes for IV pili

biosynthesis were noted (Marcy *et al.*, 2007), which could be used for cell adhesion (a possible virulence factor), biofilm formation (another possible virulence factor), and gliding motility (Varga *et al.*, 2006). A sortase, related to sortases found in *Firmiutes* and *Actinobacteria*, and UDP-*N*-acetylmuramyl tripeptide synthetase, found in *Bifidobacteria*, are characteristic of Gram-positive bacteria (Marcy *et al.*, 2007). UDP-*N*-acetylmuramyl tripeptide synthetase is a possible virulence factor and has been linked to chronic granulomatous inflammation (Simelyte *et al.*, 2003). Genes for metabolism, glycosyl hydrolase like  $\alpha$ -amylase and oligo-1, 6-glucosidase genes, have been characterized suggesting oral TM7 can use oligosaccharides as growth substrates (Marcy *et al.*, 2007). ABC protein transporter genes and arginine deiminase pathway genes have been characterized suggesting that oral TM7 can use amino acids as growth substrates, specifically arginine (Marcy *et al.*, 2007).

**Soil TM7 metagenome.** The soil TM7 metagenome was not as complete as the oral TM7 (only 680 genes were identified out of 670,000 nucleotide bases), but the total genome was estimated to be 3-3.5 megabases large (Podar *et al.*, 2007). Common genes for glycolysis, nucleotide synthesis, lipid synthesis, amino acid transformations, and cell wall biosynthesis were identified (Podar *et al.*, 2007). A novel gene insertion in ATP synthase operon, suggesting possible viral origin, was noted (Podar *et al.*, 2007). Type IV pili genes were characterized as was found in the oral TM7 metagenome (Podar *et al.*, 2007). Genes for cell protection: a gene similar to H<sup>+</sup> antiporter 1 gene (antimicrobial drug resistance), a cytochrome p450 gene (protection against toxic compounds), and a

possible conductance mechanosensitive channel (protection against osmotic lysis) were noted (Podar *et al.*, 2007). Genes for extracellular communication: type IV secretion system genes used in biofilm formation and interaction with other bacteria (Baron, 2005) and a possible autoinducer 2 exporter for interspecific signaling were discovered (Podar *et al.*, 2007).

While outside the scope of this study, further work to characterize TM7 in activated wastewater would benefit from the above mentioned metagenomes. New PCR primers could be designed to target the newly identified TM7 genes. It is possible to perform “anchored PCR.” This would use one of the known TM7 16S rRNA gene primers paired to a newly designed primer. The known TM7 primer would provide specificity and the new primer would target a genome portion outside the 16S rRNA gene. This would provide insight into parts of the genome. Substrate tracking autoradiography fluorescent *in situ* hybridization (STAR-FISH) could be used to confirm nutrient uptake genes (Ouverney & Fuhrman, 2000). STAR-FISH combines fluorescently labeled DNA probes to target specific organisms in a mixed community and general DNA stains, such as DAPI or YO-PRO-1, with radioactive labeled, defined growth substrates to determine if organisms are accumulating those molecules.

## **Materials and Methods**

### **Sample Collection and Processing**

All samples consisted of 200 ml of secondary treatment activated wastewater obtained from six sites at the San Jose/Santa Clara Water Pollution Control Plant. For sample collection dates, see Table 1. The six sampled sites were located in four secondary treatment aerator tanks with the following designations: A1Q1 East, A1Q1 Middle, A1Q3 East, B1Q1 East, B1Q1 Middle, and B1Q3 East (Figure 1). Samples were taken from these general locations where dark, stable foam formed on eddies in the aeration tank. This foam corresponds to areas that have old sludge, high particulate content, and high nitrogen gas levels. Experimental sludge samples were removed from tanks using sampling scoops provided by the treatment plant. Samples were transferred from the scoops to sterile 500 ml whirl-pak bags (Nasco, Fort Atkinson, WI) using sterile technique. Samples were stored on ice in foam-insulated coolers during transit.

Samples were processed in the laboratory within 1 h of sampling, as described below. Wastewater aliquots (50 g) of settled large particulates from each sample site were weighed in a Mettler Toledo AB54-S balance (Mettler Toledo, Columbus, OH). Weighed aliquots were added to sterile Fisherbrand 50 ml centrifuge tubes (Fisher Scientific, Waltham, Massachusetts). Samples were centrifuged in a Sorvall RC 5C centrifuge with a KA 21.50 rotor (Fisher Scientific, Waltham, Massachusetts) at 9,000 g

for 15 min. The supernatant was discarded. Wastewater pellets were stored at 4°C until for no more than 48 h until DNA extraction.

Between February 2008 and June 2009 the secondary aeration treatment tanks on the “A” half were out of service. Experimental sludge samples were taken from the “B” half from B1Q1 East, B1Q1 Middle, and B1Q3 East that correspond to previous “A” samples sites (Figure 1). Sampling conditions and processing were the same as described above.

Table 1. Dates and sample sites where raw wastewater samples were collected and DNA was extracted using a modified bead-beater half-lysis protocol.

Sites	Dates								
	8/13 2007	3/19 2008	4/27 2008	5/14 2009	6/15 2009	9/12 2009	10/14 2009	12/22 2009	1/28 2010
A1Q1 East	X	X	X		X		X	X	
A1Q1 Middle	X	X	X				X	X	
A1Q3 East	X	X	X				X	X	
B1Q1 East				X	X	X			X
B1Q1Middle				X	X	X			X
B1Q3 East				X	X	X			X

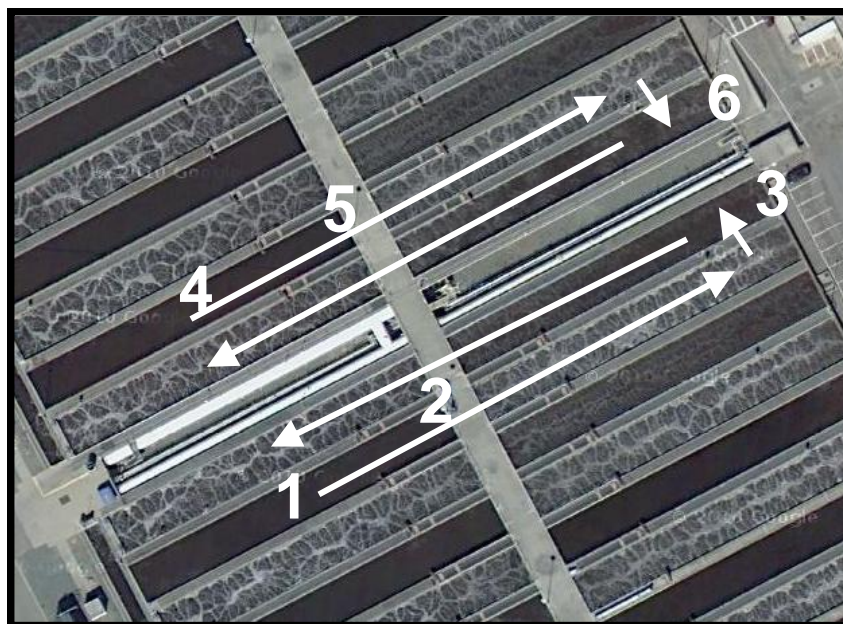


Figure 1. Secondary aerator treatment tanks at the San Jose/Santa Clara Water Pollution Control Plant (Google Maps, 2010). Sampled locations are as follows: 1) A1Q1 Mid, 2) A1Q1 East, 3) A1Q3 East, 4) B1Q1 Mid, 5) B1Q1 East, and 6) B1Q3 East. Arrows indicate direction of the activated wastewater treatment flow.

#### **DNA Extraction: Mo Bio PowerSoil Kit**

Bacterial genomic DNA was extracted from samples obtained from A1Q1 East and A1Q1 Middle using a Mo Bio PowerSoil DNA extraction kit (Mo Bio Laboratories Inc, Carlsbad, California) on 04/23/07 and 08/13/07. Pelleted sludge aliquots (0.25 g) were added to powerbead tubes provided in the Mo Bio PowerSoil DNA extraction kit. DNA was extracted according to manufacturer protocol (protocol version 10222009). DNA concentration and purity were determined using Nanodrop ND-1000



spectrophotometer with ND-1000 version 3.1.2 software (Thermo Scientific). DNA molecular size was determined via agarose gel electrophoresis. Extracted nucleic acid was stored at -20°C.

### **DNA Extraction: Bead-Beater Half-Lysis**

Bacterial genomic DNA was also extracted from each pelleted sample using a modified bead-beater half-lysis protocol as described by Roh *et al.* (2006). Pelleted sludge aliquots (0.25 g) were used for each extraction. DNA concentration and purity were determined as described above. Extracted nucleic acid was stored as described above.

Initial PCR experiments to amplify TM7 16S rRNA genes from genomic DNA prepared through the bead-beater half-lysis protocol showed PCR amplification inhibition. To overcome PCR inhibition, extracted DNA was serially diluted from 1/10 to 1/10,000 in sterile water prior to PCR. After agarose gel electrophoresis and spectrophotometer comparison between the Mo Bio PowerSoil and modified bead-beater half-lysis extracted nucleic acid, the modified bead-beater half-lysis protocol was used for all further DNA extraction. Modified bead-beater protocol gave higher DNA yields and larger DNA molecules for further use when compared to the PowerSoil kit (See Figure 6 and Table 6 in Results under “DNA Extraction” section). To improve PCR amplification, 1/100 and 1/1,000 dilutions of extracted nucleic acid were used for further experiments.

## **TM7 16S rRNA Gene Amplification**

PCR amplification of the TM7 16S rRNA genes was performed on 1/100 and 1/1,000 dilutions of extracted nucleic acid. Each PCR reaction had a total volume of 25  $\mu$ l. Each reaction consisted of: 1x PCR Buffer B (Fisher Scientific, Waltham, Massachusetts), 1.5 mM MgCl<sub>2</sub>, 0.2 mM of each deoxynucleoside triphosphate, 0.2 pmol of each forward and reverse primer,  $1.25 \times 10^{-2}$  units of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, California), and 3 to 30 ng of nucleic acid. PCR thermal cycling was carried out in an Applied Biosystems 2720 Thermal Cycler (Applied Biosystems, Foster City, California). Cycling conditions were as follows: initial denaturation step of 96°C for 3 min followed by 40 cycles of 94°C for 1 min, 61°C for 1 min, and 72°C for 1 min, with a final elongation step of 72°C for 3 min. Initial reactions used a broad-range general bacteria forward primer, BAC 8F, with a TM7 specific reverse primer, TM7 1177R (Table 2). Expected product size was approximately 1170 base-pair long.

DNA extractions were also screened for TM7 subgroup SBG3 using BAC 8F and ENV SBG3 1112R primers (Table 2). Each reaction was set up as described above with the exception that ENV SBG3 1112 primer was used instead of TM7 1177R. Cycling conditions for ENV SBG3 1112R were as follows: initial denaturation step of 96°C for 3 min followed by 40 cycles of 94°C for 1 min, 57°C for 1 min, and 72°C for 1 min, with a final elongation step of 72°C for 3 min. Expected product size was approximately 1100 base-pair.

Table 2. Primers and probes.

<b>Specificity</b>	<b>Name</b>	<b>Sequence</b>
<b>General TM7 Group Specific</b>	TM7 910F PCR Primer	CAT AAA GGA ATT GAC GGG
	TM7 1177R PCR Primer	GGA AGG AGG GGA TGA TGT
	TM7 1093F qPCR Probe	AGT CCA TCA ACG AGC GCA
	TM7 905 FISH Probe	CCG TCA ATT CCT TTA TGT TTT
<b>TM7 Subgroup Environmental SBG3 Specific</b>	SBG3 997F PCR*	TCC CGA GAA GAT TTA CG
	ENV SBG3 1112R PCR*	TCA ACT ATT CAC AAG GG
	SBG3 qPCR Probe*	ATC TGT CAC CGA GTT CCA
	SBG3_1 FISH Probe*	ATC TGT CAC CGA GTT CCA
<b>General Bacteria Specific</b>	BAC 8F PCR and qPCR	AGA GTT TGA TCC TGG CTC AG
	BAC 515R PCR Primer	KAC CGC GGC KGC TGG CA
	BAC 338I qPCR Probe	GCT GCC TCC CGT AGG AGT
	BAC 338II qPCR Probe	GCA GCC ACC CGT AGG TGT
	BAC 338III qPCR Probe	GCT GCC ACC CGT AGG TGT
	BAC 338 FISH Probe	GCT GCC TCC CGT AGG AGT
<b>Cloning Vector Specific</b>	M13 Forward PCR primer	GTA AAA CGA CGG CCA G
	M13 Reverse PCR primer	CAG GAA ACA GCT ATG AC

\*Designed for this project.

All PCR products were loaded into a 1% agarose gel (International Biotechnologies Inc., New Haven, Connecticut) containing 0.04% Ethidium Bromide. Gels were run in a Horizon 58 gel electrophoresis system (Invitrogen, Carlsbad, California) using a FB300 electrophoresis unit (Fisher Scientific, Waltham, Massachusetts) at 80 V with 400 mA of current for 25 min. Gels were visualized using a Bio-Rad Universal Hood II UV Illuminator (Bio-Rad, Hercules, California). Gel images

were processed and annotated with Quantity One version 4.6 software (Bio-Rad, Hercules, California).

### **Cloning of PCR Products**

PCR for cloning was performed on extracted genomic DNA from A1Q1 Middle, A1Q1 East, and A1Q3 East sample sites. PCR for cloning conditions were the same as described above with the exception that each reaction contained 50  $\mu$ l of the same master mix and 30 cycles, instead of 40, were performed. A lower PCR cycle number was used to avoid PCR biased amplification (Innis *et al.*, 1990). All PCR products were loaded into a 1% agarose gel containing 0.04% Ethidium Bromide. Gels were processed and visualized as above. Sample bands that correlated to the correct size, based on DNA ladder of known molecular weight, were excised and placed into sterile microfuge tubes.

PCR products were purified using QIAquick Gel Extraction Kit (Qiagen, Valencia, California). Purified PCR products were ligated into pCR2.1-TOPO vector (Invitrogen, Carlsbad, California) overnight at 14°C. Vector-insert ligation mixture was transformed into One Shot TOP10 competent cells (Invitrogen, Carlsbad, California) as per manufacturer protocol (protocol version W). Transformed cells were plated on LB agar plates with 40 mg/ml X-gal and 50  $\mu$ g/ml Kanamycin. Plates were incubated at 37°C for 16 h. Clones were screened via PCR amplification. PCR cycling conditions for clone screening were as described above with the exception that M13 Forward and M13 Reverse primers (Table 2) were used. Plasmids from representative clones were isolated

using QIAprep Spin Miniprep Kit (Qiagen, Valencia, California) per manufacturer protocol (protocol version 11/2006) and sequenced. Sequencing was performed by Geneway Research (Hayward, California) or Sequetech (Mountain View, California). Sequencing was initially done with TM7 1177R primer only. Sequences with high percent similarity to published TM7 sequences (identified via NCBI-BLAST) were re-sequenced with M13 Forward and M13 Reverse primers to acquire full-length sequences of the 16S rRNA gene PCR amplicons.

