# THE DIVERSE BACTERIAL COMMUNITY IN SAPELO ISLAND AND STUDY ON THE PATHWAY OF DMSP DEGRADATION IN MARINE BACTERIA

by

SHIYAO WANG

(Under the Direction of William B. Whitman)

## ABSTRACT

Salt marshes/estuarine systems are one of the most biologically productive ecosystems on the planet. In this study, further analyses examined the diversity of the bacterial community in a salt marsh/estuarine ecosystem in the southeastern coast of the US.

The study on bacterioplankton has been considered very important because prokaryotes are essential components of the oceanic food web and catalyze many major biogeochemical processes in the sea. Different methods to collect bacterioplankton samples from seawater may have effect on microarrays result. A whole genome microarray of *Silicibacter pomeroyi* DSS3 was analyzed as a standard to indicate the expression of collected seawater sample through different methods of collection.

*Silicibacter pomeroyi*, a member of a marine roseobacter clade, is a model system for the study of dimethylsulfoniopropionate (DMSP) metabolism. The degradation of DMSP occurs through two competing pathways, known as the cleavage and the demethylation pathways. Bioinformatics approaches were used to identify candidate genes for enzymes involved in DMSP degradation. A genetic system was applied to test the function of these genes. The growth of mutants with different carbon source was also measured.

INDEX WORDS: diversity, bacterial community, salt marsh/estuarine ecosystem, bacterioplankton, microarray data analyses, bioinformatics, genetic system, *Silicibacter pomeroyi*, dimethylsulfoniopropionate, DMSP, metabolism, growth

# THE DIVERSE BACTERIAL COMMUNITY IN SAPELO ISLAND AND STUDY ON THE PATHWAY OF DMSP DEGRADATION IN MARINE BACTERIA

by

SHIYAO WANG

B.S., University of Science and Technology of China, 2007

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2009

# THE DIVERSE BACTERIAL COMMUNITY IN SAPELO ISLAND AND STUDY ON THE PATHWAY OF DMSP DEGRADATION IN MARINE BACTERIA

by

SHIYAO WANG

| | |
|---|---|
| Major Professor: | William B. Whitman |
| Committee: | Mary Ann Moran<br>Jan Mrazek |

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2009

DEDICATION

This dissertation is dedicated to my parents and grandparents who give me their loves and supports all the time. My grandma pasted when I prepared my thesis, and didn't give me the chance to accompany with her in her last life time. This thesis is for you and I love you!

ACKNOWLEDGEMENTS

I would like to acknowledge the truly generous guidance, support, and mentoring given by Barny Whitman. His humor and intelligence have aided me greatly. Mary Ann Moran has also been a truly outstanding and supportive mentor. Both, with their very different personalities, have been invaluable in my scientific training. I would also like to acknowledge the help and guidance of my committee members: particularly Jan Mrazek for his advice in the bioinformatics, and Larry Shimkets, for his help and encouragement. And I also want to thank all of members in Whitman's lab and Moran's lab for their supports.

TABLE OF CONTENTS

Page

# LIST OF TABLES

Page

LIST OF FIGURES

Page

# CHAPTER 1

# LITERATURE REVIEW

## 1.1 ABSTRACT

Salt marshes/estuarine systems are one of the most biologically productive ecosystems on the planet. Especially the salt marshes, they are believed to provide plenty food which enhance the yield of many marine organisms and also act as a negative feedback for global warming and carbon cycle.

In order to investigate the diversity of prokaryotes, their physiological and genetic characteristics, and their geochemical activities in a salt marsh/estuarine ecosystem in the southeastern coast of the US, previous studies sampled from salt marsh and estuarine of the Sapelo Island Microbial Observatory (SIMO). Then, the 16S rRNA gene libraries were constructed based on the sampling experiment for the further study on phylogenetic diversity and composition of the bacterial community. One advantage of constructing sequence libraries is that it provides a permanent record of an organism's distribution and can keep the record for future use. In the previous study, BLAST and statistical analyses were used to assign clones to taxonomic groups and display the prokaryotic composition as well as diversity of nucleic acids from environmental sample. In this study, further analysis was applied to examine the differences of bacterial community in different ecosystems.

Seasonal differences of bacterial community in seawater, GA, were also examined and restricted to only certain taxa.

Because prokaryotes are essential components of the oceanic food web, the importance of studying on bacterioplankton have been noticed and thought as catalysts of most major biogeochemical processes in the sea (Moran MA 2004). *Silicibacter pomeroy*, a member of the marine Roseobacter clade which comprise ~10% to 20% of coastal and oceanic bacterioplankton (Moran MA 2004). To characterize the genetic diversity, a whole genome *S. pomeroyi* microarray approach was chosen. When microarray approach is chosen, the limitations of microarray become a main concern which may affect the accuracy of the result. In this study, two different treatments, "cooling" and "adding stop solution", were used when sampling. Three replicates of microarrays of each treatment were applied to compare with the control group which was sampling without any special treatment.

*S. pomeroyi* is a model system for the study of dimethylsulfoniopropionate (DMSP) metabolism. The degradation of DMSP occurs through two competing pathways, known as the cleavage and the demethylation pathways. The cleavage pathway results in the formation of dimethylsulfide (DMS), the major natural source of sulfur to the atmosphere (Bates 1987; SIMO 2001).

Microarray data were analyzed using bioinformatics methods. Differences in community composition among the groups were visualized using multidimensional scaling (MDS). And differences in expression level among the groups were visualized in heat map.

Bioinformatic approaches were used to identify candidate genes for enzymes involved in DMSP degradation. A genetic system was applied to test the function of these genes. Studies included: identification of minimum inhibitory concentrations of antibiotics and antibiotic resistance cassettes functional in *S. pomeroyi*; transformation by electroporation with DNA protected by SssI methylase;

and construction of gene replacement mutants using suicide vectors. This system was initially constructed by James R. Henriksen and was revised in this study in order to improve the efficiency. The growth of mutants with different carbon sources was also measured.

## 1.2 INTRODUCTION

Salt marshes are some of the most biologically productive ecosystems on the planet and important components of the coastal marine landscape (Pomeroy, 1981). They act as a negative feedback for global warming and the carbon cycle (Hussein, 2004). In addition, they also store a large amount of inorganic carbon and can be one of the major inorganic carbon sources to adjacent coastal waters. (Cai et al, 1998; Wang et al, 2004)

For the east coast of the United States, most salt marshes were drained to different degrees for historic reasons, such as to eradicate mosquitoes, to construct roads, and to expand residential areas. Therefore, it is difficult to choose an almost intact marsh to sample. The sampling site we chose was on Sapelo Island, which is a state-protected island and only reachable by boat with a permit from the Georgia Department of Natural Resources (DNR). Therefore, the great biodiversity is well protected on this island.

Salt marsh sediments are anaerobic, and sulfate reduction is the dominant respiratory pathway. Iron reduction, denitrification, and aerobic respiration are significant processes as well. Although bacterial and archaeal metabolic pathways have been well studied in this ecological environment taxonomic, the composition microbial community is poorly understood.

Through the polymerase chain reaction (PCR), environmental rRNA gene libraries were constructed and sequenced for this study since the sequences of rRNA genes provide an unequivocal

identification of the organism. In a prior study, the RDPquery was used to automatically identify sequences through comparison to the Ribosomal Database Project (RDP) of annotated bacterial small-subunit 16S rRNA sequences. (Lasher, et al 2009)

During studies on the ecosystems of coastal seawater of Atlantic Ocean, a member of the marine Roseobacter clade, *Silicibacter pomeroyi*, was isolated, and its genome was sequenced (Moran MA 2004). Microarrays were used to study gene expression in *S. pomeroyi*. But the limitations which may affect the accuracy of microarrays are always an issue. There are three main aspects which may limit the accuracy of microarrays for environmental samples. First, the representativeness of genes targeted by the array depends on the taxonomic composition of sample. Second, the measurement of gene expression depends on the hybridization of RNA with probes. It is possible that some species may be not detectable because the *S. pomeroyi* probes do not hybridize to their mRNA. Third, the accuracy of microarray depends on the condition of the mRNA extracted from the samples. During collection, the quality and quantity of mRNA can change very quickly. In this case, the levels of mRNA will not reflect the environmental condition if they change during sampling. In this study, two sampling treatments were applied in order to find out if they might improve the microarray accuracy. When sampling, “cooling” and “stop solution” treatments were applied and compared to no treatment in order to test if there is an effect on microarray accuracy due to improvements in mRNA preservation.

Dimethylsulfoniopropionate (DMSP) is a ubiquitous compound in the ocean and a major source of carbon and reduced sulfur for marine bacteria. Its metabolism is important in the carbon and sulfur cycles and influences global climate due to the degradation product dimethyl sulfide (DMS). In the atmosphere, DMS is oxidized to methanesulfonic acid and eventually to particulate sulfate, which then acts as hygroscopic cloud condensation nuclei. Concentrations of DMS in the atmosphere

correlate with the cloud condensation nuclei loads (Bates 1987). Increased solar radiation leads to an increase in the abundance of algae that produce DMSP, which could in turn lead to an increase in the amount of DMS formed and released into the atmosphere, which would lead to increased cloud formation and albedo. As an increase in cloud cover would then decrease the abundance of algae, this system would result in a negative feedback loop, with DMS acting as an anti-greenhouse gas. The majority of DMSP degradation is through the demethylation pathway, but the cleavage pathway is also important (Kiene 1999; Kiene 2000). *S. pomeroyi*, is a model system for the study of DMSP degradation. *S. pomeroyi* can cleave DMSP to DMS and carry out demethylation to methanethiol (MeSH) as well as degrade MeSH. To study the genetics of this pathway on the gene level, several candidate genes which may be involved in DMSP degradation were selected. The role of genes SPO1703, SPOA268, SPOA269, and SPO2134 were then tested by mutagenesis.

## 1.3 IMPORTANCE

The diversity of prokaryotes, their physiological and genetic characteristics, and their geochemical activities in a salt marsh/estuarine ecosystem on the southeastern coast of the US provide an important point of view to study the microbial distribution in salt marsh/estuarine ecosystem. Construction of sequence libraries provides a permanent record of an organism's distribution. The composition of microbial community in this important ecosystem could be understood better by studying the microbial community distribution of samples collected from different seasons.

The study on bacterioplankton has been considered very important because prokaryotes are essential components of the oceanic food web and catalyze many major biogeochemical processes in

the sea (Moran MA 2004). When study with the environmental sample, the ways of sample collection may directed effect the results of microarray. There are two common treatments to prevent the degradation of mRNA during sampling. The whole genome microarray for *S. pomeroyi* was applied in the study.

*S. pomeroyi* is a model system for the study of DMSP degradation, which includes both DMSP cleavage and demethylation pathways, and can be easily manipulated in pure culture. *S. pomeroyi* can cleave DMSP to DMS and carry out demethylation to MeSH. DMSP is a significant marine carbon and sulfur source and has a large environmental and climatic impact due to DMS release from the oceans. The degradation pathway of DMSP and the role that *S. pomeroyi* plays in this pathway is an area of active research. The understanding of the mechanism of the DMSP degradation will expand our knowledge of bacterial biochemical activities, the diversity of genes encoding for these important functions, and the details of the metabolism of this compound. To study the genetics of this pathway on the gene level, several candidate genes which may be involved in DMSP degradation were selected.

1.4 REFERENCES

Donald F. Boesch, R. Eugene Turner. (1984). "Dependence of Fishery Species on Salt Marshes: The Role of Food and Refuge" Estuaries **7**: 460 - 468

González, J. M., Moran, M. A. (1997). "Numerical dominance of a group of marine bacteria in the a-subclass of Proteobacteria in coastal seawater." Appl. Environ. Microbiol **63**: 4237 - 4242.

González, J. M. e. a. (2000). "Bacterial community structure associated with a dimethylsulfoniopropionateproducing North Atlantic algal bloom." Appl. Environ. Microbiol **66**: 4237 - 4246.

W. Soto and J.Gutierrez. (2009). Salinity and temperature effects on physiology responses of *Vibrio fischeri* from diverse ecological niches. Microb Ecol. **57**(1), 140-150.

Moran MA, B. A., González JM, Heidelberg JF, Whitman WB, Kiene RP, Henriksen JR, King GM, Belas R, Fuqua C, Brinkac L, Lewis M, Johri S, Weaver B, Pai G, Eisen JA, Rahe E, Sheldon WM, Ye W, Miller TR, Carlton J, Rasko DA, Paulsen IT, Ren Q, Daugherty SC, Deboy RT, Dodson RJ, Durkin AS, Madupu R, Nelson WC, Sullivan SA, Rosovitz MJ, Haft DH, Selengut J, Ward N (2004). "Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment." Nature **432**(7019): 910 - 913.

Pukall, R., Buntefuß, D., Fru¨ hling, A., Rohde, M., Kroppenstedt, R. M., Burghardt, J., Lebaron, P., Bernard, L., Stackebrandt, E. (1999). "*Sulfitobacter mediterraneus* sp. nov., a new sulfite-oxidizing member of the a-Proteobacteria." Int J Syst Bacteriol **49**: 513 - 519.

Thompson JD, H. D., Gibson TJ (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res **22**: 4673 - 4680.

Wagner-Do¨ bler, I., Rheims, H., Felske, A., Pukall, R. & Tindall, B. J (2003). "Jannaschia helgolandensis gen. nov., sp. nov., a novel abundant member of the marine Roseobacter clade from the North Sea." Int J Syst Bacteriol **53**: 731 - 738.

Wang ZA, C. W. (2004). "Carbon dioxide degassing and inorganic carbon export from a marsh-dominated estuary (the Duplin River): a marsh CO2 pump." Limnol Oceanogr **49**: 341 - 354.

Kiene, R. P., Linn, L. J., Bruton, J. A (2000). "New and important roles for DMSP in marine microbial communities." J Sea Res 43: 209 - 224.

Bates, T. S., R. J. Charlson, and R. H. Gammon (1987). "Evidence fof the climatic role of marine biogenic sulphur." Nature 329: 319 - 321.

Elledge, M. Z. L. S. J. (2007). "Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC." NATURE METHODS 4.

# CHAPTER 2

# A STUDY OF A DIVERSE BACTERIAL COMMUNITY

## 2.1 ABSTRACT

Salt marshes/estuarine systems are some of the most biologically productive ecosystems on the planet. Salt marshes especially are believed to provide organic carbon inputs which enhance the growth of many marine organisms and also act as a negative feedback for global warming and the carbon cycle (Boesch, 1984).

In order to investigate the diversity of prokaryotes, their physiological and genetic characteristics, and their geochemical activities in a salt marsh/estuarine ecosystem on the southeastern coast of the US, previous studies obtained samples from salt marsh and estuaries of the Sapelo Island Microbial Observatory (SIMO; table 2.1). Then, 16S rRNA gene libraries were constructed for further study of the phylogenetic diversity and composition of the bacterial communities. One advantage of constructing sequence libraries is that it provides a permanent record of an organism's distribution. In the previous study, BLAST and statistical analyses were used to assign clones to taxonomic groups and analyze the diversity of nucleic acids from environmental samples. In this study, further analyses examined the differences in the bacterial community in these samples. Seasonal differences of the bacterial communities in seawater were also examined and shown to be restricted to only certain taxa.

## 2.2 INTRODUCTION

Salt marshes are some of the most biologically productive ecosystems on the planet and important components of the coastal marine landscape (Pomeroy, 1981). They act as a negative feedback for global warming and the carbon cycle (Hussein, 2004). In addition, they also store a large amount of inorganic carbon and can be one of the major inorganic carbon sources to adjacent coastal waters. (Cai et al, 1998; Wang et al, 2004)

For the east coast of the United States, most salt marshes were drained to different degrees for historic reasons, such as to eradicate mosquitoes, to construct roads, and to expand residential areas. Therefore, it is difficult to choose an almost intact marsh to sample. The sampling site we chose was on Sapelo Island, which is a state-protected island and only reachable by boat with a permit from the Georgia Department of Natural Resources (DNR). Therefore, the great biodiversity is well protected on this island.

Salt marsh sediments are anaerobic, and sulfate reduction is the dominant respiratory pathway. Iron reduction, denitrification, and aerobic respiration are significant processes as well. Although bacterial and archaeal metabolic pathways have been well studied in this ecological environment taxonomic, the composition microbial community is poorly understood.

Through the polymerase chain reaction (PCR), environmental rRNA gene libraries were constructed and sequenced for this study since the sequences of rRNA genes provide an unequivocal identification of the organism. In a prior study, the RDPquery was used to automatically identify sequences through comparison to the Ribosomal Database Project (RDP) of annotated bacterial small-subunit 16S rRNA sequences. (Lasher, et al 2009)

In this study, DOTUR was used with a distance matrix generated by PHYLIP to calculate the diversity of the bacterial community. For the samples collected from the salt marsh, sediment, 1,186 clones formed 817 OTUs at 99% sequence similarity. For the samples collected from estuarine, seawater, 1,605 clones formed 429 OTUs at 99% sequence similarity. The sediment and seawater libraries were compared to that of a soil sample, which was collected from agricultural and forest soils in Kansas, Michigan, and Georgia. The total richness of soil samples was also high, with 11,160 clones forming 3,934 OTUs at 99 % sequence similarity.

In order to compare bacterial composition in different locations and seasons, the Libshuff program was applied. After comparing two sampling locations, Dean Creek (a tidal channel) and Doboy Sound (an open tidal bay) with Libshuff, the population of seawater bacteria did not appear to be a significantly different in these two locations. From the output of Libshuff, different seasons showed a major shift in population composition. Studies of particular organisms' growth preferences may provide a specific rationale for the shift in seasonal distributions.

Table 2.1 Basic information on sampling strategy for the SIMO study

| Year | Date | Season | Size Fraction | | Location[2] |
|---|---|---|---|---|---|
| | | | >1 μm | 0.2-1 μm | |
| 2000 | 17-Oct | Fall | CW0102[1] | | DC |
| 2001 | 26-Jan | Winter | CW2125 | | DC |
| | 26-Apr | Spring | CW3739 | | DC |
| | 18-Jul | Summer | CW4748 | CW4950 | DC |
| | 19-Jul | Summer | SW5152 | SW5354 | DS |
| 2003 | 7-Oct | Fall | CW114115 | | DC |
| 2004 | 10-Feb | Winter | CW125126 | CW127128 | DC |
| | 11-Feb | Winter | SW129130 | SW131132 | DS |
| | 23-Apr | Spring | CW135136 | CW137138 | DC |
| | 22-Jul | Summer | CW143144 | CW145146 | DC |
| | 22-Jul | Summer | SW147148 | SW149150 | DS |
| | 1-Oct | Fall | CW151152 | | DC |
| 2005 | 18-Jan | Winter | CW159160 | | DC |
| | 26-Apr | Spring | CW167168 | | DC |

1: Names of sample libraries. CW = Creek water; SW = Sound water.

2: Sampling locations: DC stands for Dean Creek; DS stands for Doboy Sound.

## 2.3 MATERIALS AND METHODS

### *2.3.1 Sequence analysis*

Previous studies focused on microbial community composition at SIMO by the direct retrieval and analysis of rRNA from bacterioplankton (Lasher, et al 2009), and these data were reanalyzed for the present study. 16S rRNA sequences were aligned in ClustalW (Thompson, 1994). Then the DNADIST module of PHYLIP (http://evolution.genetics.washington.edu/phylip.html ) was used to calculate the pairwise evolutionary distances. Finally, rarefaction curves were calculated with DOTUR using the average neighbor option (Schloss, 2005).

### *2.3.2 Study on seasonal difference*

OTUs which included both high S_ab values and high BLAST scores were selected in the seawater dataset. MEGA 4.0 was used to generate phylogenetic trees of selected OTUs (Kumar, 2004). Then these sequences were analyzed for taxonomy as a function of sampling seasons.

## 2.4 RESULTS AND DISCUSSION

### *2.4.1 Community Diversity*

Seawater libraries were from two locations: Dean Creek and Doboy Sound. Dean Creek is a tidal channel (water depth of 1-5 m), and surface water samples were collected. Doboy Sound is an open tidal bay (water depths of 5-10 m) near the southern end of Sapelo Island, and surface water samples were also collected. The clone libraries of sequences across the two locations and four seasons were pooled. In a previous study, when defining the operational taxonomic units or OTUs at $\geqslant$99%

sequence similarity, 429 OTUs were identified in the pooled sea water libraries. The Shannon indices for the individual libraries were all close to the maximum values. Even for the library of 1,605 clones pooled across seasons and locations, the Shannon index remained at 70% of the maximum value (H/Hmax), and the coverage and Chao1 estimator were 0.777 and 873, respectively. When OTUs were defined at ⩾97% sequence similarity, the Shannon index remained at 62% of its maximum value, and coverage and Chao1 estimator were 0.846 and 537, respectively (Unpublished data).

In the previous study, when defining the operational taxonomic units or OTUs at ⩾99% sequence similarity, 871 OTUs were identified in the pooled sediment libraries. The Shannon indices for the individual libraries were all close to the maximum values. Even for the library of 1,186 clones pooled across seasons, the Shannon index remained at 91% of the maximum value (H/Hmax), and the coverage and Chao1 estimator were 0.441 and 3,290, respectively. When OTUs were defined at ⩾ 97% sequence similarity, the Shannon index remained at 82% of its maximum value, and coverage and Chao1 estimator were 0.600 and 2,290, respectively (Lasher, 2009).

In this study, when defining the operational taxonomic units or OTUs at ⩾99% similarity, 3934 OTUs were identified in the pooled soil libraries. The Shannon indices for the individual libraries were all close to the maximum values. Even for the library of 11,160 clones pooled across locations, the Shannon index remained at 87% of the maximum value (H/Hmax), and the coverage and Chao1 estimator were 0.881 and 7,272, respectively. When OTUs were defined at ⩾97% sequence similarity, the Shannon index remained at 79% of its maximum value, and coverage and Chao1 estimator were 0.952 and 3,683, respectively.

Rarefaction curves were generated for the three libraries, by plotting the number of clones against the number of OTUs at ⩾99% sequence similarity in order to find out if microbial diversity was

different in seawater, sediment (salt marsh), and agricultural and forest soils (Fig 2.1). Although it seemed that the diversity of the bacterial community in soil appeared larger than that in sediments and much larger than that in seawater, this observation mau have been influenced by several artifacts. First, diversity indices are sensitive to the sample size, which made it difficult to compare the H/Hmax and Chao 1 values (Kemp PF, 2004). Secondly, soil samples were collected from the agricultural and forest of three different states (GA, KS, and MI); seawater samples were collected from Dean Creek and Doboy Sound estuaries (GA) in two seasons, and sediment samples were collected from Dean Creak salt marsh (GA) in different seasons. Therefore, the apparent diversity may represent differences of the pooled samples and not the inherent diversity within any particular environments.

Rarefaction curves were calculated for individual libraries (Fig. 2.2). Sample size was fixed at one value to make more accurate comparisons (Fig. 2.3). The numbers of observed OTUs at sample sizes of 50, which could include the smallest library, were then compared. For sediment and soil, the mean values were very similar, at 46.2±1.4 (n=10) and 47.5±1.5 (n=42), respectively (Fig. 2.3). In contrast, the mean of libraries of estuarine water was only 28.4±7.2 (n=15). In this case, the bacterial diversity in sediments was comparable to that observed in soil by nearly identical methods, and greatly exceeded that of estuarine waters. These analyses confirmed that the bacterial communities of these sediments are as diverse as those in agricultural and forest soils (Jangid K, 2008).