### **Phylogenetic Analysis**

All 16S rRNA gene sequences were stored and organized in LABnovo, a lab-centric web service for sequence retrieval, manipulation, and analysis developed in the Ouverney lab. LABNovo automatically compared each sequence to available databases using Basic Local Alignment Search Tool (BLAST). A spreadsheet was generated with the corresponding BLAST top hits and coverage for each sequence. Only sequences showing TM7 as the top BLAST hit with at least 98% sequence coverage were further analyzed. TM7 partial sequences were imported into CodonCode Aligner software (CodonCode Corporation, version 3.0.2, Dedham, Massachusetts). Individual sequences were cleaned based on the overall quality of the bases as determined by the chromatogram. Deletion of poor quality regions, as well as PCR primer and vector DNA sequences was performed. Consensus sequences of approximately 1,140 nucleotides were generated from the cleaned partial sequences from each insert side (individual forward and reverse sequences acquired from the same plasmid and combined based on

the overlapping regions of high similarities). Consensus sequences were checked in BLAST to confirm homology to TM7 with high percent coverage. Furthermore, consensus sequences were screened for chimeras using Chimera Check with Bellerophon version 3 with default settings and 1,000 base-pair threshold (DeSantis *et al.*, 2006). Consensus sequences were aligned for ARB using Greengenes NAST aligner with default settings and 1,000 base-pair threshold (DeSantis *et al.*, 2006).

Aligned sequences were incorporated in ARB software package (Ludwig *et al.*, 2004). Alignment of each sequence was refined individually, based on 16S rRNA primary and secondary structure in their relationship to the 236,469 documented 16S rRNA gene sequences in the ARB database. Phylogenetic analysis was performed on sequences by using neighbor-joining algorithms without corrections. A filter was generated specifically for the working data set using ARB automatic filter tool and was manually refined. The generated filter, TM7\_filter03302010, consisted of 1,035 columns, covering the most 1,035 conserved nucleotide bases in the working data set. Filter TM7\_filter03302010, used for phylogenetic tree generation, gave a higher stringency for separating similar TM7 sequences used for phylogenetic analysis than ARB generated or ARB default bacterial filters. The sequence of *Escherichia coli* strain E24377A was used as an out-group to root all related TM7 sequences to one common ancestor in the tree. The strength of topology was tested by bootstrap re-sampling analysis of 1,000 replications.

## **Quantitative PCR**

Quantitative PCR was performed on extracted nucleic acid dilutions that contained TM7 DNA when PCR resulted in the expected amplification product.

Quantitative PCR was performed using an Applied Biosystems 7300 Real Time PCR System (Applied Biosystems, Foster City, California), which included the use of Taqman probe assays. Three subgroups of prokaryotes were quantified via qPCR in each sample: total bacteria (domain level), total TM7 (phylum level), and the environmental SBG3 subgroup of TM7. Each reaction is described in detail below.

Sample reactions had a final concentration of 3 to 30 ng of extracted nucleic acid from activated wastewater samples. A positive control plasmid was generated that contained perfect target sequences for all qPCR primers and probes used for total bacterial, total TM7, and environmental SBG3 quantification. A standard curve was generated using a serial dilution of this positive control plasmid. Final tube concentration ranged from 65 ng to  $6.5 \times 10^{-7}$  ng ( $1 \times 10^8$  to 1 calculated molecules of plasmid), as determined through Nanodrop ND-1000 spectrophotometer. Three separate runs were performed for each general bacterial, total TM7, and environmental SBG3 subgroup quantification. Each run had its own standard curve. Each sample and serial dilution was replicated in triplicate for each run. For setting experimental parameters, cycling conditions, and analyzing data, 7300 System SDS software version 1.4.0.25 with Relative Quantification Study plug-in was used (Applied Biosystems, Foster City, California). Data was compiled and analyzed using Microsoft Excel 2003.

Since qPCR targeted 16S rRNA genes, qPCR gene amplification values were converted to cells/g activated wastewater based on a number of corrections. Conversion included: **i**) Corrections for nucleic acid dilutions used for PCR amplification, **ii**) corrections for changes in volume during DNA extraction, **iii**) corrections for changes in volume during activated wastewater processing (see Table 3 and Table 4 in “Results” section), as well as **iv**) an estimated 3.5 average rRNA gene copy number per cell (Klappenbach *et al.*, 2001) for total bacterial quantification or an estimated 2.0 average rRNA gene copy number per cell for total TM7 and environmental SBG3 quantification (Marcy *et al.*, 2007; Podar *et al.*, 2007). These corrections were used to calculate an estimated number for total bacterial, total TM7, and environmental SBG3 cells/g of raw activated wastewater (Figure 2).

To determine stability of TM7 and environmental SBG3 populations throughout the year, qPCR was performed on extracted DNA that corresponded to the four seasons. Experimental setup and cycling conditions are described below.



$$\frac{\text{Cells}}{\text{g Raw Activated Wastewater}} = \frac{\text{rRNA Gene Copy Number}}{1 \mu\text{l DNA Solution}} \times \frac{1 \text{ Cell}}{3.5 \text{ or } 2 \text{ rRNA Gene Copy Number}} \times \frac{100}{1} \times \frac{50 \mu\text{l DNA}}{0.25 \text{ g Pelleted Activated Wastewater}} \times \frac{X \text{ g Pelleted Activated Wastewater}}{Y \text{ g Raw Activated Wastewater}}$$

**Raw qPCR Data**      **Correction For Gene Copy Number Per Cell (3.5 for Total Bacteria, 2 for TM7 and SBG3)**      **Correction for 1/100 Dilution Prepared For PCR**      **Correction From DNA Extraction**      **Correction From Pelleting Activated Wastewater (Weights from Table 3 and Table 4)**

35

Figure 2. Method for converting qPCR rRNA gene copy number/ $\mu\text{l}$  DNA solution to cells/g raw activated wastewater.

Conversions include corrections for gene copy number per cell, dilution prepared for PCR amplification, changes in volume during DNA extraction, and wastewater processing.

**Total bacterial quantification.** Each qPCR reaction mixture had a final volume of 20  $\mu$ l, which contained 10  $\mu$ l Taqman Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems, Foster City, California), 1.8 pmol of forward primer BAC 8F, 3.6 pmol of reverse primer BAC 515R, 0.1 pmol of each Taqman probe: BAC 338I, BAC 338II, and BAC 338III (Table 2) with FAM reporter and Black-Hole quencher, and an additional  $2.5 \times 10^{-2}$  units of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, California). Cycling conditions for general bacterial reactions were as follows: initial denaturation step of 95°C for 10 min followed by 50 cycles of 95°C for 30 sec, 55°C for 30 sec, 60°C for 45 sec, 65°C for 15 sec, and 72°C for 30 sec. Fluorescence levels were read at the end of each 72°C elongation step.

**Total TM7 quantification.** Each qPCR reaction mixture had a final volume of 20  $\mu$ l, which contained 10  $\mu$ l Taqman Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems, Foster City, California), 0.9 pmol of forward primer TM7 910F, 0.9 pmol of reverse primer TM7 1177R, 0.1 pmol of Taqman probe TM7 1093F with VIC fluorescent reporter and Black-Hole quencher, and an additional  $2.5 \times 10^{-2}$  units of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, California). Cycling conditions for TM7 specific reactions were as follows: initial denaturation step of 95°C for 10 min followed by 50 cycles of 95°C for 30 sec, 61°C for 30 sec, and 72°C for 30 sec. Fluorescence levels were read at the end of each 72°C elongation step.

**Environmental SBG3 quantification.** Each qPCR reaction mixture had 20  $\mu$ l final volume, which contained 10  $\mu$ l Taqman Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems, Foster City, California), 0.9 pmol of forward primer SBG3 997F, 0.9 pmol of reverse primer ENV SBG3 1112R, 0.1 pmol of Taqman probe SBG3 qPCR probe with FAM reporter and Black-Hole quencher, and an additional  $2.5 \times 10^{-2}$  units of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, California). Cycling conditions for SBG3 subgroup specific reactions were as follows: initial denaturation step of 95°C for 10 min followed by 60 cycles of 95°C for 30 sec, 57°C for 30 sec, and 72°C for 30 sec. Fluorescence levels were read at the end of each 72°C elongation step.

**Cloning of quantitative PCR products.** Cloning of SBG3 subgroup and TM7 qPCR products was performed to check that TM7 and SBG3 qPCR products were from target organisms and not contaminants or non-specific amplifications. Cloning of quantitative PCR product was performed separately from qPCR runs used for quantification as only 25 PCR thermocycles were performed to limit PCR biased amplification (Innis *et al.*, 1990). Each qPCR run for cloning contained one 25  $\mu$ l positive control sample, using TM7 positive control plasmid (which matches SBG3 qPCR primers and probe sequences) at  $6.5 \times 10^4$  gene copy number per  $\mu$ l.

Each 25  $\mu$ l TM7 specific qPCR cloning reaction had at final volume: 12.5  $\mu$ l Taqman Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems, Foster

City, California), 1.13 pmol of forward primer TM7 910F, 1.13 pmol of reverse primer TM7 1177R, 0.1 pmol of Taqman probe TM7 1093F with VIC reporter and Black-Hole quencher, and an additional  $2.5 \times 10^{-2}$  units of Taq DNA polymerase (Applied Biosystems, Foster City, California). Cycling conditions for TM7 specific qPCR cloning reactions were as described above.

Each 25  $\mu$ l environmental SBG3 subgroup specific qPCR cloning reaction had at final volume: 12.5  $\mu$ l Taqman Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems, Foster City, California), 1.13 pmol of forward primer SBG3 997F, 1.13 pmol of reverse primer ENV SBG3 1112R, 0.1 pmol of Taqman probe SBG3 qPCR probe with FAM reporter and Black-Hole quencher, and an additional  $2.5 \times 10^{-2}$  units of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, California). Cycling conditions for environmental SBG3 specific qPCR cloning reactions were as described above

The AmpliTaq DNA polymerase provided in Taqman Universal PCR Master Mix, No AmpErase UNG, is a high fidelity polymerase with proof-reading abilities that does not add an “A” to the 3’ amplicon end. AmpliTaq DNA polymerase generated PCR products were incompatible with TA cloning kits without additional processing. Samples were amplified for an additional 10 cycles after the addition of 5 units of Flexi Go Taq DNA polymerase (Promega, Madison, Wisconsin) to create A-overhangs. Final qPCR products were loaded into E-Gel SizeSelect 2% gel (Invitrogen, Carlsbad, California) in an E-Gel Safe-Imager real-time transilluminator (Invitrogen, Carlsbad, California). Products were run using preprogrammed SizeSelect 2% running conditions till positive control bands had migrated into collection wells.

Isolated qPCR products were ligated into pCR2.1-TOPO vector (Invitrogen, Carlsbad, California) and transformed into One Shot TOP10 competent cells (Invitrogen, Carlsbad, California) as described above. The resultant clones were screened by PCR as described above with the exception that vector specific (qPCR) primers were used. Plasmids were extracted as described above. Sequencing was performed by Sequetech, Mountain View, California. Sequencing was done with M13 Forward only since qPCR products were only 270 base-pair long for TM7 and 125 base-pair long for environmental SBG3 reactions.

### **Preliminary Microscopy of Bacteria in Wastewater**

Prior to DNA extraction, raw wastewater was removed (2.5 ml) and placed into sterile 15 ml conical tubes. These samples were mixed with 0.45  $\mu\text{m}$  filtered 100% methanol (2.5ml), vortexed, and allowed to fix overnight at 4°C. Methanol fixed wastewater samples, when not disturbed, exhibited two separate phases. The bottom phase contained a loose collection of large settled particulates. Above this was an aqueous layer, which was slightly more turbid than distilled water when compared using the human eye. Samples used for microscopy were taken from the aqueous portion just above the settled large particulates. Each sample was added to 24 well Teflon etched, ADD-Cell coated slides (Fisher Scientific, Waltham, Massachusetts) in 5  $\mu\text{l}$  amounts. Samples were air dried for 10 min at room temperature.

Bacterial cells were fluorescently stained with YO-PRO-1 by adding 10  $\mu$ l of a working solution of 1  $\mu$ M YO-PRO-1 DNA dye (Invitrogen, Carlsbad, California) to each well. Staining was performed at room temperature for 10 min in the dark. Slides were washed by submersion into room temperature MilliQ distilled water for 3 min. Immersol 518F microscopy oil (Zeiss, Oberkochen, Germany) was added in 2  $\mu$ l amounts to each well. Corning No. 1, 60 mm cover glass was glued to each slide with nail polish. Slides were visualized using an Axioscope 2 microscope with HBO 100 light source (Zeiss, Oberkochen, Germany). Images were captured using a MC80 DX camera and analyzed with Axiovision software (Zeiss, Oberkochen, Germany).

Original microscopic results on unaltered wastewater samples showed that bacteria aggregated on suspended particulates called flocs. Large amounts of bacteria were observed to be attached to these flocs. Pre-treatment experiments were performed on wastewater prior to performing FISH to dissociate bacteria from flocs and obtain clear images of whole bacterial cells without obstructions from particulate matter.

### **Sonication/Emulsifier Optimization**

To test the best means for breaking up flocs, pelleted wastewater sludge was separated into 250  $\mu$ l aliquots. To 3 aliquots, PBS at pH 6.2 was added (1.5 ml). To another 3 aliquots, PBS pH 6.2 with 0.1% Tween 80 was added (1.5 ml). All samples were vortexed for 2 min. All samples were sonicated in a Branson 1510 sonicator at 60 hertz. One pair of samples with and without Tween 80 was sonicated for 2 min. Another pair was sonicated for 3 min and another for 5 min. Samples were vortexed and large particulates were allowed to settle. Supernatant above the settled particulates was removed and centrifuged at 5,000 g for 5 min. Pellets were resuspended in 250  $\mu$ l of PBS at pH 7.0.

Resuspended separated supernatant was added to 24 well Teflon etched ADD-Cell coated slides in 5  $\mu$ l amounts and allowed to air dry for 20 min. Slides were processed as described above.

### **Fluorescent *in situ* Hybridization**

Probes for FISH were commercially synthesized and 5' labeled with Cy3, Cy5, and Alexa-532. The probes used were: BAC 338 Cy3, TM7 905 Cy3, TM7 905 Cy5 (Operon, Huntsville, Alabama), and SBG3\_1 FISH Alexa-532 (Invitrogen, Carlsbad, California) (Table 2). All probes were diluted to a stock concentration of 400 ng/ $\mu$ l prior to use. Raw wastewater was added (2.5 ml) to sterile 15 ml conical tubes. To make an overall 0.1% solution of Tween 80, Tween 80 was added (2.5  $\mu$ l). Samples were

vortexed for 2 min. Samples were sonicated at 60 hertz for 3 min. Cells were fixed as described above. Samples were added to slides and dried as described above.

Initial FISH was performed using only BAC 338 Cy3 and TM7 905 Cy3 to determine FISH feasibility on wastewater samples. A 20% formamide hybridization buffer was prepared with 300  $\mu\text{l}$  of 0.2  $\mu\text{m}$  filtered MilliQ water, 100  $\mu\text{l}$  of 0.45  $\mu\text{m}$  filtered formamide, 90  $\mu\text{l}$  of 5 M NaCl, 9.5  $\mu\text{l}$  of 1 M Tris-HCl pH 7.2, and 0.5  $\mu\text{l}$  of 10% SDS per 500  $\mu\text{l}$  of hybridization buffer. Hybridization buffer was separated and stock probe solution added to give a working probe hybridization buffer with probe concentration at 5 ng/ $\mu\text{l}$ . Each well was hybridized with 5  $\mu\text{l}$  of probe hybridization buffer except blank wells, which were hybridized with hybridization buffer only as an autofluorescence control. A total of 15 wells were hybridized with TM7 905 Cy3 probe buffer solution, six wells with BAC 338 Cy3 probe buffer solution, and three wells with hybridization buffer with no probe. Slides were hybridized for 2 h at 43°C.