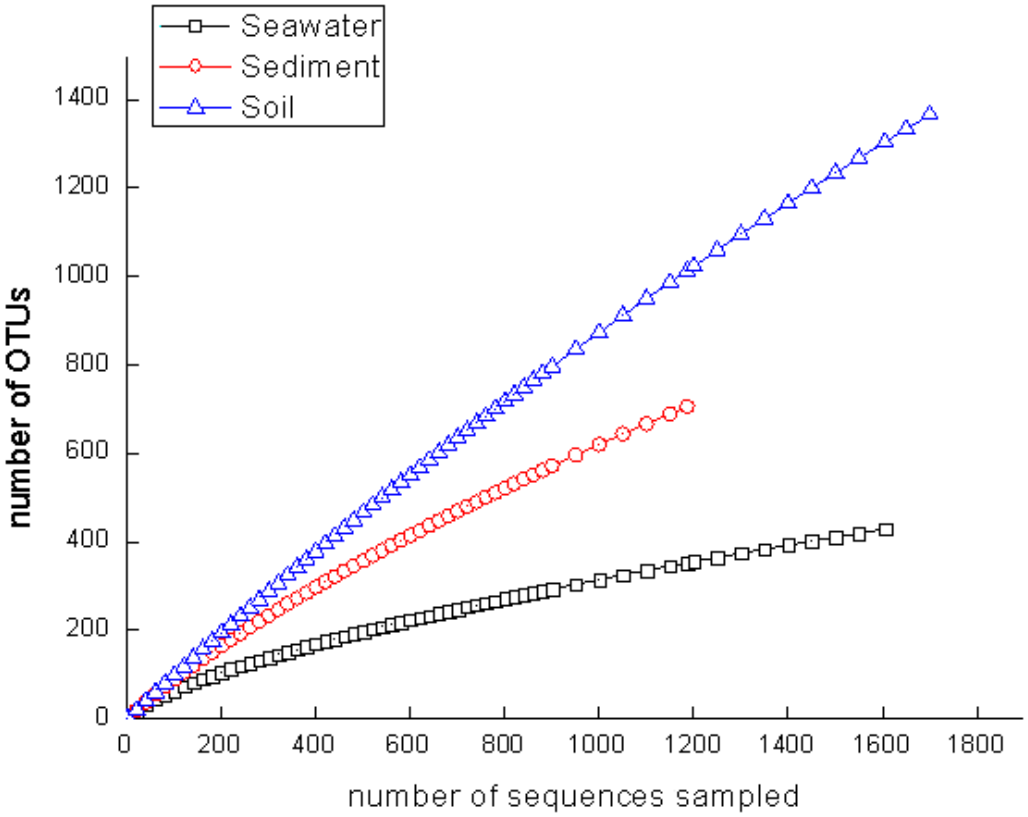



Fig 2.1 Rarefaction curves of the clone libraries: estuarine water (□), salt marsh sediment (○), and agricultural and forest soil ( ∆ ).

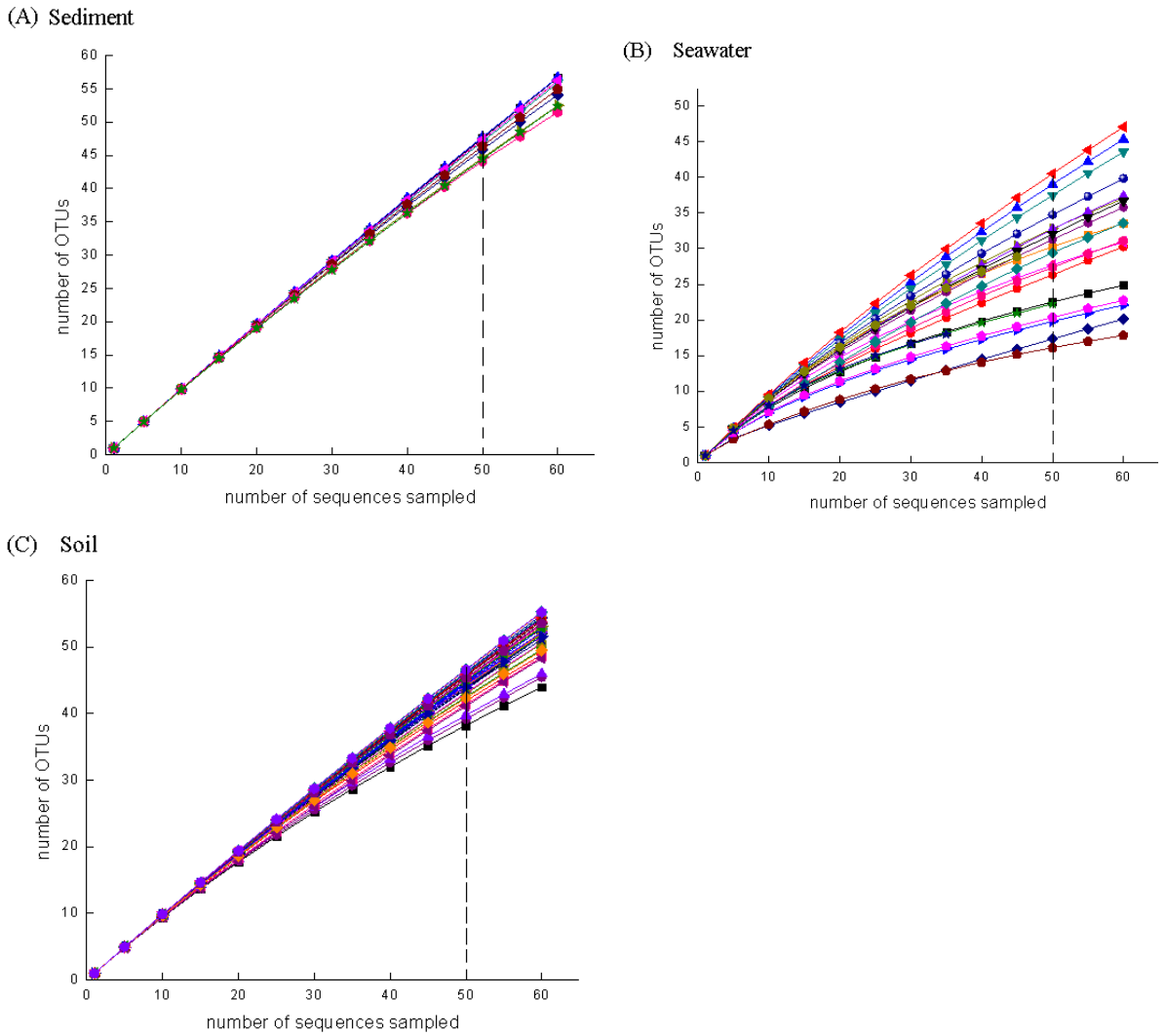



Fig 2.2 Rarefaction curves of samples collected form (A) SIMO salt marsh (Lasher, 2009) (B) SIMO estuarine (Chris Lasher, unpublished), and (C) agricultural and forest soil at KA, GA, MI. Each curve represented an individual library from a unique sample of each of the environments. The number of libraries correlate for the sediment, seawater, and soil were 9, 15, and 42.

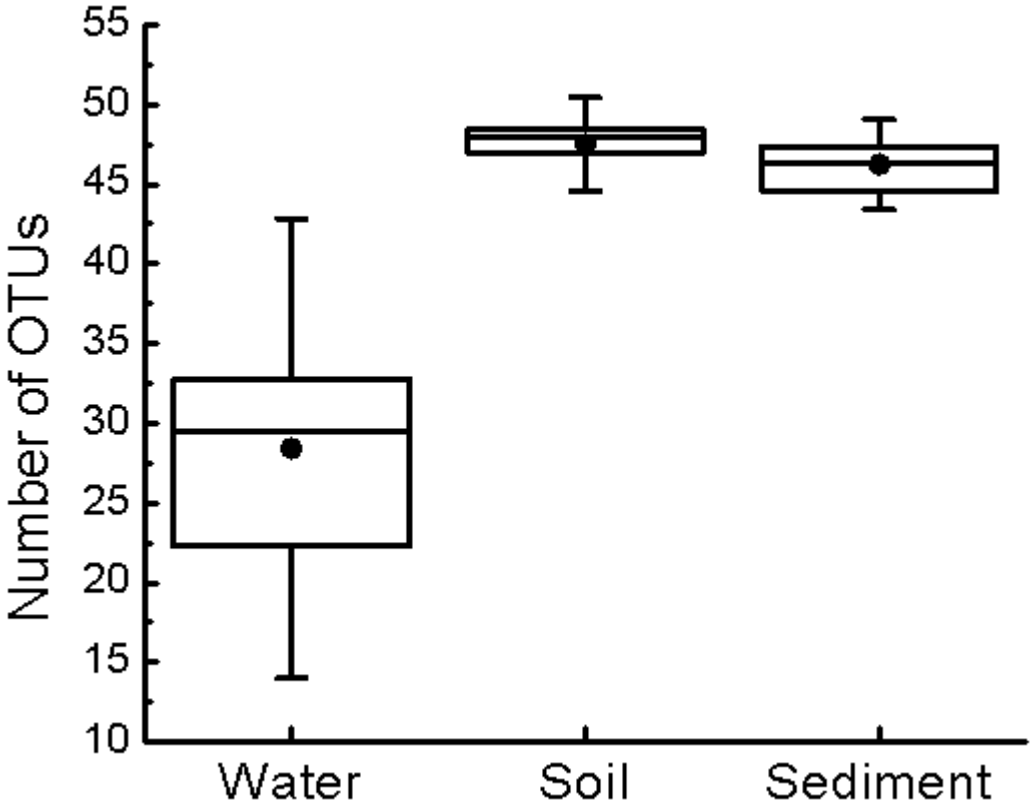



Fig. 2.3 Comparison of the diversity 16S rRNA gene libraries of seawater, agricultural soil and estuarine sediment. Libraries were all constructed by low cycle PCR, and rarefaction curves were calculated using DOTUR. Descriptive statistics were then calculated for number of OTUs observed at a sample size of 50. The box includes the middle two quartiles, the line is the median, ● is the mean, and bars denote two standard deviations about the mean.

*2.4.3 Specific phylogenetic groups*

For the SIMO seawater libraries, one quarter of the clones were related to one or more undescribed orders within the *γ- Proteobacteria*. Other abundant groups included the *δ-Proteobacteria*, *Bacteroidetes*, and *Cyanobacteria.* In order to understand more about the OTUs, OTUs (similarity $\geqslant$ 97%) which were closely related to characterized bacterial species by the criteria of both high S_ab values and high BLAST scores were selected for further study. Among 312 OTUs, there are 4 OTUs totaling 194 sequences (12% of the library) that fit these requirements (Fig 2.4).

Organisms in Group 1 were clustered within the α subclass of the *Proteobacteria.* Organisms in group 2 were clustered within the *Cyanobacteria.* Organisms in group 3 were clustered within the γ subclass of the *Proteobacteria.* The phylogenetic tree of taxa of these four groups was generated with MEGA (Fig 2.4), and the evolutionary relationships of the main bacterial community of SIMO seawater was shown.

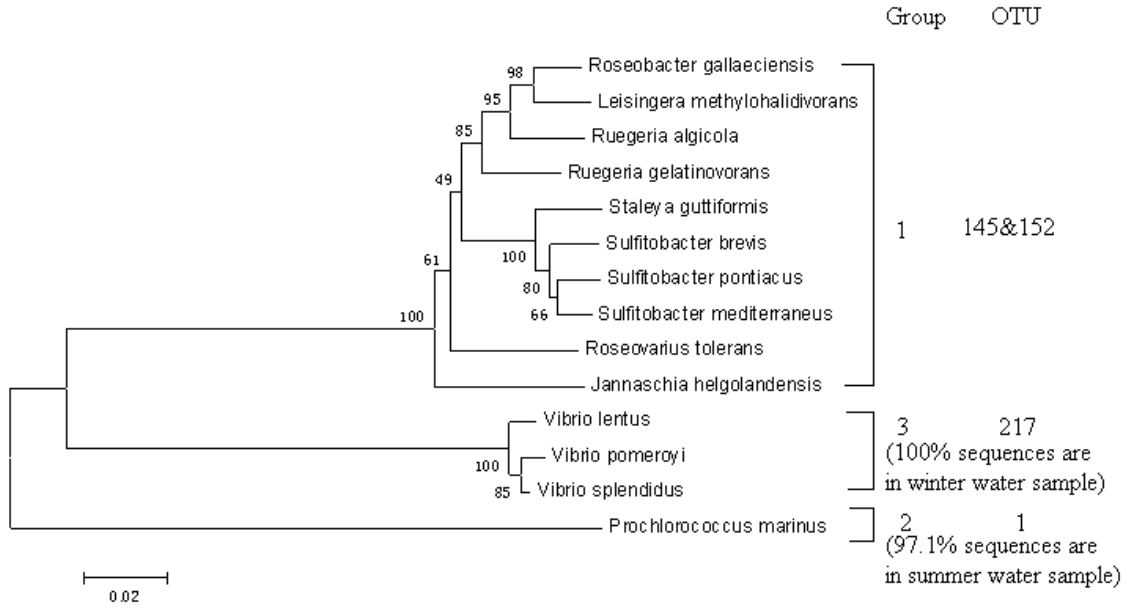



Fig 2.4 Phylogenetic tree of taxa which dominated the bacterial community of seawater library.

From the result, we can tell that for group 1, which clustered in the α subclass of the *Proteobacteria*, occurred in seawater samples in all 4 seasons. Group 2, which clustered in the *Cyanobacteria* occurred frequently (97.1%) in the summer water samples. Group 3, which clustered with the *Vibrionaceae*, a family of γ subclass of the *Proteobacteria*, occurred frequently (~100%) in the winter water samples (Fig 2.4). In summary, the seawater bacterial community responded to season and appeared to include several specific phylogenetic groups with consistent temporal patterns. This suggests that some phylogenetic groups may have been responding to specific geochemical properties of seawater.