A wash buffer was prepared with 46.8 ml MilliQ water, 2.15 ml 5 M NaCl, 1.0 ml 1 M Tris-HCl pH 7.2, and 50  $\mu\text{l}$  10% SDS. This concentration of NaCl gives a wash stringency equivalent to 20% formamide concentration. Wash buffer and distilled water were warmed to 48°C for 2 h in the hybridization oven. Slides were washed by rinsing remaining probe buffer solution with warmed water and were submerged into 50 ml 48°C wash buffer for 15 min. Wash buffer was rinsed off with 48°C water. Slides were dried at 43°C for 10 min. Slides were stained and coverslips mounted as described above. Slides were visualized with a Zeiss Axioscope A.1 epifluorescence microscope with X-cite 120Q light source (Zeiss, Oberkochen, Germany). Zeiss filter 43 was used for Cy3



imaging and filter 38 for YO-PRO-1 imaging. Pictures were registered in Axiovision 4.7 software (Zeiss, Oberkochen, Germany). Pictures were compiled and annotated in Microsoft PowerPoint 2003.

Additional FISH used TM7 905 Cy5 and SBG3\_1 FISH Alexa-532 to target environmental SBG3 subgroup of TM7. A 30% formamide hybridization buffer was prepared with 250  $\mu$ l of 0.2  $\mu$ m filtered MilliQ water, 150  $\mu$ l of 0.45  $\mu$ m filtered formamide, 90  $\mu$ l of 5 M NaCl, 9.5  $\mu$ l of 1 M Tris-HCl pH 7.2, and 0.5  $\mu$ l of 10% SDS. Slide hybridization was performed as described above with the exception that 21 wells were hybridized with TM7 905 Cy5 and SBG3\_1 FISH Alexa-532 and three wells with hybridization buffer with no probe. A wash buffer was prepared with 48 ml MilliQ water, 1.02 ml 5 M NaCl, 1.0 ml 1 M Tris-HCl pH 7.2, and 50  $\mu$ l 10% SDS. This concentration of NaCl gives a wash stringency equivalent to 30% formamide concentration. Slides were washed as described above. Slides were stained and coverslips mounted as described above with the exception that a 1  $\mu$ M DAPI dye (Invitrogen, Carlsbad, California) was used in place of YO-PRO-1 dye. Slides were visualized and images processed as above. Zeiss filter 50 was used for Cy5 imaging, filter 43 for Alexa 532 imaging, and filter 49 for DAPI imaging.

## Results

### Sludge Processing

Each wastewater sample contained a mixed particulate suspension. After 5 min, samples would separate into an aqueous layer and semi-solid layer. About 50 g of this semi-solid layer would yield on average 6.13 g pelleted sludge with a range from 3.88 g to 9.68 g as seen in Tables 3 and 4.

To detach bacteria from observed particulate flocs in wastewater, a protocol with Tween 80 addition, shaking, and sonication was used. Detaching and separating bacteria from flocs assisted FISH microscopic imaging in three ways. First, separating bacteria from flocs facilitated access for probes and dyes to enter cells. Second, visualization of cells for counting was easier, as there was less background clutter and less three dimensional structures that overlapped cells. Third, loss of sample was prevented. During early FISH experimentation, flocs would clump as the mounted sample dried, creating a mesh that peeled off while performing FISH. This decreased the number of wells that could have been used for FISH analysis. By separating bacteria from flocs, less particulate material was mounted on slides. This would reduce the chance of a mesh forming and sloughing off during FISH, which increased the number of wells that can be used for visualization.

Table 3. Raw wastewater and wastewater pellet weights for A1Q1 East, A1Q1 Middle, and A1Q3 East. Raw wastewater was centrifuged at 9,000 g for 15 min to pellet large particulates. Aliquots from pelleted wastewater were used for DNA extraction. Extracted DNA was used for PCR and qPCR applications.

Samples	Dates															
	8/13 2007			3/19 2008			4/27 2008			6/15 2009	10/14 2009			12/22 2009		
	A1Q1 East	A1Q1 Middle	A1Q3 East	A1Q1 East	A1Q1 Middle	A1Q3 East	A1Q1 East	A1Q1 Middle	A1Q3 East	A1Q1 East	A1Q1 East	A1Q1 Middle	A1Q3 East	A1Q1 East	A1Q1 Middle	A1Q3 East
Raw Wastewater (g)	49.41	50.10	50.00	50.42	49.86	51.68	52.30	52.08	51.56	26.09	50.01	49.80	50.05	50.68	50.71	52.23
Wastewater Pellet (g)	9.36	8.03	7.41	5.86	6.02	6.58	5.25	6.52	9.68	2.08	6.59	5.32	3.88	5.53	4.09	5.60

Table 4. Raw wastewater and wastewater pellet weights for B1Q1 East, B1Q1 Middle, and B1Q3 East. Raw wastewater was centrifuged at 9,000 g for 15 min to pellet large particulates. Aliquots from pelleted wastewater were used for DNA extraction. Extracted DNA was used for PCR and qPCR applications.

<b>Samples</b>	<b>Dates</b>											
	<b>5/14/2009</b>			<b>6/15/2009</b>			<b>9/12/2009</b>			<b>1/28/2010</b>		
	<b>B1Q1 East</b>	<b>B1Q1 Middle</b>	<b>B1Q3 East</b>	<b>B1Q1 East</b>	<b>B1Q1 Middle</b>	<b>B1Q3 East</b>	<b>B1Q1 East</b>	<b>B1Q1 Middle</b>	<b>B1Q3 East</b>	<b>B1Q1 East</b>	<b>B1Q1 Middle</b>	<b>B1Q3 East</b>
Raw Wastewater (g)	49.46	50.15	49.98	25.94	26.02	25.84	51.60	52.14	51.66	50.53	49.80	49.82
Wastewater Pellet (g)	9.40	8.05	7.40	2.62	3.73	4.29	7.16	5.56	4.49	6.85	4.93	6.88

Addition of PBS with sonication for 3 and 5 min gave higher cell counts (Table 5) and separation (Figure 3) when compared to the no treatment control (Figure 5). However, all samples with the addition of Tween 80 had a greater amount of filamentous bacteria present in the liquid portion (Figure 4), as compared to the no treatment control (Figure 5) and samples with the addition of PBS only (Figure 3). Filaments seen in Figure 4 had similar morphology to previously published TM7 filaments from activated wastewater (Hugenholtz *et al.*, 2001). Solid floc particulates from Tween 80 samples (Figure 4) were smaller than those in the no treatment control (Figure 5) and in PBS only samples (Figure 3). Although 5 min sonication with Tween 80 gave the highest cell count in the liquid portion (Table 5), it was noticed that more and smaller flocs particles were in the liquid phase sample (Figure 4). These small floc particles would cover a larger area of well space, limiting the ability to see individual cells. Addition of 0.1% Tween 80 PBS solution with sonication for 3 min at 60 hertz gave the necessary combination of removing bacterial adhesion from flocs and separating large particulates from the liquid phase. Therefore a 3 min sonication treatment on the liquid phase, with the addition of PBS and Tween 80, was performed for further wastewater processing for microscopy.

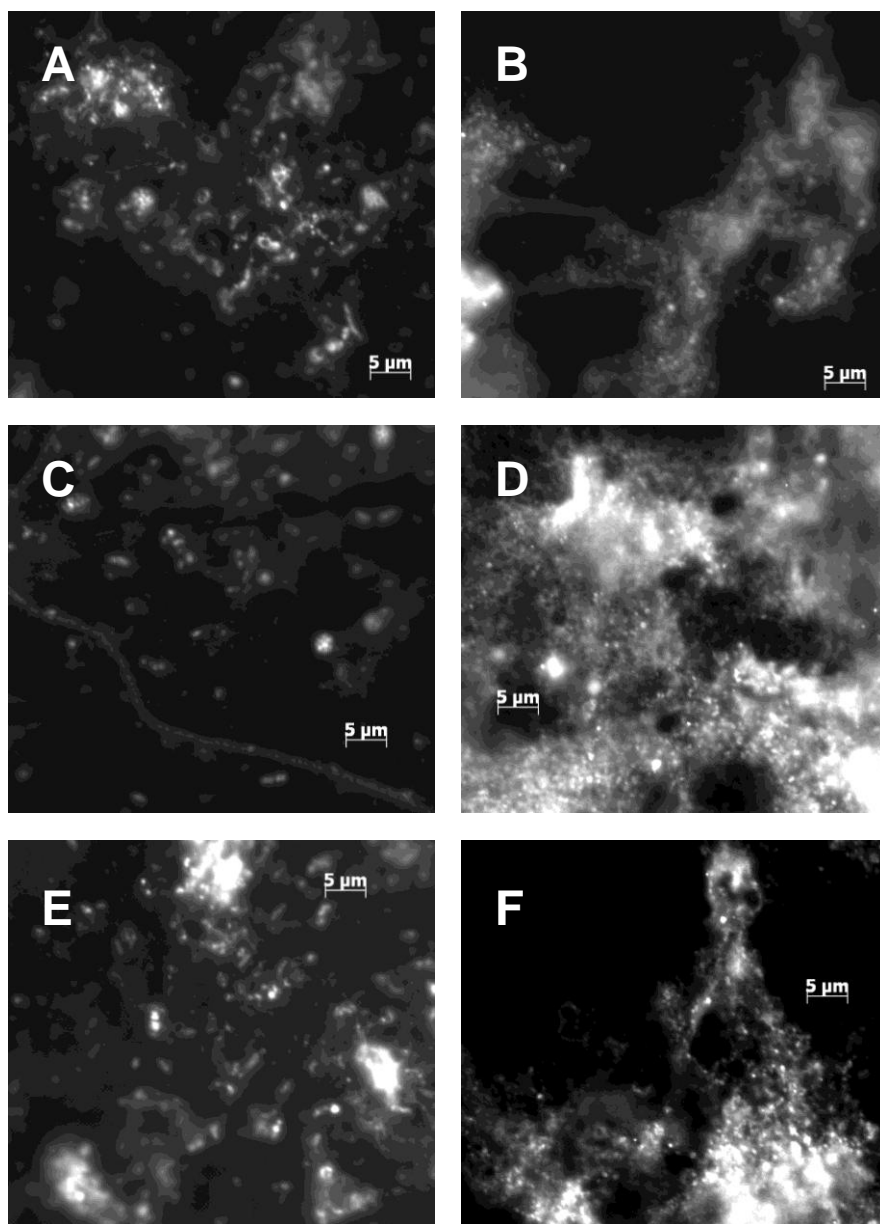


Figure 3. Sonication results on wastewater samples with the addition of PBS only.

(A) Supernatant from 2 min sonication. (B) Solid phase from 2 min sonication.

(C) Supernatant from 3 min sonication. (D) Solid phase from 3 min sonication.

(E) Supernatant from 5 min sonication. (F) Solid phase from 5 min sonication.

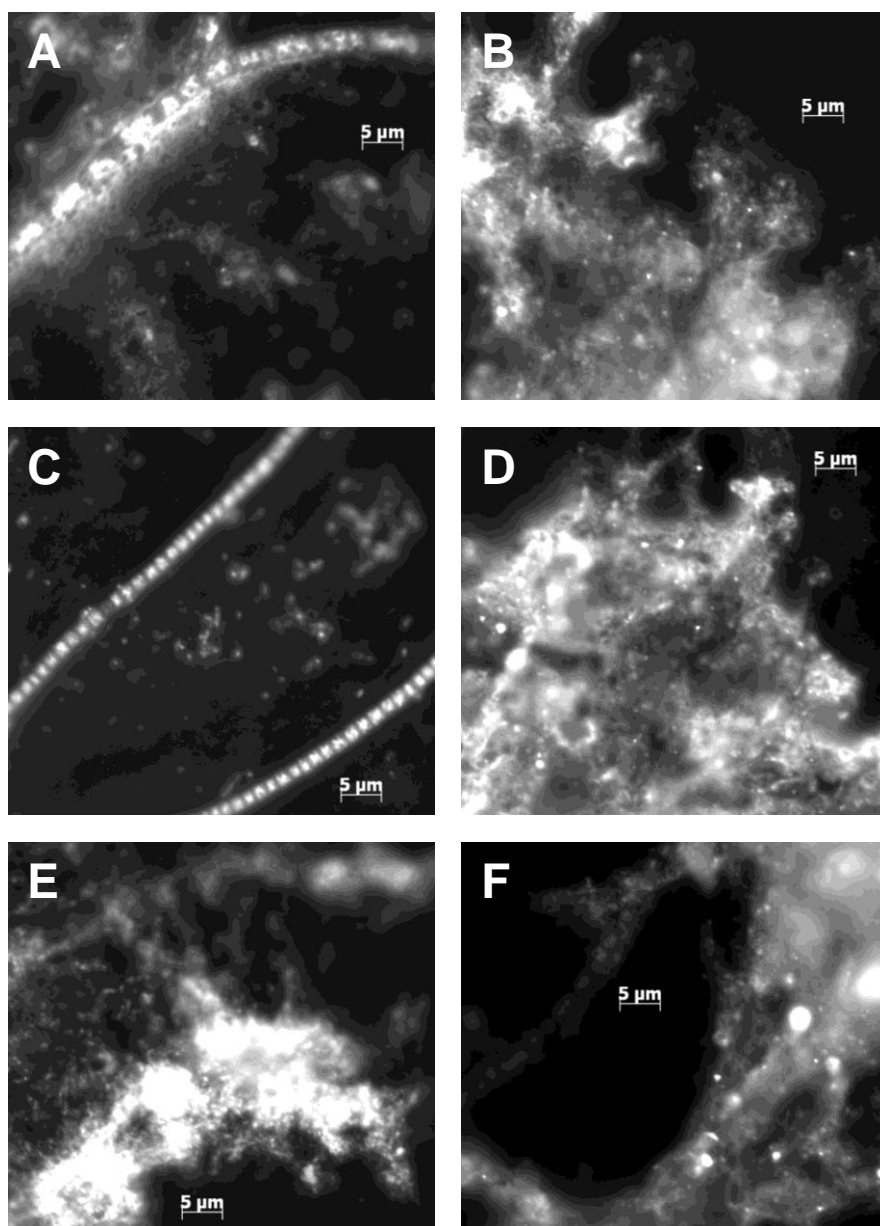


Figure 4. Sonication results on wastewater samples with the addition of PBS and Tween 80 emulsifier. (A) Supernatant from 2 min sonication. (B) Solid phase from 2 min sonication. (C) Supernatant from 3 min sonication. (D) Solid phase from 3 min sonication. (E) Supernatant from 5 min sonication. (F) Solid phase from 5 min sonication.