For group 2, the representative species is *Prochlorococcus marinus* (Table 2.2), a small chlorophyll b-containing cyanobacterium in the ocean that is ubiquitous between 40$^{o}$N and 40$^{o}$S and abundant in nutrient poor regions of the oceans (Partensky, 1999). To explain its abundance in SIMO during summer, we can hypothesize that the seawater carried by the Gulf stream current to SIMO which may contain fewer nutrients in summer, and the stronger sun light may also lead to the increase in *Prochlorococcus*.

For group 3, the representative species are *Vibrio pomeroyi*, *Vibrio splendidus*, and *Vibrio lentus* (Table 2.2). The specific occurrence of *Vibrio* in winter SIMO water may indicate a lower optimal temperature of these three species of *Vibrio.* Although there is no record of the optimum temperature of these particular species*,* other species of *Vibrio* (*V. fischeri, V. vulnificus,* and *V. tapetis*) have an optimum range of 7 — 22 $^{o}$C, which is lower than species in group 1 (Soto 2009; Kaspar 1993; Borrego 1996). Group 1 has a larger range of optimum temperature, salinity and pH than group 2 or group 3. This may also contribute to why species in group 1 occurred in seawater samples in all four seasons. From Table 2.2, we can find that there is no obvious preference of these three groups for

carbon and nitrogen source utilization, or for the susceptibility to penicillin. Moreover, the G+C contents are different in all of these three groups. For group 1, all of G+C contents are above 55 (mol %). For group 2, G+C content is 32 (mol %). For group 3, G+C contents are around 44 (mol %). Although difference in G+C contents can not be used to explain the seasonal difference, it supports the approach used here for group classification.

| Group (G) | G1 | | | | | | | | | | G2 | G3 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Species | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| Morphology | Ovoid rods | Ovoid rods | Rods/ovoid | Rod | Rod | Rod | Rod/ovoid | Rod | Rod | Rod | Rod | | Curved Rod | Rod |
| Flagella | Polar | Subpolar | Subpolar | NF | subpolar | | | | + | + | | | ND | polar |
| Motility | + | + | + | +/− | + | + | | − | + | + | | | + | + |
| Temperature optimum(° C) | 23--27 | 25--30 | 27 | ND | 17--28 | 22--25 | 26 | 25--30 | 12--20 | 3--26 | NF | NF | NF | NF |
| Salinity optimum (%(w/v)) | 0.6--12 | 0.6--12 | 1.5--6.5 | 1.0--8.0 | 1.5-2.0 | 2.0--2.5 | ND | 1.0--7.0 | 1.0 | 1.0--2.0 | NF | 6.0--8.0 | 2.0--6.0 | |
| pH optimum | 7.7 | 7.5 | 7.7 | 6.2--9.0 | 7.0--7.5 | 7.5--7.8 | ND | 7.0--8.0 | 8.09 | | NF | | | |
| Bchl a | − | − | − | +/- | − | − | − | | +/- | − | − | | | − |
| Utilization of Carbon and energy sources: | | | | | | | | | | | | | | |
| D-glucose | + | + | − | + | + | w | | + | w | + | + | | | |
| D-Mannitol | | ND | | + | | | | | − | | ND | + | + | + |
| Fructose | | + | − | | | | | + | w | | | | | |
| Acetate | | + | − | + | + | + | + | | + | + | | | | − |
| Methanol | ND | − | − | + | | − | ND | | − | − | | | | |
| Pyruvate | | + | − | + | + | + | + | | + | | + | | | + |
| Maltose | + | + | | | + | | | | − | | | | + | + |
| Sucrose | + | + | | | | | | | w | | + | + | − | − |
| Trehalose | + | + | | | | | | | w | | | | | |
| Citrate | − | + | − | − | + | w | | + | − | + | | − | + | + |
| Butyrate | + | − | | w | + | + | ND | | − | + | | | | |
| 2-Ketoglutarate | + | + | | | | | | + | − | | | | | + |
| Utilization of nitrogen sources: | | | | | | | | | | | | | | |
| Ammonia | | | | | | + | | | | | + | | | |
| Urea | | | | | − | | | | | | + | − | − | |
| Nitrate | | | | | | | | | | | − | | | |
| Nitrate reduced to nitrite | − | − | − | − | − | + | + | − | + | − | − | + | + | |
| Sulfite oxidation | | | | | + | + | | | | − | | | | |
| Oxidase | + | + | | w | + | + | + | w | + | + | | + | + | |
| Susceptibility to penicillin (10U) | − | + | − | +/- | | | | | + | | | | | |
| Non-diffusible pigment | Brownish | Beige | Brown | Red | − | − | Beige | | Beige | | Kelly | | − | − |
| Diffusible pigment | Brown | − | | | | | | | | | | | | |
| G-C content (mol%) | 57.6--58.0 | 60 | 60.5 | 63.3--63.4 | 59 | 61.7--62.5 | 59 | 63 | 55.0--56.3 | 57.9--58.1 | 32 | | 45--46 | 44 |

"+": positive
"-": negative
"ND": not detactive
"NF": not found
"w": weak reaction

Table 2.2 Characteristics of species in 3 groups. Column 1 to 10 are species in group 1; column 11 is species in group 2; column 12 to 14 are species in group 3.

1 *Roseobacter gallaeciensis* BS107T (Ruiz-Ponte, 1998); 2 *Ruegeria algicola* ATCC 51440T (Lafay, 1995); 3 *Leisingera methylohalidivorans* MB2T (Schaefer, 2002); 4 *Roseovarius tolerans* (T) Ekho Lake172 Y11551 (Labrenz, 1999); 5 *Sulfitobacter mediterraneus* (Pukall, 1999); 6 *Sulfitobacter pontiacus* (Sorokin 1995; Ivanova 2004); 7 *Ruegeria gelatinovorans* (Uchino 1998); 8 *Jannaschia helgolandensis* (T) Hel 10 AJ438157 (Wagner-Do¨ bler 2003); 9 *Staleya guttiformis* (T) Ekho Lake_38 Y16427 (Matthias Labrenz 2000); 10 *Sulfitobacter brevis* (T) Ekho Lake_162 Y16425 (Pukall, 1999); 11 *Prochlorococcus marinus* subsp. Pastoris (Rippka R 2000); 12 *Vibrio pomeroyi* (Thompson 2003); 13 *Vibrio splendidus* (Macián 2001); 14 Vibrio lentus (Macián 2001). For four groups of microorganisms, the differences in the metabolism of carbon compounds as revealed by the biology system.

## 2.5 REFERENCES

Donald F. Boesch, R. Eugene Turner. (1984). "Dependence of Fishery Species on Salt Marshes: The Role of Food and Refuge" Estuaries **7**: 460 - 468

González, J. M., Moran, M. A. (1997). "Numerical dominance of a group of marine bacteria in the a-subclass of Proteobacteria in coastal seawater." Appl. Environ. Microbiol **63**: 4237 - 4242.

González, J. M. e. a. (2000). "Bacterial community structure associated with a dimethylsulfoniopropionateproducing North Atlantic algal bloom." Appl. Environ. Microbiol **66**: 4237 - 4246.

W. Soto and J.Gutierrez. (2009). Salinity and temperature effects on physiology responses of *Vibrio fischeri* from diverse ecological niches. Microb Ecol. **57**(1), 140-150.

C. W. Kaspar and M. L. Tamplin. (1993). Effects of Temperature and Salinity on the survival of *Vibrio vulnificus* in seawater and shellfish. Appl. Environ. Microbiol **59**, 2425-2429.

Borrego, J. J., Castro, D., Luque, A., Paillard, C., Maes, P., García, M. T. & Ventosa, A. (1996). *Vibrio tapetis* sp. nov., the causative agent of the brown ring disease affecting cultured clams. Int. J. Syst. Bacteriol **46**, 480-484.

Ivanova, E. P., Gorshkova, N. M., Sawabe, T., Zhukova, N. V., Hayashi, K., Kurilenko, V. V., Alexeeva, Y., Buljan, V., Nicolau, D. V., other authors (2004). "*Sulfitobacter delicatus* sp. nov. and *Sulfitobacter dubius* sp. nov., respectively from a starfish (Stellaster equestris) and sea grass (Zostera marina)." Int. J. Syst. Evol. Microbiol **54**: 475 - 480.

Kiene, R. P., L. J. Linn, J. González, M. A. Moran, and J. A. Bruton (1999). "Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton." Appl. Environ. Microbiol **65**: 4549 - 4558.

Kiene, R. P. a. L. J. L. (2000). "Distribution and turnover of dissolved DMSP and its relationship with bacterial production and dimethylsulfide in the Gulf of Mexico." Limnology and Oceanography **45**: 849 - 861.

Chris Lasher, Glen Dyszynski, Karin Everett, Jennifer Edmonds, Wenying Ye, Wade Sheldon, Shiyao Wang, Samantha B. Joye, Mary Ann Moran, William B. Whitman. (2009). "The Diverse Bacterial Community in Intertidal, Anaerobic Sediments at Sapelo Island, Georgia". Environ. Microbiol **58**: 244 - 261.

Jangid K, W. M., Franzluebbers AJ, Sanderlin JS, Reeves JH, Jenkins MB, Endale DM, Coleman DC, Whitman WB (2008). "Relative impacts of land-use, management intensity and fertilization upon soil microbial community structure in agricultural systems." Soil Biol Biochem **40**: 2843 -2853

Macián, M. C., Ludwig, W., Aznar, R., Grimont, P.D.A., Schleifer, K.H., Garay, E. and Pujalte, M.J. (2001). "*Vibrio lentus* sp. nov., isolated from Mediterranean oysters." Int. J. Syst. Evol. Microbiol **51**: 1449 - 1456.

Matthias Labrenz, B. J. T., Paul A. Lawson, Matthew D. Collins, Peter Schumann and Peter Hirsch (2000). "*Staleya guttiformis* gen. nov., sp. nov. and Sulfitobacter brevis sp. nov., a-3-Proteobacteria from hypersaline, heliothermal and meromictic antarctic Ekho Lake." Int J Syst Bacteriol **49**: 137 - 147.

Rippka R, C. T., Hess W, Lichtlé C, Scanlan D J, Palinska K A, Iteman I, Partensky F, Houmard J, Herdman M (2000). "*Prochlorococcus marinus* Chisholm et al. 1992 subsp. pastoris subsp. nov. strain PCC 9511, the first axenic chlorophyll a2/b2-containing cyanobacterium (Oxyphotobacteria)." Int. J. Syst. Evol. Microbiol **50**: 1833- 47.

Sambrook, J., E. F. Fritsch, and T. Maniatis (2001). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory, New York, 3rd edition.

Rafel Simo (2001). "Production of atmospheric sulfur by oceanic plankton: biogeochemical, ecological and evolutionary links." Trends Ecol Evol **16**: 287 - 294.

Sorokin, D. Y. (1995). "*Sulfitobacter pontiacus* gen. nov., sp. nov. a new heterotrophic bacterium from the Black Sea, specialised on sulfite oxidation." Microbiology (English translation of Mikrobiologiya) **64**: 295 - 305.

Suzuki, M. T., Be´ja`, O., Taylor, L. T. & DeLong, E. F (2001). "Phylogenetic analysis of ribosomal RNA operons from uncultivated coastal marine bacterioplankton." Environ. Microbiol **3**: 323 - 331.

Thompson, F. L., Thompson, C. C., Li, Y., Gomez-Gil, B., Vandenberghe, J., Hoste, B., Swings, J. (2003). "*Vibrio kanaloae* sp. nov, *Vibrio pomeroyi* sp. nov. and *Vibrio chagasii* sp. nov., from sea water and marine animals." Int. J. Syst. Evol. Microbiol **53**: 753 - 759.

Todd, J. D., R. Rogers, Y. G. Li, M. Wexler, P. L. Bond, L. Sun, A. R. J. Curson, G. Malin, M. Steinke, and A. W. B. Johnston (2007). "Structural and regulatory genes required to make the gas dimethyl sulfide in bacteria." Science **315**: 666 - 669.

Uchino, Y., Hirata, A., Yokota, A., Sugiyama, J. (1998). "Reclassification of marine Agrobacterium species: proposals of *Stappia stellulata* gen. nov., comb. nov., *Stappia aggregata* sp. nov., nom. rev., *Ruegeria atlantica* gen. nov., comb. nov., *Ruegeria gelatinovora* comb. nov., *Ruegeria algicola* comb. nov., and *Ahrensia kieliense* gen. nov., sp. nov., nom. rev." J Gen Appl Microbiol **44**: 201 - 210.