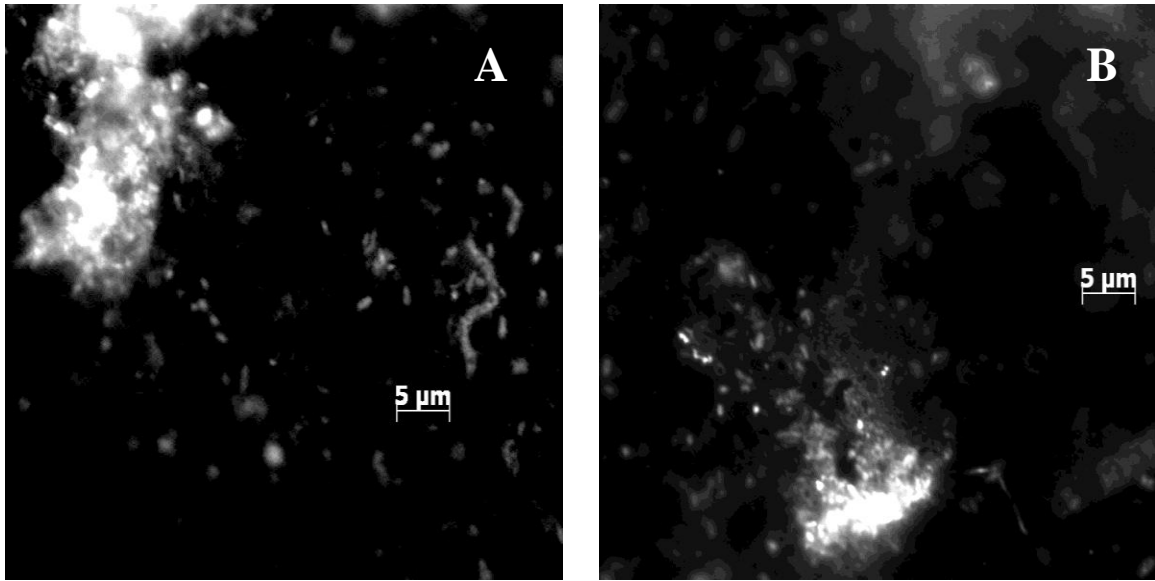


Figure 5. Raw wastewater from secondary treatment tank. (A) Supernatant after large particles were allowed to settle. (B) Solid phase after large particles were allowed to settle.

### **DNA Extraction**

Initial DNA extractions of activated wastewater showed that a modified bead-beater half-lysis protocol gave a consistently higher yield as seen when comparing Table 6, Table 7, and Table 8. Modified bead-beater half-lysis protocol yielded DNA concentrations in excess of 700 ng/ $\mu$ l. PowerSoil extracted DNA concentrations were less than 100 ng/ $\mu$ l per sample. In terms of DNA/protein purity, as determined by A260/A280 ratio, on average, PowerSoil extracted DNA consistently was closer to the ideal 1.80 ratio ( $1.81 \pm 0.17$ ) compared to modified bead-beater half-lysis extracted DNA ( $1.67 \pm 0.26$ ).



Table 5. Average cell count per ml of raw activated wastewater. Liquid numbers based on average from 10 separate counts from  $2.03 \times 10^3 \mu\text{m}^2$  area on slide. Solid numbers based on average from 10 separate counts from  $225 \mu\text{m}^2$  area on slide.

<b>Sample</b>	<b>Ave. Cells/ml</b>
control liquid	$1.00 \times 10^8$
control solid	$2.42 \times 10^8$
PBS only 2min liquid	$8.46 \times 10^7$
PBS only 2min solid	$3.15 \times 10^8$
PBS only 3min liquid	$1.08 \times 10^8$
PBS only 3min solid	$3.49 \times 10^8$
PBS only 5min liquid	$1.24 \times 10^8$
PBS only 5min solid	$2.27 \times 10^8$
PBS + Tween 80 2min liquid	$7.48 \times 10^7$
PBS + Tween 80 2min solid	$1.59 \times 10^8$
PBS + Tween 80 3min liquid	$1.39 \times 10^8$
PBS + Tween 80 3min solid	$1.78 \times 10^8$
PBS + Tween 80 5min liquid	$1.34 \times 10^8$
PBS + Tween 80 5min solid	$1.31 \times 10^8$

In terms of DNA molecule size as seen in Figure 6, the PowerSoil kit (bottom two wells) extracted large DNA molecules (>2,000 bp). The modified bead-beater half-lysis protocol (Figure 6, second and third wells from top) extracted a large number of small fragments (about 300 bp). However, there are larger DNA molecules visible (>2,000 bp) in the wells. Although the PowerSoil DNA extraction kit yielded purer DNA, it had a short shelf-life of about 6 months. After comparing the two methods, the modified bead-

beater half-lysis protocol was used for all DNA extractions performed after 08/13/2007. Modified bead-beater half-lysis extracted DNA was used for cloning as well as total bacterial, total TM7, and environmental SBG3 qPCR temporal studies.

Table 6. DNA extracted from “A” aerator tank group using modified bead-beater half-lysis protocol and Mo Bio PowerSoil kit. DNA amount expressed as ng/μl as determined with a Nanodrop spectrophotometer. DNA purity expressed as A260/A280 ratio. Pure DNA with no protein contamination will have an ideal A260/A280 ratio of 1.80. PowerSoil extracted DNA had better A260/A280 ratios than modified bead-beater half-lysis protocol extracted DNA, suggesting less protein contamination.

Extracted DNA	Dates				
	4/23/2007		8/13/2007		
	A1Q1 East	A1Q1 Middle	A1Q1 East	A1Q1 Middle	A1Q3 East
Modified Bead- Beater Half-Lysis ng/ul (A260/A280 ratio)	1366.08 (1.37)	956.28 (1.38)	1886.00 (1.85)	2311.00 (1.88)	1235.00 (1.86)
Mo Bio PowerSoil ng/ul (A260/A280 ratio)	10.13 (1.58)	20.82 (1.98)	37.20 (1.81)	26.90 (1.89)	

Table 7. DNA extracted from “A” aerator tank group samples from 3/19/2008 to 12/22/2009 using a modified bead-beater half-lysis protocol. DNA amount expressed as ng/μl as determined with a Nanodrop spectrophotometer. DNA purity expressed as A260/A280 ratio. Pure DNA with no protein contamination will have an ideal A260/A280 ratio of 1.80. Extracted DNA used for total bacterial, total TM7, and environmental SBG3 qPCR temporal studies.

Extracted DNA	Dates												
	3/19/2008			4/27/2008			6/15/2009	10/14/2009			12/22/2009		
	A1Q1 East	A1Q1 Middle	A1Q3 East	A1Q1 East	A1Q1 Middle	A1Q3 East	A1Q1 East	A1Q1 East	A1Q1 Middle	A1Q3 East	A1Q1 East	A1Q1 Middle	A1Q3 East
Modified Bead-Beater Half-Lysis ng/ul (A260/A280 ratio)	793.20 (2.61)	1282.21 (2.62)	1023.54 (3.41)	1437.36 (1.91)	2332.10 (1.86)	1796.63 (1.91)	1566.97 (1.76)	1632.34 (1.89)	2090.92 (1.89)	1815.38 (1.90)	1866.00 (1.21)	504.22 (1.33)	2603.00 (1.33)

Table 8. DNA extracted from “B” aerator tank group samples using a modified bead-beater half-lysis protocol. DNA amount expressed as ng/μl as determined with a Nanodrop spectrophotometer. DNA purity expressed as A260/A280 ratio. Pure DNA with no protein contamination will have an ideal A260/A280 ratio of 1.80. Extracted DNA used for total bacterial, total TM7, and environmental SBG3 qPCR temporal studies.

54

Extracted DNA	Dates											
	5/14/2009			6/15/2009			9/12/2009			1/28/2010		
	B1Q1 East	B1Q1 Middle	B1Q3 East	B1Q1 East	B1Q1 Middle	B1Q3 East	B1Q1 East	B1Q1 Middle	B1Q3 East	B1Q1 East	B1Q1 Middle	B1Q3 East
Modified Bead-Beater Half-Lysis ng/ul (A260/280 ratio)	1366.08 (1.37)	956.28 (1.38)	935.45 (1.37)	1079.00 (1.65)	780.47 (1.77)	1018.26 (1.94)	923.86 (1.95)	1031.60 (1.96)	923.86 (1.95)	1070.90 (1.77)	1126.30 (1.76)	1102.70 (1.69)

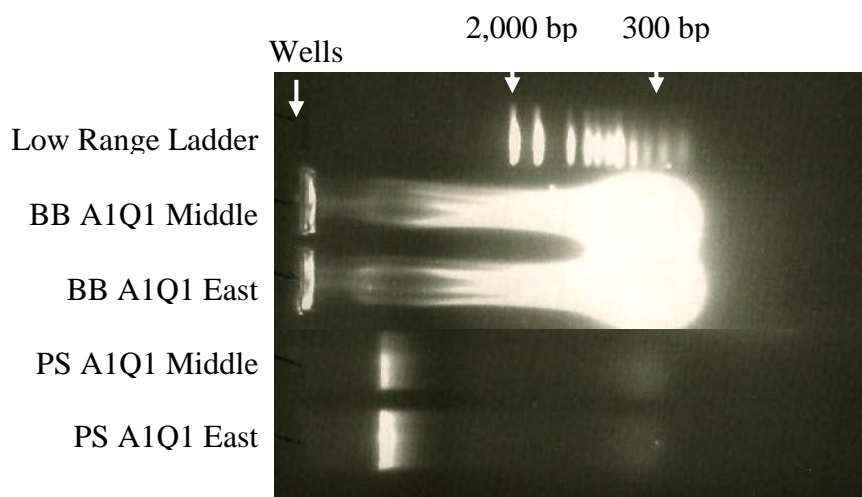


Figure 6. Comparison between modified bead-beater half-lysis extracted nucleic acid (BB) and Mo Bio PowerSoil extracted nucleic acid (PS). Each well was loaded with 8 $\mu$ l of undiluted extracted DNA. Modified bead-beater half-lysis samples showed higher concentration of DNA than Mo Bio PowerSoil extracted DNA per  $\mu$ l; although, modified bead-beater half-lysis protocol produced fragmentation of DNA. Image was cropped to exclude samples not relevant to this thesis.

### **TM7 16S rRNA Gene Amplification Results**

Initial PCR screening showed that there was PCR inhibition in modified bead-beater half-lysis extracted DNA (Figure 7). Dilution of the extracted nucleic acid in sterile water removed PCR inhibition. A 1/100 dilution of extracted nucleic acid consistently diminished inhibition such that PCR products were detected (Figure 7). All three sample sites from “A” aerator tank group and all three sample sites from “B” aerator tank group were found to be presumptive positive for TM7 based on PCR

amplification of the 16S rRNA gene, when using BAC 8F and TM7 1177R. Additional screening showed that at least one sample from all “A” aerator tank sites yielded PCR products with SBG3 group primers. Extracted DNA from “B” aerator tank sites did not produce PCR products with SBG3 group primers (Table 9).

### **Cloning TM7 PCR Products**

After samples were determined to be presumptive positive for TM7, PCR for cloning was performed on “A” aerator tank group sites. All three sample sites had PCR products that corresponded to the correct base pair size when compared to the positive control in a 1% agarose gel. Band extracted DNA concentrations ranged from about 10 ng/ul to 30 ng/ul. Based on PCR clone screening with M13 Forward and M13 Reverse primers, 12 clones from A1Q1 East, 18 clones from A1Q1 Middle, and 20 clones from A1Q3 East had PCR products of the correct size (approximately 1,170 bp).

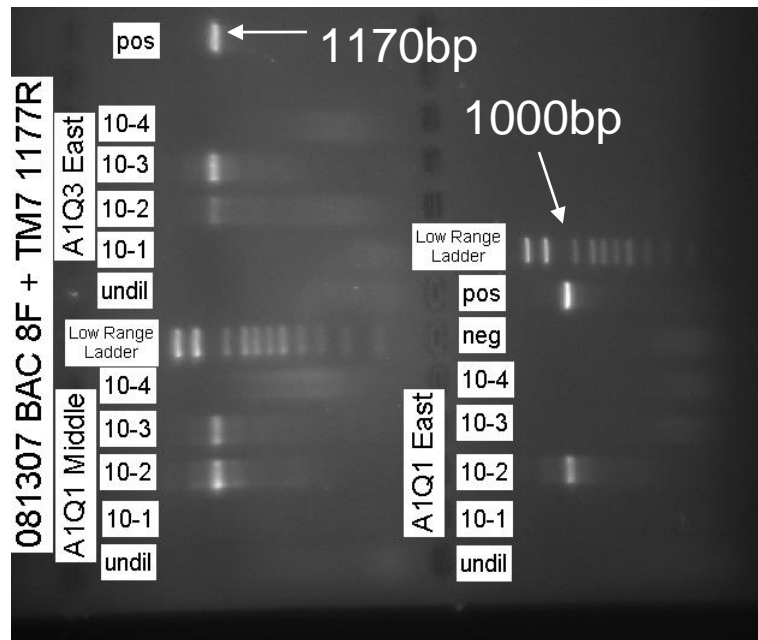


Figure 7. PCR results from BAC 8F and TM7 1177R screening of nucleic acid extracted with modified bead-beater protocol. Undiluted and diluted samples from A1Q1 East, A1Q1 Middle, and A1Q3 East extracted on 08/13/2007 were used. Undiluted samples were inhibited while 1/100 dilutions produced PCR products.

Table 9. Combined “A” and “B” aerator tank group results for DNA extraction, PCR amplification, and FISH. Samples were positive for TM7 and SBG3 based on PCR amplification of the 16S rRNA gene. TM7 positives included both “A” and “B” aerator tank group samples. SBG3 positives only came from “A” aerator tank samples. FISH was performed on both “A” (12) and “B” (5) aerator tank samples.

Number of Samples DNA Extracted	Number of Samples Positive For TM7	Number of Samples Positive For SBG3	Number of Samples Analyzed by FISH
31	27	7	17

## **Phylogenetic Analysis**

BLAST results showed that sequences of seven clones from A1Q1 East, eight clones from A1Q1 Middle, and nine clones from A1Q3 East (24 clones total) had TM7 sequences as their top hits with 80% or greater coverage when compared to previously published sequences in the NCBI database. Sequences from A1Q1 East and A1Q3 East clones matched to environmental TM7 with 86% or greater coverage and 91% identity based on BLAST reports. Sequences matched with either candidate division TM7 enrichment culture clone hao21 (accession number FJ386548) or uncultivated candidate division TM7 bacterium clone Cystalbog 2KD8 (accession #AY792306). Sequences from the eight clones from A1Q1 Middle matched to 16S rRNA sequences from human oral sub-gingival crevice derived uncultivated candidate division TM7 bacterium clone SBG 3 (accession number AY144355) with 99% coverage and 98% similarity based on BLAST reports. All other clones matched to uncultivated, non-TM7 clone sequences.

ARB phylogenetic tree generation was performed on eight representatives, six environmental SBG3 clones and two environmental TM7 clones that represent the two top hit environmental TM7 BLAST results (Figure 8). The two environmental TM7 clones grouped together and showed similarity to other previously published environmental TM7 clones, but did not group closely with previously published human-associated TM7 clones. The six environmental SBG3 clones grouped very close together and are represented on this tree by clone HM208132. Environmental SBG3 clone HM208132 showed 98.60% similarity with human oral clone SBG3 (Table 10) and showed similarity to other human-associated TM7 clones (Figure 8 and Table 10).



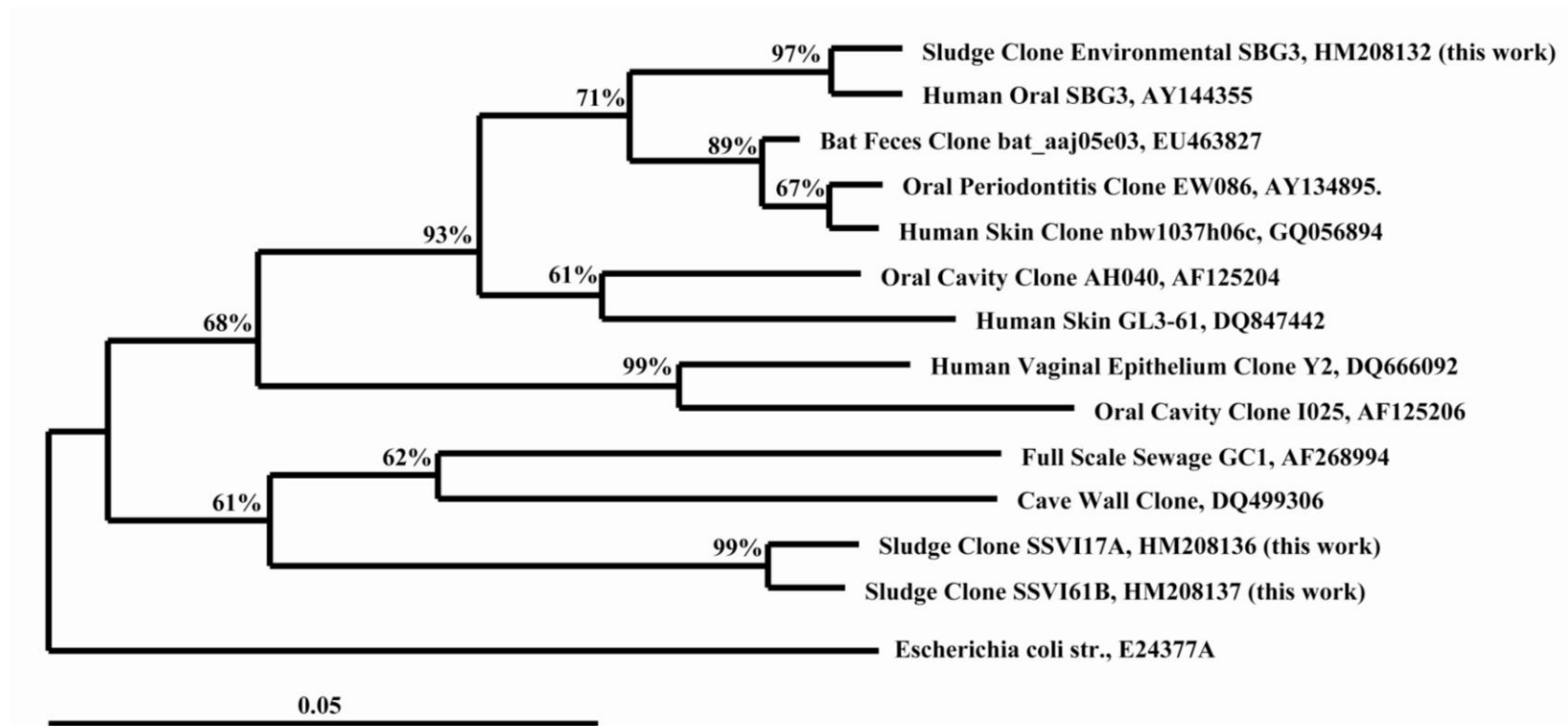


Figure 8. Evolutionary distance of human-associated and environmental TM7 phylotypes based on 16S rRNA gene phylogenetic analysis. Sludge clone environmental SBG3, HM208132 (top node) grouped to the published human oral TM7, SBG3 AY144355. Tree created in ARB software with no evolutionary corrections, bootstrap of 1,000 replicates (bootstrap values  $\geq 50\%$  shown) and *E. coli* E24377A as the outgroup.

Table 10. Similarity of human-associated and environmental TM7 phylotypes to environmental SBG3 HM208132 (this work) based on 16S rRNA gene sequence phylogenetic analysis. Sludge clone environmental SBG3 showed 98.60% similarity to the published human oral TM7, SBG3 AY144355.

	<b>% Similarity To Environmental SBG3 (this work)</b>
<b>Human Oral SBG3, AY144355</b>	<b>98.60</b>
Bat Feces Clone bat_aaj05e03, EU463827	95.55
Human Skin Clone nbw1037h06c, GQ056894	95.36
Oral Periodontitis Clone EW086, AY134895	94.81
Oral Cavity Clone AH040, AF125204	91.72
Human Skin Clone GL3-61, DQ847442	90.87
Human Vaginal Epithelium Clone Y2, DQ666092	87.51
Sludge Clone SSVI17A, HM208136 ( <b>this work</b> )	85.93
Oral Cavity Clone IO25, AF125206	85.92
Sludge Clone SSVI61B, HM208137 ( <b>this work</b> )	86.23
Cave Wall Clone, DQ499306	84.67
Full Scale Sewage GCI, AF268994	84.23
<i>E. coli</i> str., E24377A	76.12

### **Quantitative PCR**

Real time quantitative PCR was used to detect and quantify the 16S rRNA gene copy numbers. Separate qPCR assays were optimized and applied to quantify the total bacterial, total TM7, and TM7 subgroup environmental SBG3 16S rRNA gene copy numbers. Threshold was set to obtain a standard curve as close to the ideal -3.33 slope as

possible. This normalized data accumulated over the three runs and gave amplification efficiencies nearing the ideal value of one (Table 11). The  $R^2$  values for all runs were greater than 99% (100% being ideal), except for two general bacterial runs which had an  $R^2$  of greater than 98% but less than 99% (Table 11). This shows that the standard points fit closely to the generated standard curve. Although, all runs had a serial dilution series that ranged from  $10^8$  to 1 positive control copy number per  $\mu\text{l}$ , the 7300 software produced the best line with five sequential dilution points. The five point standard curve ranges used for total bacterial, total TM7, and TM7 subgroup environmental SBG3 quantification runs are shown in Table 11. From “A” and “B” sample sites, 87% of samples had detectable levels of TM7 16SrRNA genes (Table 9). Only seven samples from “A” sample sites contained detectable levels of SBG3 16S rRNA genes when using SBG3 specific primers (Table 2) as seen in Table 9. No samples from “B” sample sites had detectable levels of SBG3 16S rRNA copy numbers when using SBG3 specific primers which was consistent with PCR of these samples.

The highest TM7 count was recorded in August 2007 at  $5.43 \times 10^7$  cells/g and the lowest in December 2009 at  $2.5 \times 10^5$  cells/g. The highest total bacterial count was recorded in August 2007 at  $2.47 \times 10^9$  cells/g and the lowest in March 2008 at  $7.12 \times 10^6$  cells/g. These values were used to provide a temporal look (Figure 9, 10, 11) at how the total bacterial and total TM7 populations fluctuated over time. The percentage of TM7 bacteria versus the total bacterial population size was calculated (Table 12).

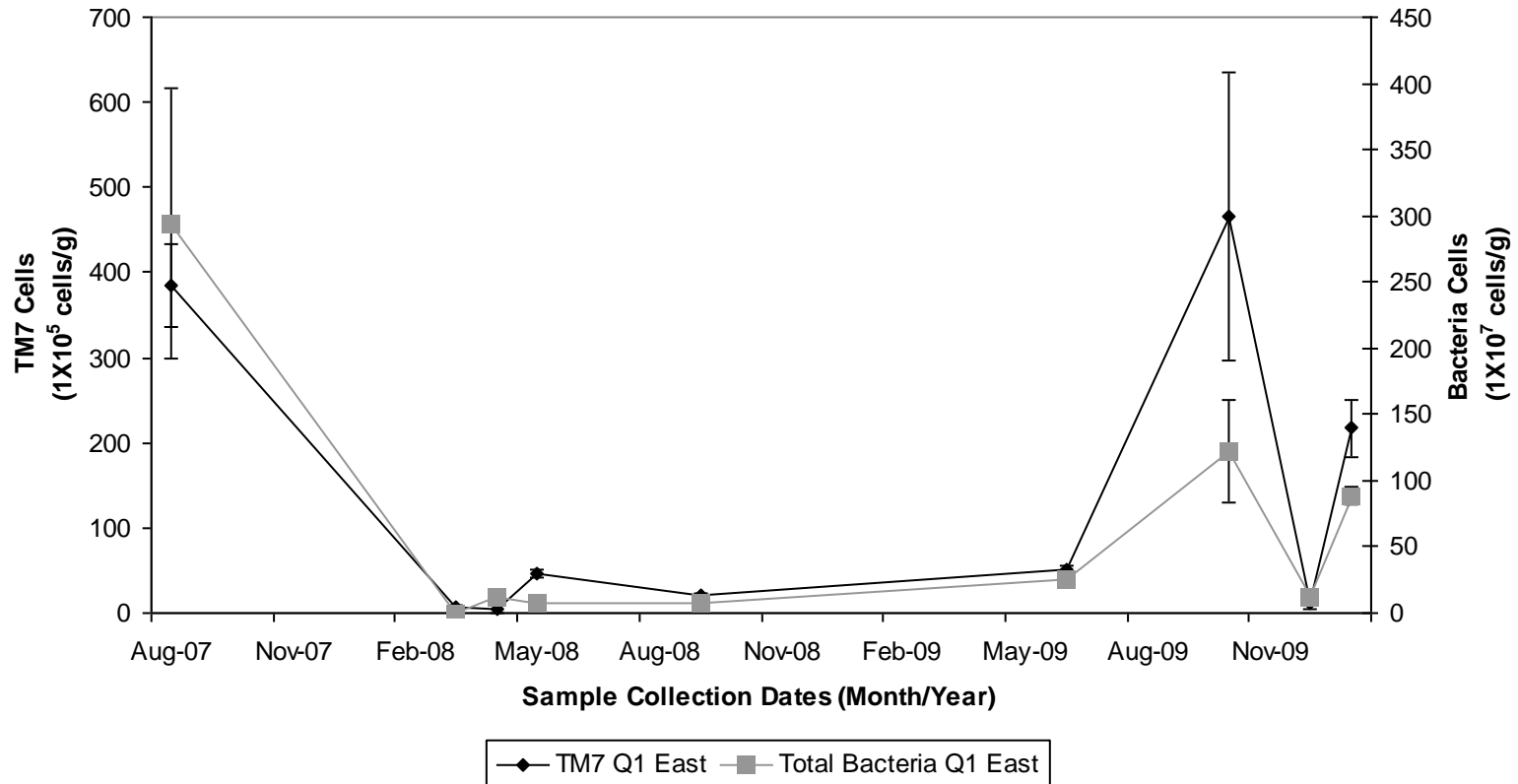
Table 11. Experimental values included in qPCR analysis from temporal study of total bacterial, total TM7, and TM7 subgroup, environmental SBG populations in activated wastewater. All assays consisted of three runs and each run of three triplicates per sample, total sample population (n=9).

<b>qPCR Assay</b>	<b>Amplification Parameters</b>			<b>Standard Range</b>
	<b>Amplification Efficiency (Standard Deviation)</b>	<b>Slope (Standard Deviation)</b>	<b>R<sup>2</sup></b>	
Total Bacterial	1.02 (±0.025)	-3.29 (±0.056)	>98%	10 <sup>4</sup> -10 <sup>8</sup>
Total TM7	0.96 (±0.019)	-3.44 (±0.051)	>99%	10 <sup>3</sup> -10 <sup>7</sup>
Environmental SBG3	0.92 (±0.020)	-3.53 (±0.056)	>99%	10 <sup>1</sup> -10 <sup>5</sup>

Table 12. Average TM7 percentage of the total bacterial population at the three sample sites. “A” and “B” tank groups were combined.

	<b>Q1 East</b>	<b>Q1 Middle</b>	<b>Q3 East</b>
<b>Percent TM7 of Total Bacterial Population</b>	<b>3.44%</b>	<b>3.75%</b>	<b>2.72%</b>

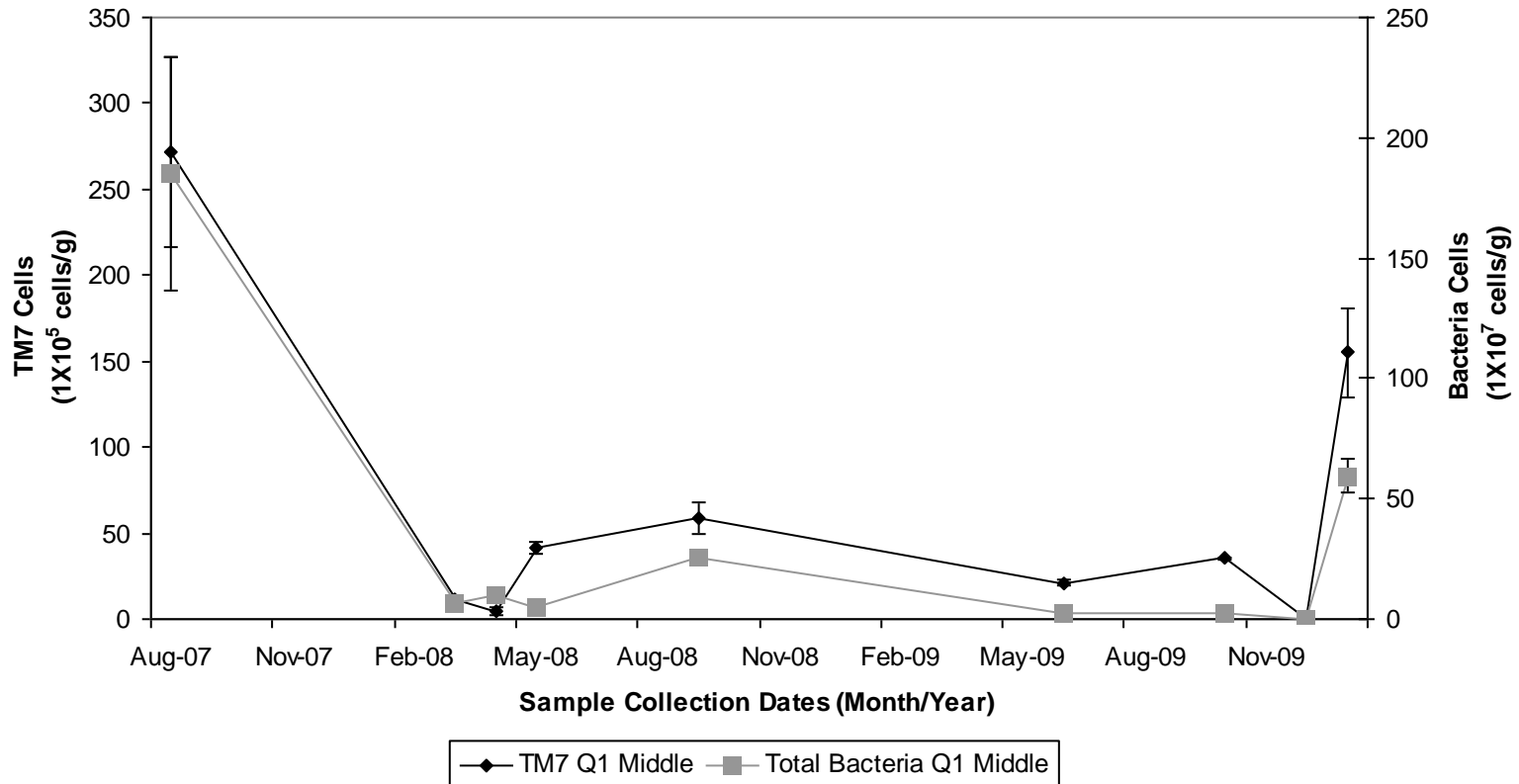
### Temporal Quantification of Total Bacteria and TM7 in Activated Wastewater Sample Site Q1 East



69

Figure 9. Temporal study of Q1 East total bacterial population and TM7 population from combined “A” and “B” aerator tank groups.