Wagner-Do¨ bler, I., Rheims, H., Felske, A., Pukall, R. & Tindall, B. J (2003). "*Jannaschia helgolandensis* gen. nov., sp. nov., a novel abundant member of the marine Roseobacter clade from the North Sea." Int. J. Syst. Evol. Microbiol **53**: 731 - 738.

F. Partensky, W. R. Hess & D. Vaulot (1999). "*Prochlorococcus*, a marine photosynthetic prokaryote of global significance". *Microbiology and Molecular Biology Reviews* **63**: 106–127

# CHAPTER 3

# ANALYSIS OF MICROARRAY DATA FOR RELATIONSHIPS BETWEEN TREATMENTS AND QUALITY OF mRNA

## 3.1 ABSTRACT

The study on bacterioplankton has been considered very important because prokaryotes are essential components of the oceanic food web and catalyze many major biogeochemical processes in the sea (Moran MA 2004). In this study, we focus on the effect of methods to collect bacterioplankton samples from seawater on microarrays. A whole genome microarray of *Silicibacter pomeroyi* DSS3 was used as a standard to indicate the expression of collected seawater sample. *S. pomeroyi* is a member of the marine Roseobacter clade which comprises ~10% to 20% of coastal and oceanic bacterioplankton ((Moran MA 2004; González 1997; González 2000; Suzuki 2001). In this study, two different treatments, "cooling" and "stop solution", were used when sampling for seawater mRNA. Three replicates of each treatment were compared with the control group, which was sampling seawater without any special treatment.

Microarray data were analyzed using bioinformatics methods. Differences in expressed gene composition among the groups were visualized using multidimensional scaling (MDS), and differences in expression level among the groups were visualized with a heat map.

## 3.2 INTRODUCTION

During studies on the ecosystems of coastal seawater of Atlantic Ocean, a member of the marine Roseobacter clade, *Silicibacter pomeroyi*, was isolated, and its genome was sequenced (Moran MA 2004). Microarrays were used to study gene expression in *S. pomeroyi*. In addition, environmental RNA was hybridized against the whole genome microarray for *S. pomeroyi*. But the limitations which may affect the accuracy of microarrays are always an issue. There are three main aspects which may limit the accuracy of microarrays for environmental samples. First, the representativeness of genes targeted by the array depends on the taxonomic composition of sample. Second, the measurement of gene expression depends on the hybridization of RNA with probes. It is possible that some species may be not detectable because the *S. pomeroyi* probes do not hybridize to their mRNA. Third, the accuracy of microarray depends on the condition of the mRNA extracted from the samples. During collection, the quality and quantity of mRNA can change very quickly. In this case, the levels of mRNA will not reflect the environmental condition if they change during sampling. In this study, two sampling treatments were applied in order to find out if they might improve the microarray accuracy. When sampling, “cooling” and “stop solution” treatments were applied and compared to no treatment in order to test if there is an effect on microarray accuracy due to improvements in mRNA perservation.

## 3.3 METHODS

### *3.3.1 Sample collection and microarray data obtained*

Seawater samples were collected at SIMO with a carboy and then pumped through tubing into two filter rigs (3 μm pre-filter and 0.22 μm filter). For each library, 10 L of seawater was filtered, and RNA was collected from the 0.22 μm filter for each array.

Cooling treatment and stop solution treatments were used to test if they can prevent mRNA degradation during sampling. The cooling treatment consisted of inserting an in-line cooling coil in the tube prior to the filters. The cooling coil was chilled with ice water. In addition, the 3 μm filter was placed in an ice bath, and the 0.22 μm filter rig was placed on the top of ice. The stop solution treatment included adding stop solution into the carboy containing the seawater for about one hour prior to filtration. The stop solution (500 ml of 95% ethanol and 5% phenol) was added into the carboy with 10 L of seawater to stabilize mRNA and stop the degradation. Sample collection and RNA extraction were done by Scott Gifford.

### *3.3.2 Microarray data analysis*

Theoretically, mRNA for ribosomal RNA protein synthesis is more stable than mRNA representing other proteins. In this study, mRNA of ribosomal RNA proteins was treated as an internal standard to measure the degradation of other mRNAs. Fluorescence level of every probe was divided by the average fluorescence of mRNA for ribosomal RNA proteins under each treatment.

Every treatment included three arrays as parallel tests. Therefore, including the control group (no treatment used) there should be 9 arrays. Unfortunately, three of them were failed during sampling (two for the stop solution treatment, one for cooling treatment). Therefore, only limited analysis of these data could be performed. In this case, expression data which obtained from the six remaining arrays were treated individually.

The arrays' data were combined into one Excel file. Then, the file was divided into two parts: (1) expression level of mRNA of proteins other than ribosomal RNA proteins; (2) expression level of mRNA of ribosomal RNA proteins. The average expression level of mRNA of ribosomal RNA proteins under different treatments and control were calculated. Then every value of mRNA expression under a particular treatment was divided by the average of the three replicate mRNA of ribosomal RNA proteins expression values under that treatment. Results were saved in a .txt file as the input file of R (http://www.r-project.org/).

A Multidimensional scaling (MDS) method was applied to calculate the relatedness among the six arrays. In this study, R (http://www.r-project.org/) was used to execute the MDS analysis. For each sample, every pairwise comparison made for each gene to rate similarity. The code is attached in the supplementary material.

A heat map was programmed to present a two-color visual comparison of each gene expression level under the six conditions. The heat map is a 2D color representation of the data, which shows the relationships of the data. For the data set, log values were calculated for every calculation of normalized mRNA expression, and the mean of the control group (Normal1, 2, 3) fluorescence level of each probe was calculated. Log ratios were normalized by subtracting the mean of the Normal group. Therefore, if the expression level of genes were upregulated, the values were larger than 0. In

this study, R (http://www.r-project.org/) was used to program the heat map analysis. Two colors were defined: green, if the value was smaller than 0; red, if the value was larger than 0.

## 3.4 RESULTS AND DISCUSSION

### *3.4.1 Cluster and MDS*

Microarray data was obtained from Scott Gifford. During the cooling treatments, two sampling failed. Therefore, only one microarray was preformed for the cooling treatment. For the stop solution treatment, one sample was also failed, and only two microarrays were performed from the "stop solution" treatment. Averages could not be calculated because of the missing replicates. Therefore, each dataset has to be treated as an individual. Differences in expression levels among the groups were visualized using MDS of the similarity matrix to produce a two-dimensional figure (Fig 3.1). From this figure we could not identify groups that segregated based on treatment. Furthermore, a hierarchical cluster analysis was applied based on the distance matrix of each group (Fig 3.2), which indicated that the effect of cooling and stop solution treatments were different compared to the normal one. However, the analysis did not separate one normal group and one stop solution group.

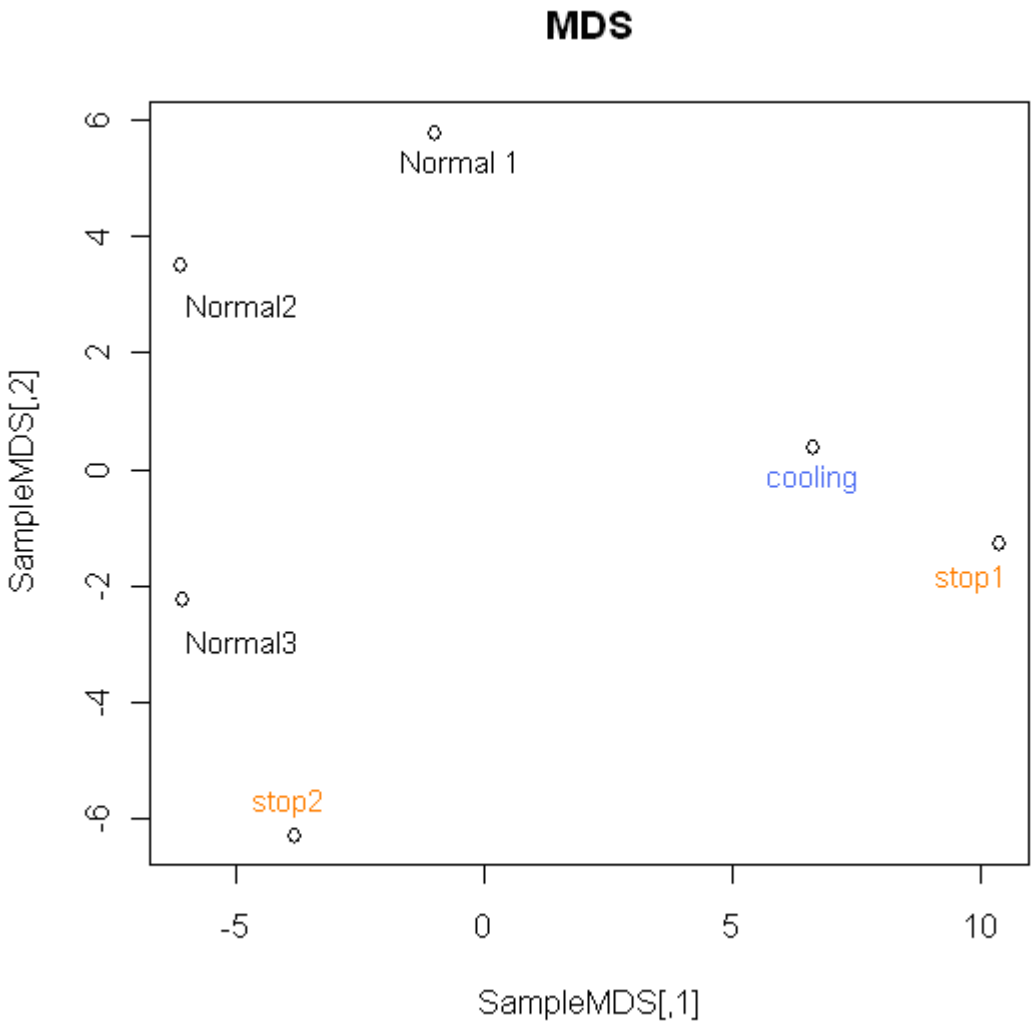



Fig 3.1 MDS plot of six data sets. Normal 1, 2, and 3: no treatment; Stop 1 and 2: stop solution treatment; Cooling: cooling treatment.

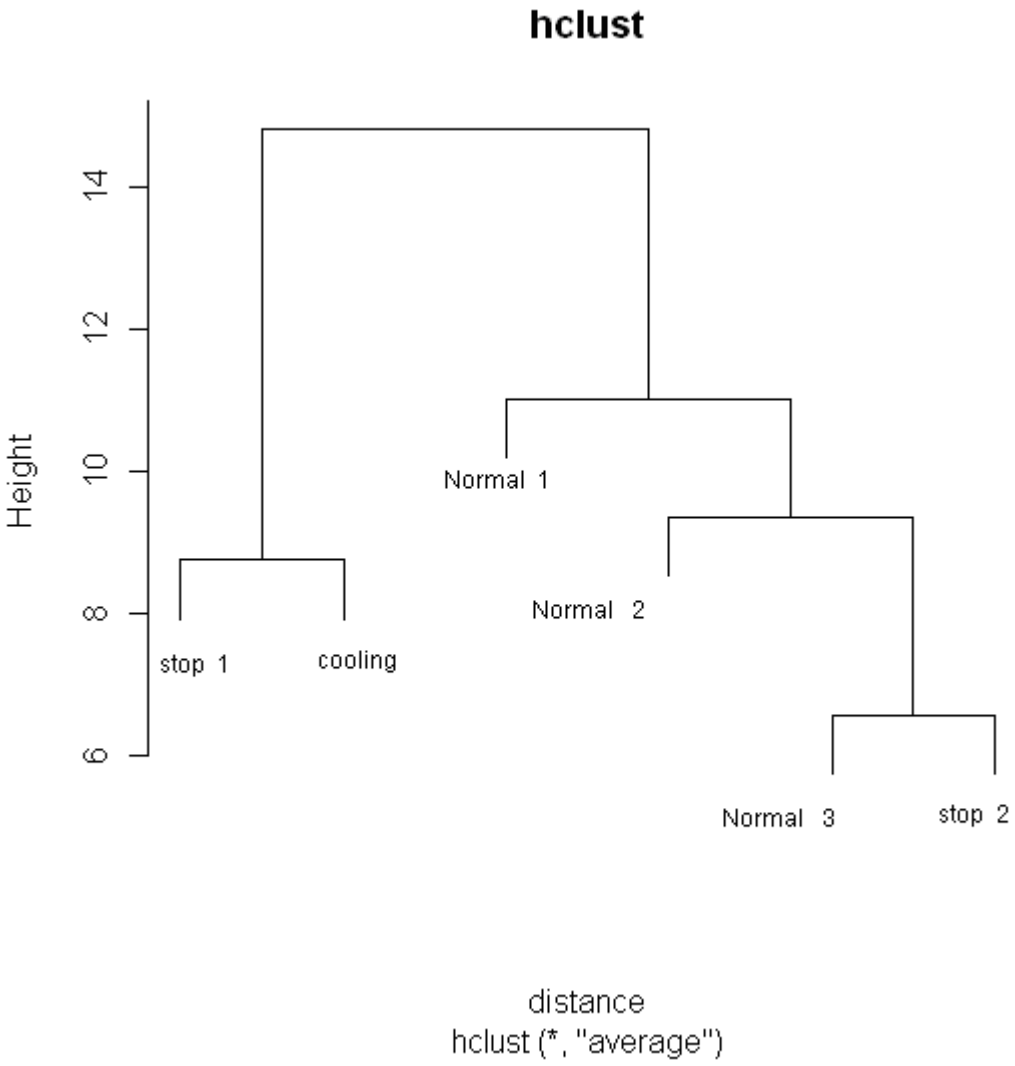



Fig 3.2 Hierarchical cluster analysis of gene expression level under stop solution treatment (“stop 1” and “stop 2”), cooling treatment (“cooling”), and no treatment (“Normal 1”, “Normal 2”, and “Normal 3”). The heights show the distance among these six groups based on the distance matrix obtained from non-normalized fluorescence values.

*3.4.2 Heat map*

The gene expression levels for a particular normal array might not representative. Therefore, the normalization was also applied by calculating the mean of the normal group and comparing the expression values under each treatment to this mean. The result was visualized using a heat map (Fig. 3.3): when the value is much larger than the mean, it will be given a much brighter red color; when the value is much smaller than the mean, it will be given a much brighter green color; and the values in between will be given a color in the red-green gradient depending on the difference between that value and the mean.

After the normalization, we can find that although the "cooling" group was still clustered with the normal group, the three normal arrays tended to cluster together first, and cluster with the single cooling array later from the heat map. However, it does not change the result of hcluster, that is, one of the stop solution treatment group still clustered with cooling group and the other stop solution group clustered with normal groups (Fig 3.4). These results lead to the conclusion that, the stop solution treatment may have a slight and not significant effect on the mRNA quality when compared to the group without treatment. And the cooling treatment may not have an effect on the mRNA quality when compare to the normal groups.

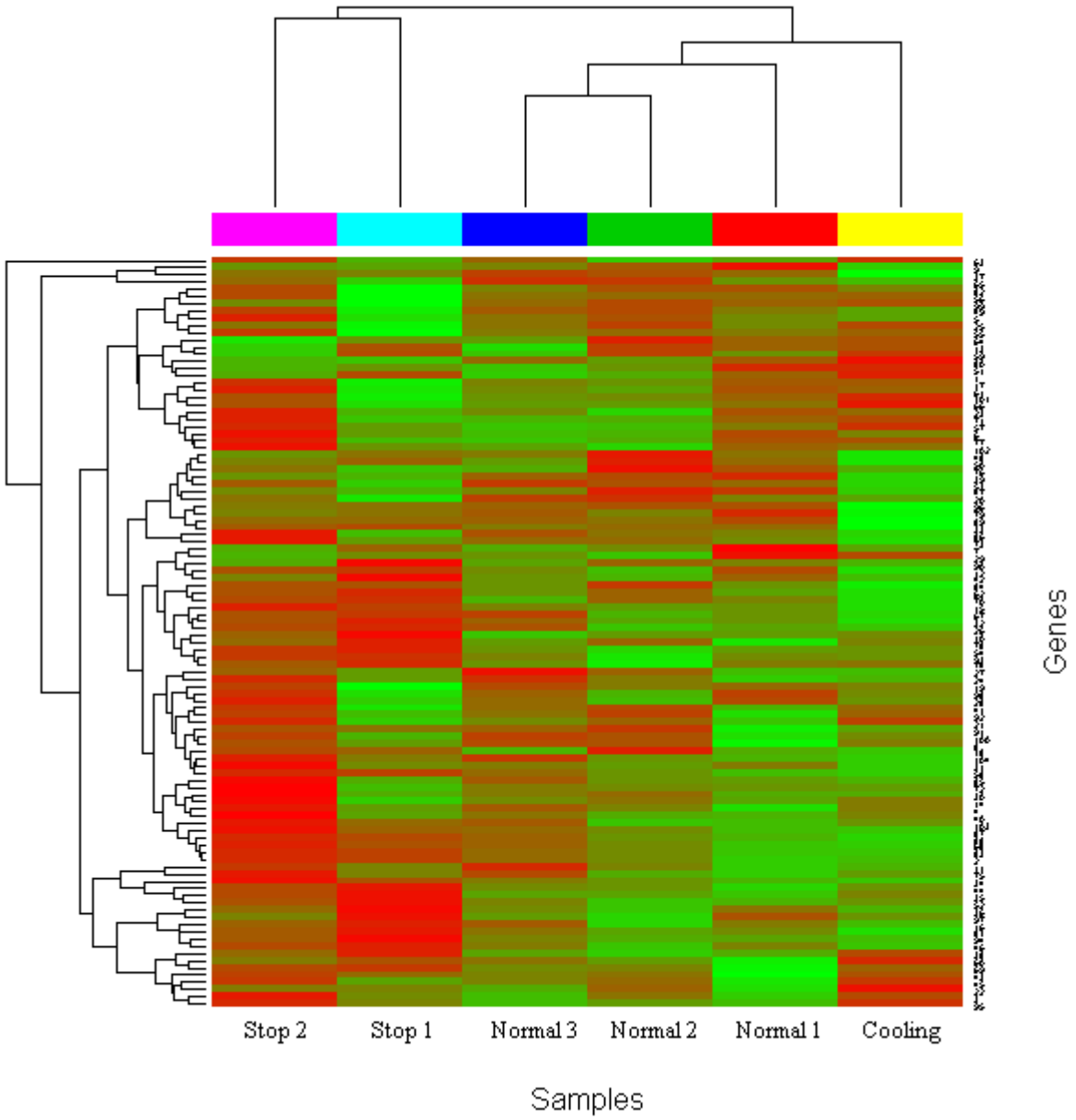



Fig 3.3 Heat map of gene expression levels. Arrays 1, 2, and 3 are the normal treatment; arrays 4 and 5 are the stop solution treatment; array 6 is the cooling treatment. Fluorescence levels were normalized as described in the text. A red color means the expression level is higher compared to the mean of Normal group. A green color means the expression level is lower compared to the mean of Normal group.

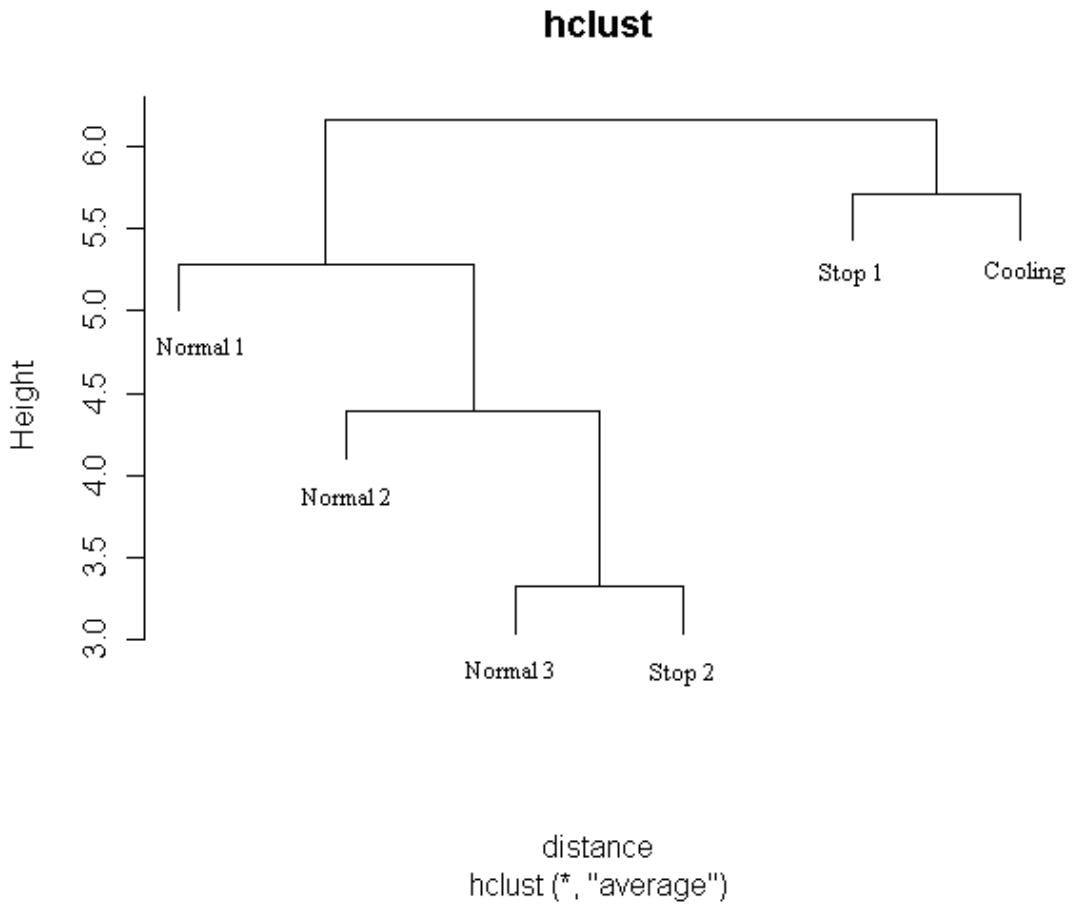



Fig 3.4 Hierarchical cluster analysis of gene expression level under stop solution treatment (“stop 1” and “stop 2”), cooling treatment (“cooling”), and no treatment (“Normal 1”, “Normal 2”, and “Normal 3”). The heights show the distance among these six groups based on the distance matrix obtained from normalized fluorescence values.

3.5 REFERENCES

Moran MA, B. A., González JM, Heidelberg JF, Whitman WB, Kiene RP, Henriksen JR, King GM, Belas R, Fuqua C, Brinkac L, Lewis M, Johri S, Weaver B, Pai G, Eisen JA, Rahe E, Sheldon WM, Ye W, Miller TR, Carlton J, Rasko DA, Paulsen IT, Ren Q, Daugherty SC, Deboy RT, Dodson RJ, Durkin AS, Madupu R, Nelson WC, Sullivan SA, Rosovitz MJ, Haft DH, Selengut J, Ward N (2004). "Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment." Nature **432**(7019): 910 - 913.

Pomeroy LR, W. R. (1981). "Ecology of a salt marsh." Springer-Verlag.

Pukall, R., Buntefuß, D., Fru¨ hling, A., Rohde, M., Kroppenstedt, R. M., Burghardt, J., Lebaron, P., Bernard, L., Stackebrandt, E. (1999). "*Sulfitobacter mediterraneus* sp. nov., a new sulfite-oxidizing member of the a-Proteobacteria." Int J Syst Bacteriol **49**: 513 - 519.

Thompson JD, H. D., Gibson TJ (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res **22**: 4673 - 4680.

Wagner-Do¨ bler, I., Rheims, H., Felske, A., Pukall, R. & Tindall, B. J (2003). "Jannaschia helgolandensis gen. nov., sp. nov., a novel abundant member of the marine Roseobacter clade from the North Sea." Int J Syst Bacteriol **53**: 731 - 738.

Wang ZA, C. W. (2004). "Carbon dioxide degassing and inorganic carbon export from a marsh-dominated estuary (the Duplin River): a marsh CO2 pump." Limnol Oceanogr **49**: 341 - 354.

Kiene, R. P., Linn, L. J., Bruton, J. A (2000). "New and important roles for DMSP in marine microbial communities." J Sea Res 43: 209 - 224.

# CHAPTER 4

# A GENETIC STUDY OF DMSP METABOLISM IN *Silicibacter pomeroyi DSS-3*

## 4.1 ABSTRACT

*Silicibacter pomeroyi*, a member of a marine roseobacter clade, is a model system for the study of dimethylsulfoniopropionate (DMSP) metabolism. The degradation of DMSP occurs through two competing pathways, known as the cleavage and the demethylation pathways. The cleavage pathway results in the formation of dimethylsulfide (DMS), the major natural source of sulfur to the atmosphere (Bates 1987; SIMO 2001).