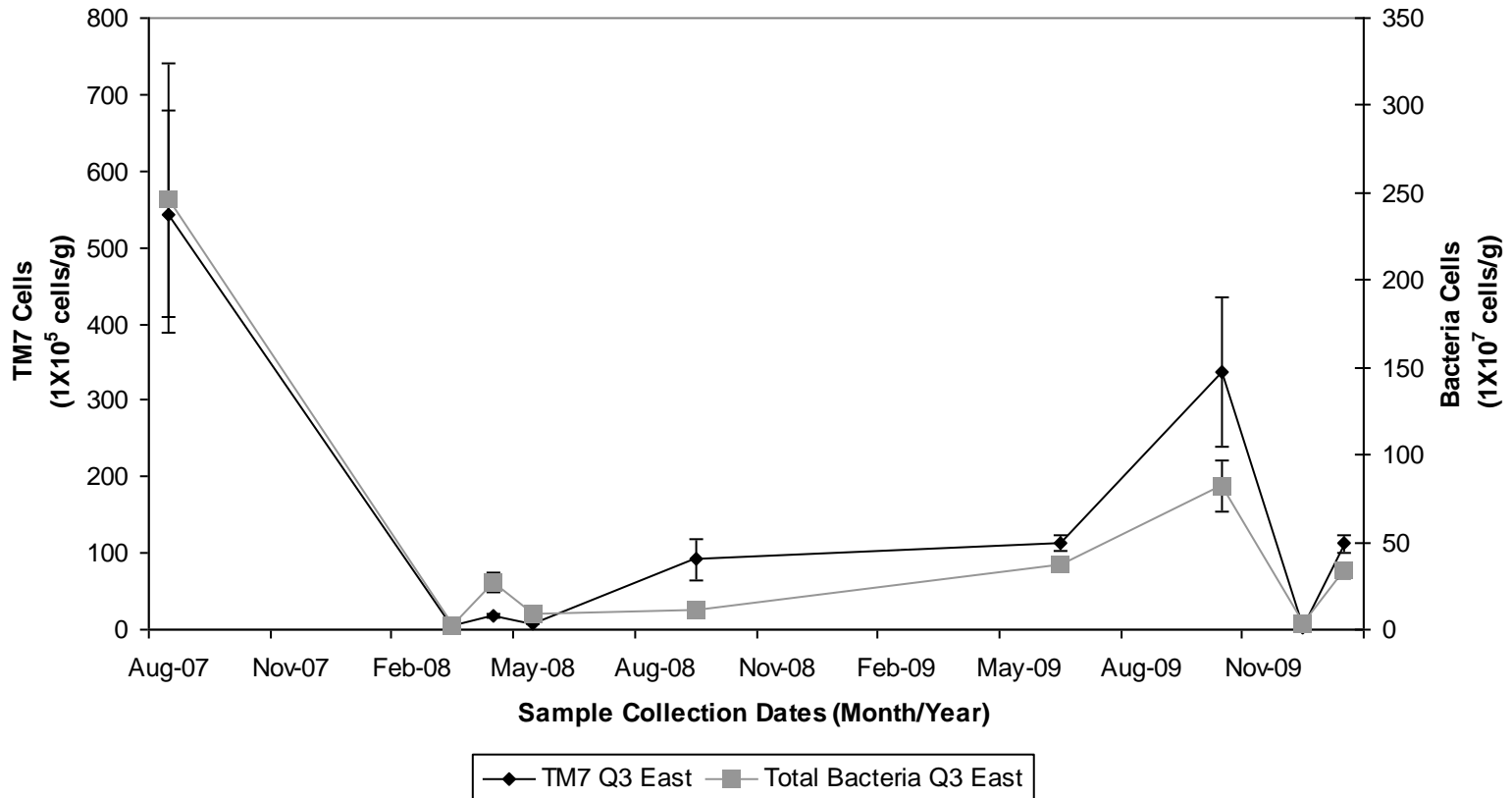
### Temporal Quantification of Total Bacteria and TM7 in Activated Wastewater Sample Site Q1 Middle



64

Figure 10. Temporal study of Q1 Middle total bacterial population and TM7 population from combined “A” and “B” aerator tank groups.

### Temporal Quantification of Total Bacteria and TM7 in Activated Wastewater Sample Site Q3 East



69

Figure 11. Temporal study of Q3 East total bacterial population and TM7 population from combined “A” and “B” aerator tank groups.

TM7 bacteria were present at all sample collection dates throughout the year (Figure 9, 10, 11) suggesting that TM7 is a constant resident of the activated wastewater bacterial community at sampled sites. Factors that affected bacterial populations could not be determined as many changes were performed in wastewater processing during the study period by the San Jose/Santa Clara Water Treatment Facility. However, fall and early winter collection points had, on average, higher total bacterial counts and total TM7 counts as compared to spring and summer collection points. Changes in total TM7 bacterial counts were similar to changes in total bacterial counts, suggesting that total TM7 percentage of the total mixed bacterial community remains relatively constant at 2.75% to 3.75% (Table 13).

Environmental SBG3 were not present at all sample collection dates throughout the year (Figure 12). Additionally, environmental SBG3 levels varied between the three “A” sample collection sites depending on the time of year (Figure 12). The highest environmental SBG3 level identified was in August 2007 at  $8.39 \times 10^4$  cells/g (Table 13). Environmental SBG3 percentage of the total TM7 population ranged from  $3.31 \times 10^{-2}\%$  to  $7.64 \times 10^{-1}\%$  (Table 13). Environmental SBG3 percentage of the total bacterial population ranged from  $1.00 \times 10^{-3}\%$  to  $9.21 \times 10^{-3}\%$  with a spike of  $1.03 \times 10^{-1}\%$  seen in March 2008 at sample site A1Q1 East (Table 13). Factors that caused the variation in environmental SBG3 levels could not be determined.



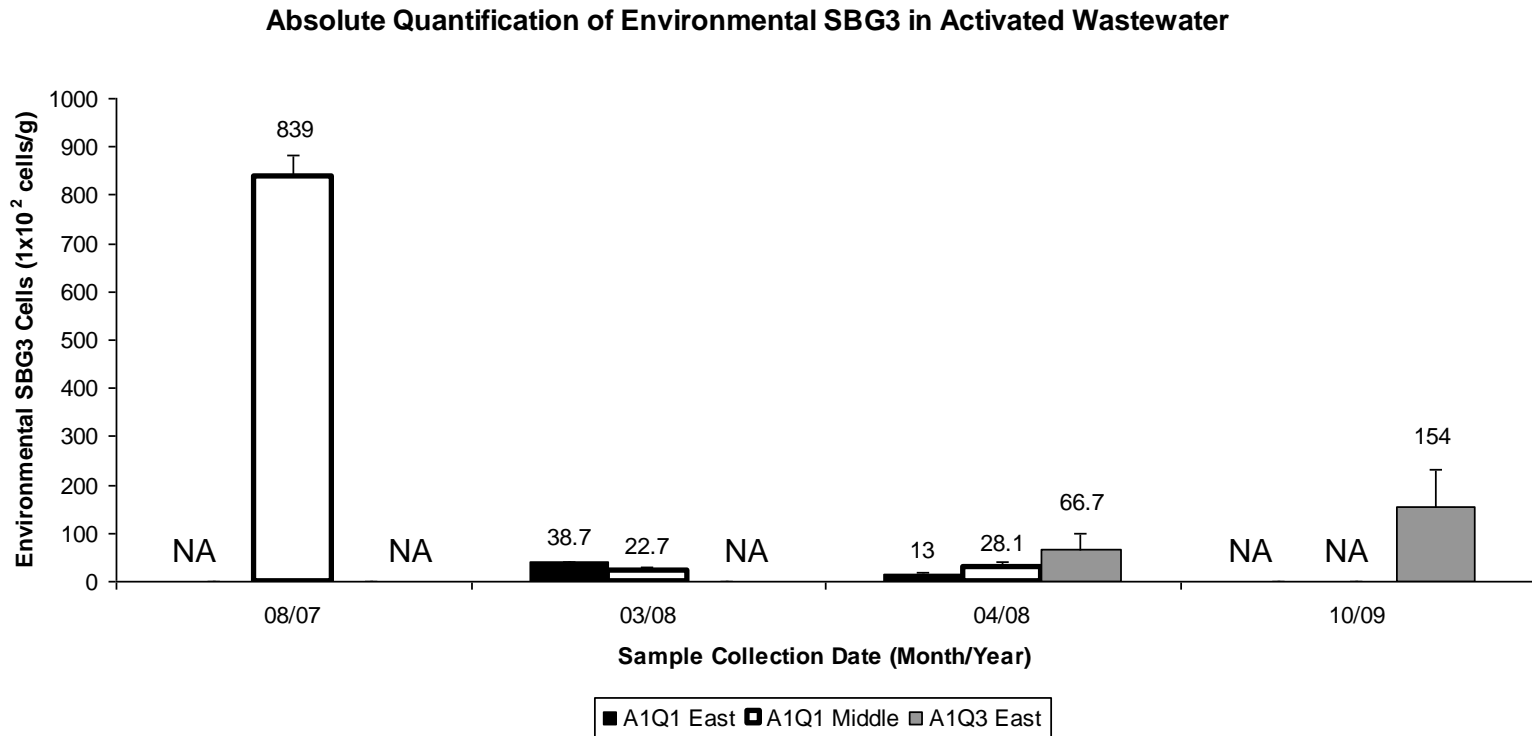


Figure 12. Average number of environmental SBG3 in activated wastewater from the three “A” aerator tank group sample sites from two year period. “NA” equals sample site did not have a quantifiable level of SBG3 DNA through ABI 7300 qPCR machine with Taqman assay.

Table 13. Quantification of environmental SBG3 and environmental SBG3 population percent of the total TM7 and total bacterial populations for samples that were positive for environmental SBG3 16S rRNA gene presence. Detection and quantification performed via quantitative PCR (qPCR) and expressed as cells per gram activated wastewater. Environmental SBG3 percent of the total bacterial population determined by comparing environmental SBG3 qPCR results to total TM7 and total bacterial qPCR results.

	<b>Dates</b>						
	<b>A1Q1 Middle 8/13/2007</b>	<b>A1Q1 East 3/19/2008</b>	<b>A1Q1 Middle 3/19/2008</b>	<b>A1Q1 East 4/27/2008</b>	<b>A1Q1 Middle 4/27/2008</b>	<b>A1Q3 East 4/27/2008</b>	<b>A1Q3 East 10/14/09</b>
Environmental SBG3 Cells/g Activated Wastewater (Standard Deviation)	8.39x10 <sup>4</sup> (±0.43)	3.87x10 <sup>3</sup> (±0.23)	2.27x10 <sup>3</sup> (±0.72)	1.30x10 <sup>3</sup> (±0.39)	2.81x10 <sup>3</sup> (±1.17)	6.67x10 <sup>3</sup> (±3.13)	1.54x10 <sup>4</sup> (±0.77)
Percent Environmental SBG3 of Total TM7 Population	3.08x10 <sup>-1</sup> %	7.64x10 <sup>-1</sup> %	3.31x10 <sup>-1</sup> %	2.05x10 <sup>-1</sup> %	5.56x10 <sup>-1</sup> %	3.75x10 <sup>-1</sup> %	3.31x10 <sup>-2</sup> %
Percent Environmental SBG3 of Total Bacterial Population	4.54x10 <sup>-3</sup> %	1.03x10 <sup>-1</sup> %	9.21x10 <sup>-3</sup> %	1.00x10 <sup>-3</sup> %	2.95x10 <sup>-3</sup> %	3.40x10 <sup>-3</sup> %	2.55x10 <sup>-3</sup> %

### **Cloning of qPCR Products**

For the TM7 qPCR cloning reaction, 20 colonies were observed. Seven colonies gave an insert of the correct size when screened by PCR. The environmental SBG3 cloning reaction gave 17 colonies with PCR screening showing that one had an insert of the correct size. All other colonies had short sequences suggesting vector closed on primer dimers or that vector closed upon itself. Yield was low due to high fidelity *Taq* polymerase in the ABI master mix not being compatible with TA cloning. Sequences of the seven TM7 cloning reaction colonies had the same sequence to environmental TM7 clone HM208137 in the coverage area. The sequence from the one SBG3 cloning reaction colony was the same as environmental SBG3 clone HM208132 in the coverage area. Cloning of qPCR products strongly suggested that TM7 specific and environmental SBG3 specific qPCR primers target TM7 and environmental SBG3 members.

### **Fluorescent *in situ* Hybridization**

FISH was performed on samples from both “A” and “B” wastewater aerator tank groups to identify TM7 cell presence in activated wastewater and wastewater TM7 cell morphology. The use of general DNA stain was used to identify objects as cells versus autofluorescence or non-cellular material. The visualization of general DNA stain confirms that objects labeled with probes were cells.

As seen in Figure 13, TM7-905 probes bound to filaments and bacilli cells. Filaments were differentiated into long, thick filaments, and short, thinner filaments. Long filaments grew up to 60  $\mu\text{m}$  long and 2.5  $\mu\text{m}$  wide with individual cells 2.5  $\mu\text{m}$  in

length. Some long filaments also appeared to have a sheath surrounding them. Short filaments grew 4-15  $\mu\text{m}$  long and 1.5  $\mu\text{m}$  wide with cells 3-5  $\mu\text{m}$  in length. Large bacilli were observed that were 2-4  $\mu\text{m}$  in length and 1.5  $\mu\text{m}$  wide. Small bacilli were 1-2  $\mu\text{m}$  long and 0.5  $\mu\text{m}$  wide. As seen in Figure 14, the environmental SBG3 subgroup probe bound to diplo-bacilli that were together 1.5-3.5  $\mu\text{m}$  in length and 1  $\mu\text{m}$  in width.