Bioinformatic approaches were used to identify candidate genes for enzymes involved in DMSP degradation. A genetic system was applied to test the function of these genes. Studies included: identification of minimum inhibitory concentrations of antibiotics and antibiotic resistance cassettes functional in *S. pomeroyi*; transformation by electroporation with DNA protected by SssI methylase; and construction of gene replacement mutants using suicide vectors. This system was initially constructed by James R. Henriksen and was revised in this study in order to improve the efficiency. The growth of mutants with different carbon sources was also measured.

## 4.2 INTRODUCTION

Dimethylsulfoniopropionate (DMSP) is a ubiquitous compound in the ocean and a major source of carbon and reduced sulfur for marine bacteria. Its metabolism is important in the carbon and sulfur cycles and influences global climate due to the degradation product dimethyl sulfide (DMS). In the atmosphere, DMS is oxidized to methanesulfonic acid and eventually to particulate sulfate, which then acts as hygroscopic cloud condensation nuclei. Concentrations of DMS in the atmosphere correlate with the cloud condensation nuclei loads (Bates 1987). Increased solar radiation leads to an increase in the abundance of algae that produce DMSP, which could in turn lead to an increase in the amount of DMS formed and released into the atmosphere, which would lead to increased cloud formation and albedo. As an increase in cloud cover would then decrease the abundance of algae, this system would result in a negative feedback loop, with DMS acting as an anti-greenhouse gas. The majority of DMSP degradation is through the demethylation pathway, but the cleavage pathway is also important (Fig 4.1) (Kiene 1999; Kiene 2000).

*Silicibacter pomeroyi*, a member of the marine Roseobacter clade, is a model system for the study of DMSP degradation. *S. pomeroyi* can cleave DMSP to DMS and carry out demethylation to methanethiol (MeSH) as well as degrade MeSH. To study the genetics of this pathway on the gene level, several candidate genes which may be involved in DMSP degradation were selected. The role of genes SPO1703, SPOA268, SPOA269, and SPO2134 were then tested by mutagenesis.

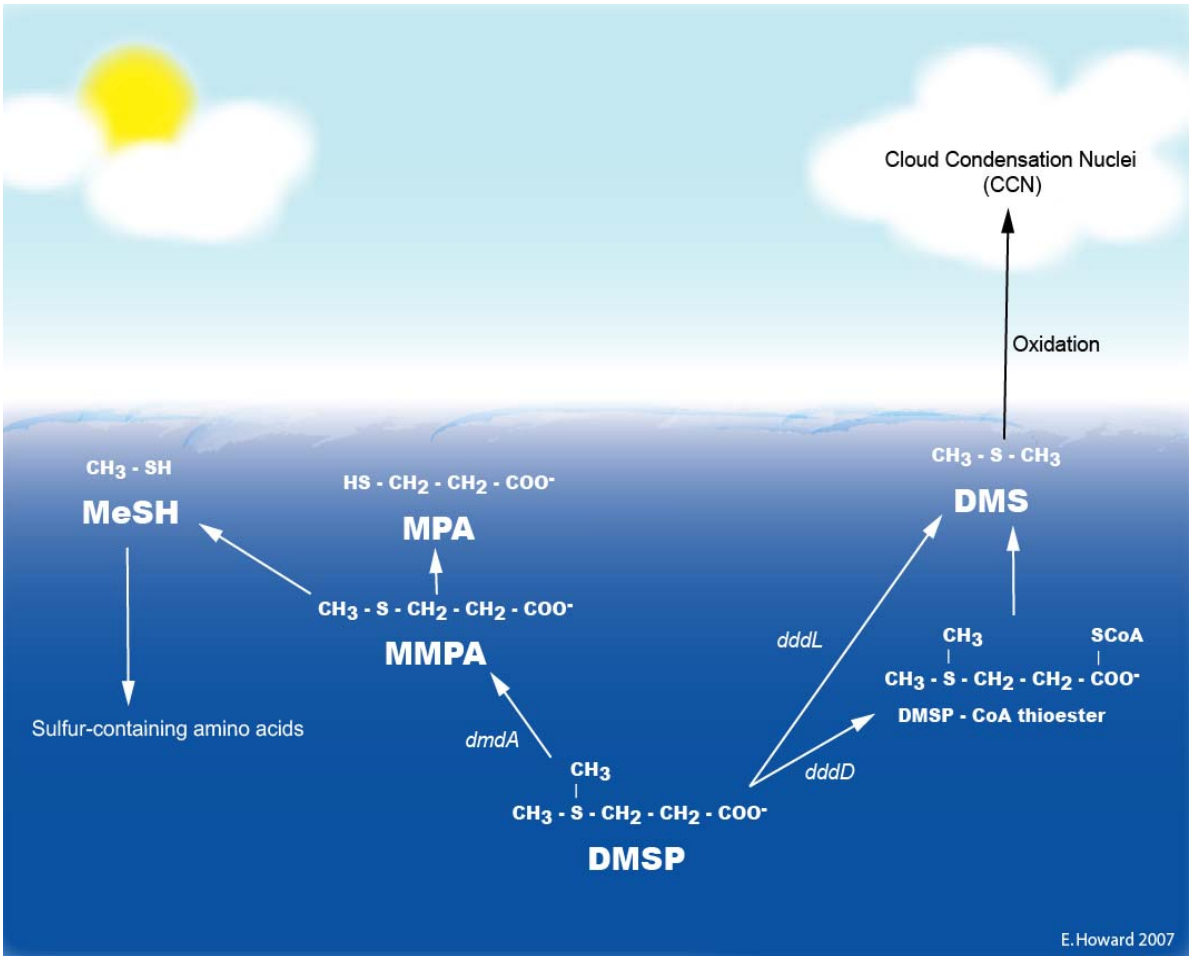



Fig 4.1 DMSP degradation pathways. Modified from Erinn Howard.

## 4.3 MATERIALS AND METHODS

### *4.3.1 Bioinformatic analysis*

To identify candidate genes which play a role in DMSP metabolism, microarray and proteomics data were analyzed. BLAST was also applied to find homologous to proposed genes in this pathway.

*BLAST*

BLASTp was used with the proposed genes (Todd 2007) involved in DMSP degradation pathway in other organisms as the query sequence against the *S. pomeroyi* genome sequence.

*MEGA*

MEGA4 was used to generate trees of the homologous genes from an evolutionary perspective.

At the same time, data sets obtained from previous microarray (Scott Gifford, Dr. Moran's lab; Christopher Reisch, Dr. Whitman's lab) and proteomic (James R. Henriksen, Dr. Whitman's lab) studies on DMSP degradation in *S. pomeroyi* were integrated with the result of BLAST (Table 4.1).

### *4.3.2 Insertional inactivation of target genes*

CONSTRUCTION OF RECOMBINANT PLASMID THROUGH SLIC

A shuttle plasmid with regions of up- and down-stream sequence of SPO1703 flanking the $tet^R$ genes from pRK415 was constructed by Simple Ligation Independent Cloning (SLIC; Elledge 2007).

The upstream region was amplified with primers SPO1703_UF HindIII and SPO1703_UR SpeI, which included ~20 bp overlap of the vector pCR2.1 region and $tet^r$ region and also included HindIII

and SpeI digestion sites for later screening. The *tetR* region of pRK415 was PCR amplified with primers TetR_F SpeI and TetR_R XhoI, which included SpeI and XhoI digestion sites for later screening. The downstream region was amplified with SPO1703_Df XhoI and SPO1703_DR XbaI, which included ~20 bp overlap of vector pCR2.1 region and $tet^r$ region and also included XhoI and XbaI digestion sites for later screening. (Fig 1.2) Vector pCR2.1 was cut with HindIII and XbaI, as well as EcoRI to reduce the background. After T4 DNA polymerase treatment, the overlapping single strands were annealed, and the reaction transformed into chemically competent DH5_α *E. coli* via heat shock (Sambrook 2001; Fig 4.2). The strain with recombinant plasmid pSLIC1703 was maintained in 35% glycerol and frozen at -80 °C.

*4.3.3 Mutagenesis following electroporation and recombination*

ELECTROCOMPETENT *S. pomeroyi*

A 2 L *S. pomeroyi* culture was grown at 30 °C with shaking in 1/2 YTSS medium separately in two flasks (size 2 L) to mid-log growth stage (0.15–0.20 OD600) and chilled at 4 °C for one hour. The cells were harvested by centrifugation at 8,000 rpm for 10 min, the supernatant was decanted, and the cells were washed three times with 250 mL, 150 mL and 100 mL of 4 °C sterile 10 % glycerol (diluted with water). Then the pellet was resuspended in 2 mL of 10 % glycerol as mentioned above. After the final wash, 110 µL of the cell suspension were transferred into microcentrifuge tubes, frozen immediately with dry ice and ethanol, and then stored at -80 °C.

Table 4.1 Candidate genes for mutagenesis analysis.

| Gene ID | Product | 1 | 2 | 3 | Predicted Function |
|---|---|---|---|---|---|
| SPO1703 | CaiB/BaiF family | +/- | - | + | cleavage pathway |
| SPOA0268 | transcriptional regulator | - | + | - | None |
| SPOA0269 | conserved hypothetical | - | + | - | None |
| SPO2134 | trimethylamine methyltransferase family | + | - | - | DMSP demethylation |
| SPO2299 | adenylosuccinate lyase | - | - | + | DMSP cleavage |
| SPO0207 | aminomethyl transferase family | + | - | - | DMSP-->MMPA |
| SPO1648 | aminomethyl transferase family | + | - | - | DMSP-->MMPA |
| SPOA0318 | methionine gamma-lyase | + | - | - | MeSH-->methionine |
| SPO0633 | gamma-glutamyltranspeptidase | - | + | - | Detoxifying |
| SPO2203 | methylmalonate-semialdehyde dehydrogenase | - | + | + | cleavage pathway |
| SPO3016 | methylenetetrahydrofolate reductase | + | - | - | methyl group transfer |
| SPO1914 | alcohol dehydrogenase | + | + | + | DMSP demethylation |
| SPO3850 | glutathione-dependent formaldehyde dehydrogenase | - | + | - | Detoxifying |
| SPOA0073 | arsenate reductase | + | - | - | DMSP uptake regulation |

1: Microarry result. “+” means probes were up-regulated during growth with DMSP. “–” means probes were not up-regulated during growth with DMSP. “+/-” means only one of two probes was up-regulated during growth with DMPS.

2: Proteomic result. “+” means proteins were up-regulated during growth with DMSP. “–” means the proteins were not up-regulated during growth with DMSP.

3: A homolog to genes which are proposed to have a function in DMSP degradation pathway in other species was found in the *S. pomeroyi* genome. “+” means a homologous gene was found. “–” means a homologous gene was not found.