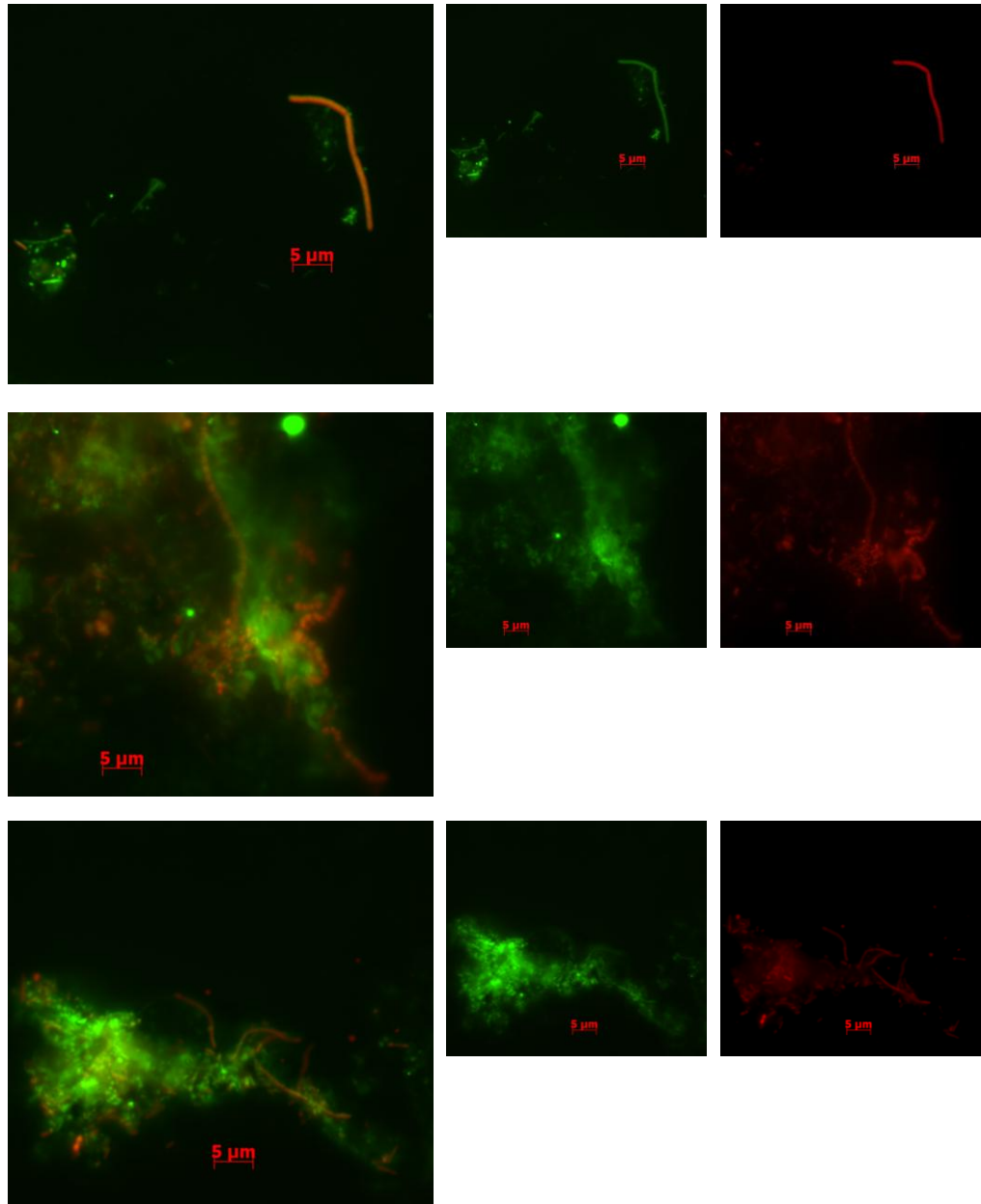


Figure 13. FISH results on wastewater samples using TM7 specific probe TM7 905 CY3 (red) and YO-PRO-1 general DNA stain (green). TM7 probe bound to long and short filaments as well as bacilli cells. Individual channels shown to compare TM7 probe labeled cells and cells stained with general DNA stain.

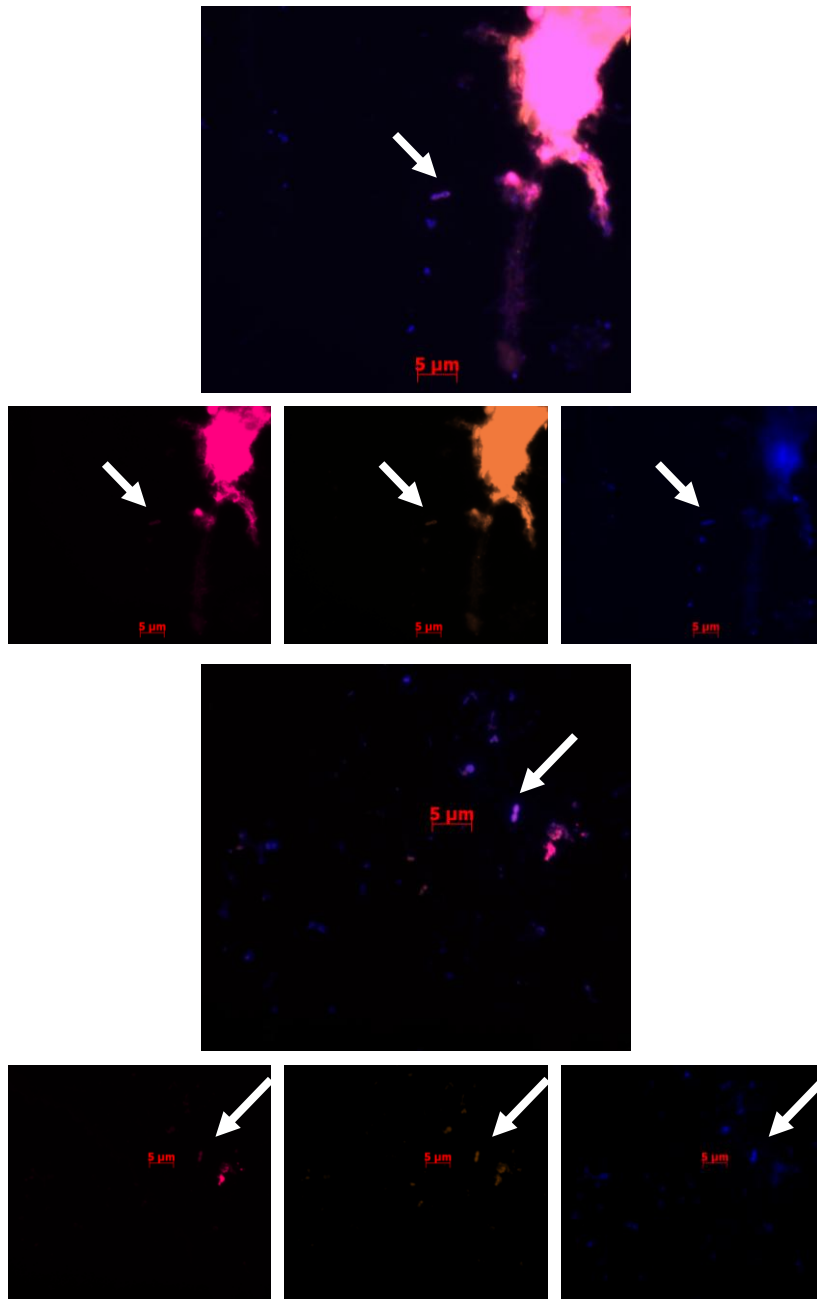


Figure 14. FISH results on wastewater samples using TM7 specific probe TM7 905 CY5 (pink), SBG3 Subgroup probe SBG3\_1 FISH CY3 (orange), and DAPI general DNA stain (blue). SBG3 and TM7 labeled cells were diplo-bacilli cell morphology. Individual channels shown to compare probe labeled cells and cells stained with general DNA stain.

## Discussion

### Identification of Environmental, Human-Like TM7 Bacteria

The currently accepted theory regarding the TM7 phylum is that it is clustered as distinct groups based on their niche habitats (Hugenholtz *et al.*, 2001). The two main TM7 clusters are grouped as environmental and human-associated TM7 groups. Previous studies were able to identify environmental TM7 bacteria from a variety of environmental sources or human-associated TM7 from a variety of human sites. It is generally accepted that if two organisms' 16S rRNA gene sequences are 98.5% similar or higher, that those two organisms belong to the same genus. None of the currently published environmental TM7 and human-associated TM7 shows high sequence similarities to each other. This study is first in identifying an environmental TM7 with such high similarity (98.60%) to a human-associated TM7. The use of TM7 16S rRNA gene amplification primers allowed for the specific targeting of the TM7 group without having to resort to amplifying the entire microbial community 16S rRNA genes. Of the 49 BAC 8F – TM7 1177R generated PCR products that were submitted for sequencing, 25 matched closely to previously published TM7 bacteria based on NCBI-BLAST reports. Among the 24 TM7 sequences, 15 were environmental TM7 that were split between two aquatic TM7 groups. The remaining nine sequences represent six phlotypes related to the SBG3 subgroup of TM7 which was originally identified in the human sub-gingival crevice. ARB phylogenetic analysis confirmed that the six phlotypes present were unique but highly related to human oral TM7 subgroup SBG3. The identification of an

environmental SBG3 refutes the claim that the TM7 phylum is segregated into only human-associated TM7 and only environmental TM7.

In a separate study by Ghadiri *et al.* (2009), sequences from cloned environmental SBG3 PCR products differed from sequences obtained from screened samples provided by members in the laboratory. Additionally, environmental SBG3 sequences differed from human oral cavity SBG3 sequences published by Brinig *et al.* (2003) (Ghadiri *et al.*, 2009). This reduces the possibility that results were due to laboratory members contaminating activated wastewater samples with SBG3 DNA from their microbial communities.

Environmental SBG3 cells visualized by FISH had strong fluorescence signal from both TM7 and SBG3 probes. Fluorescent *in situ* hybridization targeted the 16S rRNA molecule found in the small ribosomal subunit. Whole cell visualization was possible as there were numerous target ribosomes present in the cell. Ribosomes are used to indicate the health of a cell (Wagner, 1994). Actively growing bacteria create and sustain more ribosomes than bacteria that are in stationary phase or dying (Wagner, 1994). Strong fluorescence from environmental SBG3 suggests that large numbers of ribosomes were present in environmental SBG3. The visualization of so many ribosomes in environmental SBG3 cells strongly suggests that environmental SBG3 were present and alive in activated wastewater. Environmental SBG3 cells visualized in activated wastewater had similar morphology to previously published oral TM7 morphologies.



## **Evaluation of Wastewater Sample Sites for Further Study**

A steady supply of environmental SBG3 is needed to better understand environmental SBG3 and identify qualities that could perpetuate TM7 pathogenesis in humans. A temporal study was performed using TM7 and SBG3 specific 16S rRNA gene primers and highly sensitive qPCR techniques to ensure a clear definition of TM7 and environmental SBG3 prevalence and abundance in activated wastewater samples.

The presence of TM7 DNA in most of the samples suggests that this phylum is a consistent member of sampled activated wastewater site communities. This is supported by the presence of TM7 probe labeled cells in all FISH samples. TM7 cells visualized by FISH conform to previously published morphologies present in activated wastewater. The presence of numerous TM7 morphologies suggests that activated wastewater is a good habitat for TM7.

Environmental SBG3 bacteria were identified in 7 out of 31 samples through qPCR 16S rRNA gene amplification. It is possible that environmental SBG3 bacteria were present in “negative” samples but were below the detection threshold for qPCR. This can be explained by the need to dilute extracted DNA in order to obtain PCR products. All DNA used for qPCR analysis were 1/100 dilutions of the original extracted DNA. The smallest amount of environmental SBG3 that was quantified was  $1.30 \times 10^3$  cells/g activated wastewater. A 1/100 dilution would represent 13 cells/g of activated wastewater which would be near the threshold for detection. Samples that contained less environmental SBG3 than  $1.30 \times 10^3$  cells/g activated wastewater would be below the

threshold for detection using qPCR after dilutions necessary for optimal PCR amplification to occur.

Environmental SBG3 were not localized to one sample location but were identified within the three “A” sample sites (Figure 12). However, environmental SBG3 were not detected at all sample collection dates. Collection dates were a month to several months apart. Samples were collected from tanks that were approximately 8 m wide by 45 m long by 15 m deep. The wastewater flow in the secondary aerator treatment process was from A1Q1 Middle to A1Q1 East to A1Q3 East (Figure 1). The entire treatment process would take about two weeks.

It is possible environmental SBG3 were not detected at all times because they moved through the secondary aerator water treatment process. Environmental SBG3 may have been localized in a micro-environment outside the sampled locations. It is possible environmental SBG3 need to be “seeded” in the aerator tanks by the constant influx of raw sewage. The need to be “seeded” could explain why environmental SBG3 were detected in sample site A1Q1 East which is against the wastewater flow compared to A1Q1 Middle (Figure 1). This can also explain the large gap between environmental SBG3 identification in March 2008 and October 2009. Between these two dates no environmental SBG3 were quantified in the three “A” sample locations.

TM7 were detected in higher percentages of the total bacterial population in activated wastewater when compared to TM7 quantified from human oral sites by Brinig *et al.* (2003). On average more TM7 were quantified in the human oral cavity, but there were more total bacteria in the human oral cavity as well (Brinig *et al.*, 2003). TM7

bacteria only make up 0.3% of the total bacterial population in the human mouth except during early stages of periodontitis where it is 0.5% (Brinig *et al.*, 2003). In comparison to human oral sites, TM7 abundance is 7.5 times higher (2.75-3.75% of the total bacterial community) in activated wastewater. Environmental SBG3 percentages in activated wastewater ( $1.00 \times 10^{-3}\%$  to  $9.21 \times 10^{-3}\%$ ) are comparable to the total TM7 percentage in the human mouth. In the human sub-gingival crevice, SBG3 is one of only seven sub-gingival TM7 groups. Combined with the knowledge that the total TM7 population is only 0.3-0.5% of the total bacterial population in the human oral cavity, environmental SBG3 is a larger percentage of the total TM7 population in activated wastewater compared to the human oral cavity.

### **Possible Further Studies**

Characterization of environmental SBG3 to identify any features that would aid TM7 in human pathogenicity would follow this study. Substrate tracking autoradiography fluorescent *in situ* hybridization (STAR-FISH) and whole genome amplification are possible approaches to distinguish possible pathogenic genes found in environmental SBG3.

From the two TM7 metagenomes that have been sequenced, radioactive amino acids would be an excellent starting point for STAR-FISH. Both the oral and soil TM7 metagenomes had genes that code for amino acid use. In particular, genes for an arginine deiminase pathway were characterized from the oral TM7 metagenome. The identification of TM7 using other amino acids could indicate the need for a host organism

to provide necessary growth substrates. The need to acquire amino acids from outside sources could indicate possible TM7 pathogenicity.

Another approach would be whole gene amplification of an environmental SBG3 to construct an environmental SBG3 metagenome. This would be accomplished by separating a single environmental SBG3 cell using FISH, combined with either a microfluidic “chip” device (Marcy *et al.*, 2007) or a flow cytometer (Podar *et al.*, 2007), to ideally isolate a single cell in microliter volumes. In this approach the use of Tween 80 with sonication would aid cell separation from floc material. This would increase the available cells to be isolated. However, there is still the floc material in suspension that would clog any cell separation device. A likely approach to remove floc material would be to use a density gradient, which would isolate cells not attached to flocs.

Environmental SBG3 are good candidates to be isolated this way since no environmental SBG3 have been attached to flocs of large size when visualized using FISH. The cell separating device would only separate cells that are the approximate morphology of environmental SBG3 combined with cells fluorescing with both SBG3\_1 FISH and TM7 905 probes. This would be followed by multiple displacement strand amplification where random primers amplify the genome and the resulting products sequenced (Marcy *et al.*, 2007). A metagenome would be constructed and genes identified in this way.

## **Conclusions**

Since TM7 bacteria are not cultivable, frequent sampling from the original sample source is required. If samples were obtained from patients, this would require an

Institutional Review Board (IRB) approval prior to sampling, a qualified physician to extract samples, and a large number of individuals who carry TM7 bacteria from whom to sample routinely. In addition, changes in diet and medication affect the microbial flora in the sample site, such that some sample donors may lose TM7 over time (Oliver *et al.*, 1991). This would require that more sample donor candidates be screened to provide a steady source. Because obtaining human samples is difficult and potentially unreliable, it would be better to have an environmental source that is easily sampled from and consistently contains a TM7 bacterium closely related to human-associated TM7 bacteria.

The two goals of this study were to identify a human-like environmental TM7 bacterial population and qualify the environmental source for further study. The completion of these goals would aid future studies in understanding the role TM7 bacteria play in human diseases.

Activated wastewater provides an accessible sample source that contains TM7. Microscopic FISH data provided evidence that numerous TM7 morphotypes were present in activated wastewater. Our qPCR results supported FISH in that 87% of samples had TM7 16S rRNA genes. TM7 16S rRNA genes were identified throughout the year suggesting that TM7 is a constant resident in activated wastewater. A human-like TM7 bacterium, environmental SBG3, was identified through cloning and ARB software analysis. ARB analysis suggested that environmental SBG3 bacteria were 98.60% similar to human oral SBG3. This suggests that environmental SBG3 are closely related to human oral SBG3. Since it is possible that environmental SBG3 are closely related to oral SBG3, they may have similar genomes and characteristics.

This study identified an environmental SBG3 to use as a model to study human periodontal TM7 bacteria. The goal for further research is to study environmental SBG3 to better understand human periodontal TM7 bacteria. This environmental bacterial model to study human disease-associated bacteria can be extended to encompass other uncultivated bacteria. It is highly likely that other human-associated TM7 bacteria, such as those that are linked to inflammatory bowel disease (Kuehbacher *et al.*, 2008), have a member or close relative in an environmental source that can be used to understand their role in disease. Several uncultivated bacterial groups have been linked to other human oral diseases (Kumar *et al.*, 2006) and human vaginal diseases (Fredricks & Marrazzo, 2005). As additional disease-associated uncultivated bacteria are identified, environmental sources will be found that contain closely related members that can be used to understand their potential role in human diseases.

## References

- Acinas, S., Marcelino, L., Klepac-Ceraj, V., & Polz, M. (2004). Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *Journal of Bacteriology*, *186*(9), 2629-2635.
- Agarwal, A., Clements, J., Sedmak, D., Imler, D., Nahman, N., Orsinelli, D., & Hebert, L. (1997). Subacute bacterial endocarditis masquerading as type III essential mixed cryoglobulinemia. *Journal of the American Society of Nephrology*, *8*(12), 1971-1976.
- Albandar, J., & Kingman, A. (1999). Gingival recession, gingival bleeding, and dental calculus in adults 30 years of age and older in the United States, 1988-1994. *Journal of Periodontology*, *70*(1), 30-43.
- Altschul, S., Madden, T., Schäffer, A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search prog. *Nucleic Acids Research*, *25*(17), 3389-3402.
- Amalfitano, S., & Fazi, S. (2008). Recovery and quantification of bacterial cells associated with streambed sediments. *Journal of Microbiological Methods*, *75*(2), 237-243.
- Amann, R., Ludwig, W., & Schleifer, K. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiology and Molecular Biology Reviews*, *59*(1), 143-169.
- Applied Biosystems (2005). Absolute quantification getting started guide for the applied biosystems 7300/7500/7500 fast real-time PCR system (Rev. D). Foster City, California: Applied Biosystems.
- Barkovskii, A., & Fukui, H. (2004). A simple method for differential isolation of freely dispersed and particle-associated peat microorganisms. *Journal of Microbiological Methods*, *56*(1), 93-105.
- Baron, C. (2005). From bioremediation to biowarfare: On the impact and mechanism of type IV secretion systems. *FEMS Microbiology Letters*, *253*(2), 163-170.

- Bik, E., Eckburg, P., Gill, S., Nelson, K., Purdom, E., Francois, F., Perez-Perez, G., *et al.* (2006). Molecular analysis of the bacterial microbiota in the human stomach. *Proceedings of the National Academy of Sciences of the United States of America*, 103(3), 732-737.
- Bond, P., Erhart, R., Wagner, M., Keller, J., & Blackall, L. (1999). Identification of some of the major groups of bacteria in efficient and nonefficient biological phosphorus removal activated sludge systems. *Applied and Environmental Microbiology*, 65(9), 4077-4084.
- Branson (2002). Operator's manual: Ultrasonic cleaners, models 1510, 2510, 3510, 5510, 8510 (Rev. B). Danbury, Connecticut: Branson.
- Brinig, M., Lepp, P., Ouverney, C., Armitage, G., & Relman, D. (2003). Prevalence of bacteria of division TM7 in human subgingival plaque and their association with disease. *Applied and Environmental Microbiology*, 69(3), 1687-1694.
- CodonCode Corporation (2009). CodonCode aligner user manual. Dedham, Massachusetts: CodonCode Corporation.
- Cole, J., Chai, B., Marsh, T., Farris, R. J., Wang, Q., Kulam, S., Chandra, S., *et al.* (2003). The ribosomal database project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Research*, 31(1), 442-443.
- Crosby, L., & Criddle, C. (2003). Understanding bias in microbial community analysis techniques due to rrn operon copy number heterogeneity. *BioTechniques*, 34(4), 790-794, 796, 798 passim.
- DeLong, E. (1992). Archaea in coastal marine environments. *Proceedings of the National Academy of Sciences of the United States of America*, 89(12), 5685-5689.
- DeSantis, T., Hugenholtz, P., Keller, K., Brodie, E., Larsen, N., Piceno, Y., Phan, R., *et al.* (2006). NAST: A multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Research*, 34(suppl\_2), W394-399.
- Fredricks, D. (2001). Microbial ecology of human skin in health and disease. *Journal of Investigative Dermatology Symposium Proceedings*, 6(3), 167-169.
- Fredricks, D., & Marrazzo, J. (2005). Molecular methodology in determining vaginal flora in health and disease: Its time has come. *Current Infectious Disease Reports*, 7(6), 463-470.



- Fredricks, D., Fiedler, T., Thomas, K., Oakley, B., & Marrazzo, J. (2007). Targeted PCR for detection of vaginal bacteria associated with bacterial vaginosis. *Journal of Clinical Microbiology*, 45(10), 3270-3276.
- Fuchsman, C., & Rocap, G. (2006). Whole-genome reciprocal BLAST analysis reveals that planctomycetes do not share an unusually large number of genes with *Eukarya* and *Archaea*. *Applied and Environmental Microbiology*, 72(10), 6841-6844.
- Gao, Z., Tseng, C., Pei, Z., & Blaser, M. (2007). Molecular analysis of human forearm superficial skin bacterial biota. *Proceedings of the National Academy of Sciences*, 104(8), 2927-2932.
- Gavrilova, H., Dinis, J., & Ouverney, C. (2009). Validation of a fluorescent probe for visualizing a novel environmental bacterium closely associated with a human oral uncultured TM7. Poster presented at College of Science Day, San Jose State University, CA.
- Ghadiri, J., Dinis, J., & Ouverney, C. (2009). The fine line between human and sludge microbiomes. Poster presented at College of Science Day, San Jose State University, CA.
- Google Maps. San Jose/ Santa Clara water treatment plant secondary aerator treatment tanks [map]. Retrieved, June 10, 2010, from <http://maps.google.com>
- Heymas, R., van der Helm, S., deVries, H., Fennema, H., Coutinho, R., Bruisten, S. (2010). Clinical Value of *Treponema pallidum* real-time PCR for diagnosis of syphilis. *Journal of Clinical Microbiology*, 48(2), 497-502.
- Hugenholtz, P., Tyson, G., Webb, R., Wagner, A., & Blackall, L. (2001). Investigation of candidate division TM7, a recently recognized major lineage of the domain bacteria with no known pure-culture representatives. *Applied and Environmental Microbiology*, 67(1), 411-419.
- Hugenholtz, P., Tyson, G., & Blackall, L. (2002). Design and evaluation of 16S rRNA-targeted oligonucleotide probes for fluorescence in situ hybridization. *Methods in Molecular Biology*, 179, 29-42.
- Holben, W. 1994. Isolation and purification of bacterial DNA from soil. In R.W. Weaver (Ed.) *Methods of soil analysis, part 2. Microbiological and biochemical properties.* (Vol 5, pp. 727-751). Madison, Wisconsin: Soil Science Society of America.

- Innis, M., Gelfand, D., Snisky, J., & White, T. (1990). PCR protocols: a guide to methods and applications. San Diego, California: Academic Press Incorporated.
- Janssen, P., Yates, P., Grinton, B., Taylor, P., & Sait, M. (2002). Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions *Acidobacteria*, *Actinobacteria*, *Proteobacteria*, and *Verrucomicrobia*. *Applied and Environmental Microbiology*, *68*(5), 2391-2396.
- Kazor, C., Mitchell, P., Lee, A., Stokes, L., Loesche, W., Dewhirst, F., & Paster, B. (2003). Diversity of bacterial populations on the tongue dorsa of patients with halitosis and healthy patients. *Journal of Clinical Microbiology*, *41*(2), 558-563.
- Klappenbach, J., Saxman, P., Cole, J., & Schmidt, T. (2001). rrndb: The ribosomal RNA operon copy number database. *Nucleic Acids Research*, *29*(1), 181-184.
- Kuczynska, E., & Shelton, D. (1999). Method for detection and enumeration of *Cryptosporidium parvum* oocysts in feces, manures, and soils. *Applied and Environmental Microbiology*, *65*(7), 2820-2826.
- Kuehbacher, T., Rehman, A., Lepage, P., Hellmig, S., Fölsch, U., Schreiber, S., & Ott, S. (2008). Intestinal TM7 bacterial phylogenies in active inflammatory bowel disease. *Journal of Medical Microbiology*, *57*(Pt 12), 1569-1576.
- Kumar, P., Griffen, A., Barton, J., Paster, B., Moeschberger, M., & Leys, E. (2003). New bacterial species associated with chronic periodontitis. *Journal of Dental Research*, *82*(5), 338-344.
- Kumar, P., Griffen, A., Moeschberger, M., & Leys, E. (2005). Identification of candidate periodontal pathogens and beneficial species by quantitative 16S clonal analysis. *Journal of Clinical microbiology*, *43*(8), 3944-3955.
- Lane, D., Pace, B., Olsen, G., Stahl, D., Sogin, M., & Pace, N. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(20), 6955-6959.
- Liesack, W., Weyland, H., & Stackebrandt, E. (1991). Potential risks of gene amplification by PCR as determined by 16S rDNA analysis of a mixed-culture of strict barophilic bacteria. *Microbial Ecology*, *21*(1), 191-198.
- Loesche, W., & Syed, S. (1973). The predominant cultivable flora of carious plaque and carious dentine. *Caries Research*, *7*(3), 201-216.

- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., *et al.* (2004). ARB: A software environment for sequence data. *Nucleic Acids Research*, 32(4), 1363-1371.
- Marcy, Y., Ouverney, C., Bik, E. M., Lösekann, T., Ivanova, N., Martin, H. G., Szeto, E., *et al.* (2007). Dissecting biological “dark matter” with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proceedings of the National Academy of Sciences*, 104(29), 11889-11894.
- Merck. (2001). Polysorbate 80. In Merck Index (Vol. 13, pp 7664). Whitehouse Station, New Jersey: Merck.
- Murray, A., Preston, C., Massana, R., Taylor, L., Blakis, A., Wu, K., & DeLong, E. (1998). Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters near Anvers Island, Antarctica. *Applied and Environmental Microbiology*, 64(7), 2585-2595.
- Oliver, R., Brown, L., & Loe, H. (1991). Variations in the prevalence and extent of periodontitis. *Journal of the American Dental Association (1939)*, 122(6), 43-48.
- Oliver, R., Brown, L., & Loe, H. (1998). Periodontal diseases in the United States population. *Journal of Periodontology*, 69(2), 269-278.
- Ouverney, C., Fuhrman, J. (2000). Marine planktonic *Archea* take up amino acids. *Applied and Environmental Microbiology*, 66(11), 4829-4833.
- Ouverney, C., Armitage, G., & Relman, D. (2003). Single-cell enumeration of an uncultivated TM7 subgroup in the human subgingival crevice. *Applied and Environmental Microbiology*, 69(10), 6294-6298.
- Paster, B., Boches, S., Galvin, J., Ericson, R., Lau, C., Levanos, V., Sahasrabudhe, A., *et al.* (2001). Bacterial diversity in human subgingival plaque. *Journal of Bacteriology*, 183(12), 3770-3783.
- Pei, Z., Bini, E., Yang, L., Zhou, M., Francois, F., & Blaser, M. (2004). Bacterial biota in the human distal esophagus. *Proceedings of the National Academy of Sciences of the United States of America*, 101(12), 4250-4255.
- Podar, M., Abulencia, C., Walcher, M., Hutchison, D., Zengler, K., Garcia, J., Holland, T., *et al.* (2007). Targeted access to the genomes of low-abundance organisms in complex microbial communities. *Applied and Environmental Microbiology*, 73(10), 3205-3214.

- Renvert, S., Pettersson, T., Ohlsson, O., & Persson, G. (2006). Bacterial profile and burden of periodontal infection in subjects with a diagnosis of acute coronary syndrome. *Journal of Periodontology*, 77(7), 1110-1119.
- Roh, C., Villatte, F., Kim, B., & Schmid, R. (2006). Comparative study of methods for extraction and purification of environmental DNA from soil and sludge samples. *Applied Biochemistry and Biotechnology*, 134(2), 97-112.
- Roose-Amsaleg, C., Garnier-Sillam, E., & Harry, M. (2001). Extraction and purification of microbial DNA from soil and sediment samples. *Applied Soil Ecology*, 18(1), 47-60.
- Sakamoto, M., Rocas, I., Siqueira, J., & Benno, Y. (2006). Molecular analysis of bacteria in asymptomatic and symptomatic endodontic infections. *Oral Microbiology and Immunology*, 21(2), 112-122.
- Simelyte, E., Rimpiläinen, M., Zhang, X., & Toivanen, P. (2003). Role of peptidoglycan subtypes in the pathogenesis of bacterial cell wall arthritis. *Annals of the Rheumatic Diseases*, 62(10), 976-982.
- Smith, G., & Thornton, E. (1993). Pathogenicity of *Fusobacterium necrophorum* strains from man and animals. *Epidemiology and Infection*, 110(3), 499-506.
- Tringe, S., & Rubin, E. (2005). Metagenomics: DNA sequencing of environmental samples. *Nature Reviews. Genetics*, 6(11), 805-814.
- Varga, J., Nguyen, V., O'Brien, D., Rodgers, K., Walker, R., & Melville, S. (2006). Type IV pili-dependent gliding motility in the Gram-positive pathogen *Clostridium perfringens* and other *Clostridia*. *Molecular Microbiology*, 62(3), 680-694.
- Wagner, R. (1994). The regulation of ribosomal RNA synthesis and bacterial cell growth. *Archives of Microbiology*, 161, 100-109
- Waites, K., & Atkinson, T. (2009). The role of *Mycoplasma* in upper respiratory infections. *Current Infectious Disease Reports*, 11(3), 198-206.
- Yeates, C., Gillings, M., Davison, A., Altavilla, N., & Veal, D. (1997). PCR amplification of crude microbial DNA extracted from soil. *Letters in Applied Microbiology*, 25(4), 303-307.
- Yeates, C., Gillings, M., Davison, A., Altavilla, N., & Veal, D. (1998). Methods for microbial DNA extraction from soil for PCR amplification. *Biological Procedures Online*, 1, 40-47.

Zengler, K., Toledo, G., Rappé, M., Elkins, J., Mathur, E., Short, J. M., & Keller, M. (2002). Cultivating the uncultured. *Proceedings of the National Academy of Sciences of the United States of America*, 99(24), 15681-15686.