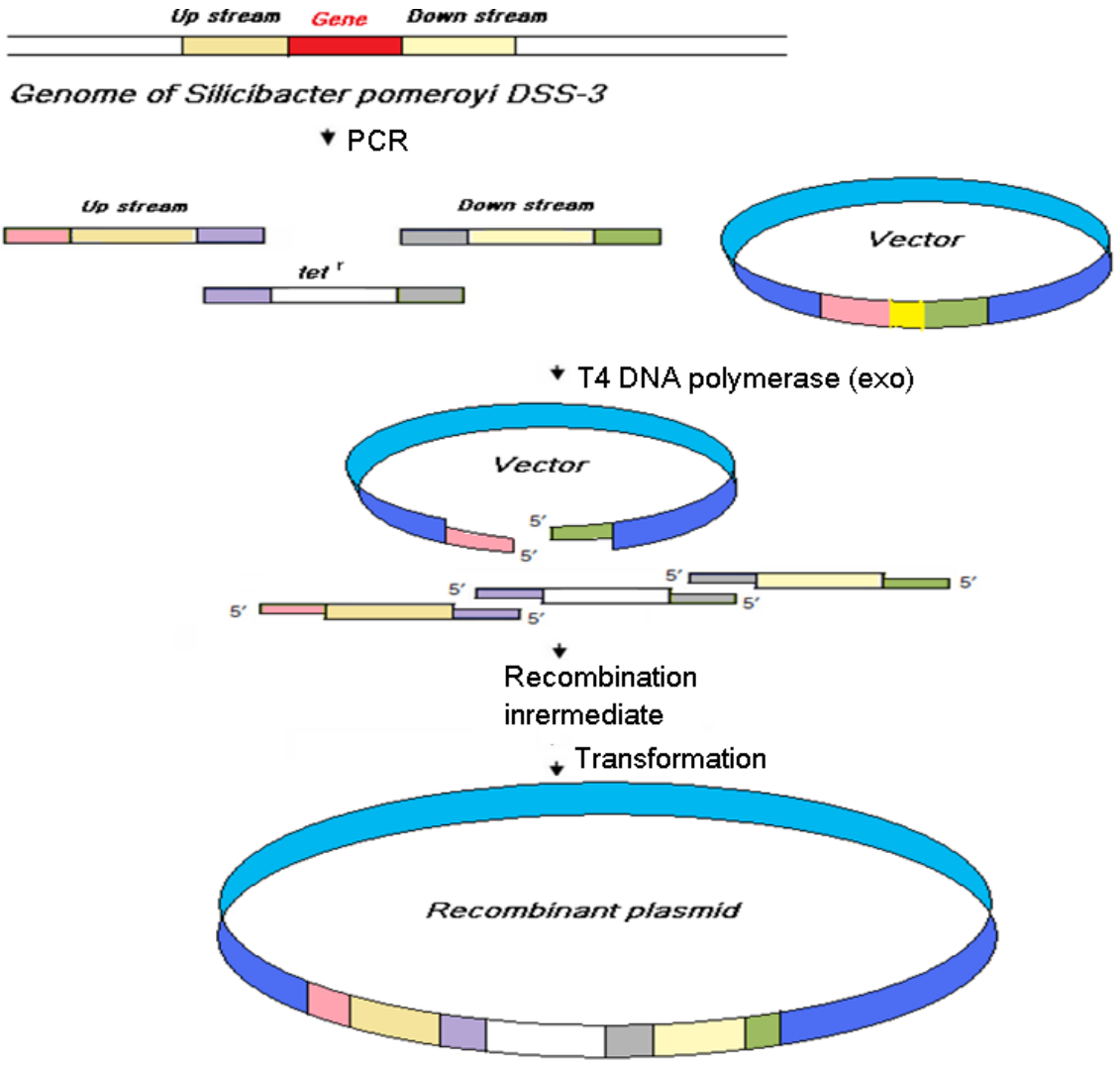



Fig 4.2 Recombinant plasmid construction by assembling three fragments using SLIC. An illustration of the four-way SLIC reaction: annealing with T4 DNA polymerase-treated fragments and the linear vector, pCR2.1.

Table 4.2 Oligonucleotides used in this study

| | Name | Sequence 5' to 3' |
|---|---|---|
| Recombinant Plasmid pSYW2x1703 Construction | TetR_F SpeI | ACTAGTACCGTATTACCGCCTTTGAGTGAG |
| | TetR_R XhoI | CTCGAGACGCTGAGTGCGCTTCAAATCATC |
| | SPO1703_Df XhoI | AGCGCACTCAGCGTCTCGAGTATGAAACAGGATCAACCCGGACC |
| | SPO1703_DR XbaI | GGGCGAATTGGGCCCTCTAGAAGCGAGAATTTCAATCTCCGGCAC |
| | SPO1703_UF HindIII | GACCATGATTACGCCAAGCTTATTTGAGGCGAGCTTTCCAACTGC |
| | SPO1703_UR SpeI | GGCGGTAATACGGTACTAGTATTTCTGCCAATGCGTTCCATCCG |
| Mutants Verification | TetR_F SpeI | ACTAGTACCGTATTACCGCCTTTGAGTGAG |
| | TetR_R XhoI | CTCGAGACGCTGAGTGCGCTTCAAATCATC |
| | Check_SPO1703_InterGene_F | AATCGCCCGCTTTCCCATATTCG |
| | Check_SPO1703_InterGene_R | TGCAAAGAACATCCAGCCATCAGC |
| | Check_SPO1703 tet_down_F | AAACCCTGTTGCATGGAATCTCCG |
| | Check_SPO1703 tet_down_R | TCAGGAAATTGAGGCCGTTCAAGG |
| | Check_SPO1703 up_tet_F | TTCGCCAATCCATCGACAATCACC |
| | Check_SPO1703 up_tet_R | TTTGACGGCATCCAAGGTTTCGG |
| | Check_SPOA269_InterGene_F | AAGAGAATTTGGAGCTTTGGCCGC |
| | Check_SPOA269_InterGene_R | TGCCATCCATGAAGCTGTTGATCC |
| | Check_SPO2134_InterGene_F | ACGATTTATGCTTGTTGCGCAGGC |
| | Check_SPO2134_InterGene_R | GCAAGTGAATGTTGAACTTGGCCC |

## RESTRICTION PROTECTION

SssI CpG methylase (New England Biolabs) was used to methylate 1 µg of DNA according to the manufacturer's directions (1 µl of SAM + 3 µl of SssI for 60 µl total volume). Protection was checked by digestion with BtuI, a CpG methylation sensitive restriction endonuclease. After methylation, DNA was isolated using a DNA Clean and Concentrator kit (Zymo Research), dissolved in 10 µl sterilized de-ionized water and immediately used for electroporation.

Shuttle vector pRK415 (1µg) was methylated and isolated through the methods mentioned above. Then vector pRK415 was diluted 10 times with sterilized de-ionized water and stored separately in 10 microcentrifuge tubes at -80 °C.

## ELECTROPORATION OF *S. pomeroyi*

Electrocompetent *S. pomeroyi* cells were thawed on ice, mixed with DNA, and immediately electroporated in a 1 cm gap cuvette using a GenePulser II Electroporation system with a PLUS Pulse controller (Bio-Rad) and the following settings: 1.45 kV, 25 µF capacitance, and 186 Ω resistance in parallel. Time constants were 6.7–6.9 ms. The electroporated cells were suspended immediately in 1 ml of 1/2 YTSS at room temperature and transferred to a sterile glass test tube. Cultures were incubated at 30 °C with shaking and allowed to double (usually within 2–3 hours) before plating on 1/2 YTSS with 20 µg $mL^{-1}$ tetracycline. Typically, the entire electroporation mixture was plated on 20 plates. SssI methylated pSLIC1703 was electroporated into electrocompetent *S. pomeroyi* DSS-3 and

plated on 1/2 YTSS with 20 μg $mL^{-1}$ tetracycline to generate *S. pomeroyi* Δ1703. Colonies were picked into 5 ml of 1/2 YTSS broth with 20 μg $mL^{-1}$ tetracycline. After growth, they were screened by PCR. Mutants which showed the expected results from PCR were maintained and stored in -80 $^{o}$C freezer with 0.5 ml of 1/2 YTSS medium plus 20 μg $mL^{-1}$ tetracycline and 0.5 ml of 70% glycerol in each cryogenic tube.

*4.3.4 Confirmation of gene replacement*

PCR amplifications of genomic DNA extracts with primers that spanned the junction of the flanking region, the antibiotic cassette (TetR_F SpeI and TetR_R XhoI) and the gene SPO1703 (Check_SPO1703_F and Check_SPO1703_R) were performed. In order to distinguish the double and single recombinants, primers upstream of the up region and the internal to the $tet^{r}$ region (Check_SPO1703 up_tet_F and Check_SPO1703 up_tet_R), and internal to the $tet^{r}$ region and downstream of down region (Check_SPO1703 tet_down_F and Check_SPO1703 tet_down_R) were designed.

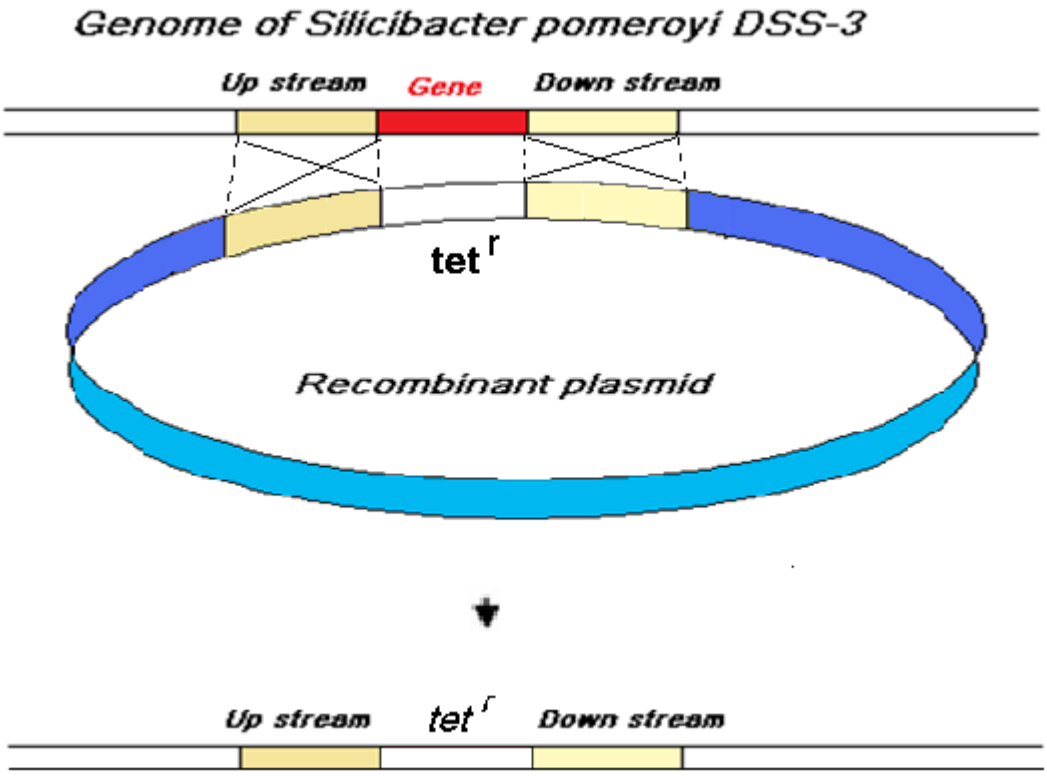



Fig 4.3 Illustration of double crossing over during electroporation.

*4.3.5 Characterization of gene replacement mutants*

GROWTH CURVE TEST

Both the wild type and mutant were grown to mid-log growth stage (0.15–0.20 $OD_{600nm}$) in ½ YTSS. The cells were diluted 1:1000 with Basal medium (MBM) and transferred into test tubes and cultured at 30 °C. Eight kinds of carbon sources were tested during the growth experiments (Table 4.3).

## 4.4 RESULTS AND DISCUSSION

*4.4.1 Recombinant plasmid construction and verification*

A recombinant plasmid with the up-stream and down-stream regions flanking gene SPO1703 was constructed through SLIC (Elledge 2007; Fig 4.2 and Fig 4.4).

*E. coli* colonies were selected from LB plates with 20 μg $mL^{-1}$ tetracycline and 20 μg $mL^{-1}$ ampicillin (the vector pCR2.1 contained *amp*$^r$ gene). Plasmids were purified and verified through digestion with restriction enzymes XbaI, SpeI, HindIII, and XhoI (Fig 4.5).

*4.4.2 Transformation of S. pomeroyi*

In a previous study, an initial *S. pomeroyi* electroporation protocol was generated by James R. Henriksen. In order to improve the protocol, the shuttle vector pRK415, which includes a *tet*$^r$ region, was used to test the efficiency of electroporation. Colonies were selected which grew on plates with tetracycline. The relationship between recovery time after electroporation and efficiency was

determined (Fig 46). From these experiments, the optimal time for recovery was 2~3 hours post electroporation.

The tetracycline resistance gene was amplified from the mutant ΔSPO1703 (Fig 47). The mutant was then shown not to result from a single cross over event (Fig 4.8), since the PCR product of a single cross over would contain the sequence of the vector, $m_r$ = 3.9Kb. In addition, a wild type gene was not present in these cells (Fig 4.9). Therefore, using this technique, replacements of many genes of interest can be undertaken in *S. pomeroyi*.

*4.4.3 Construction of mutants in S. pomeroyi*

The same method was applied to construct the deletion mutants of other candidate genes in *S. pomeroyi.* The SLIC recombinant plasmids which contained the upstream and downstream regions of candidate genes SPOA0268, SPOA0269, and SPO2134 were obtained from Christopher Reisch and James R. Henriksen. After electroporation, colonies which had tetracycline resistance were cultured, and verified by PCR amplification of the tetracycline resistance genes and the candidate genes. As a result, all the mutants contained the tetracycline resistance gene (Fig 4.10 A, C, E), and showed an absence of the candidate genes (Fig 4.10 B, D, F).

All the mutants were maintained in 35% glycerol (0.5 ml 1/2 YTSS medium and 0.5 ml 70% glycerol) and frozen at -80 °C.

Table 4.3 Composition of carbon sources which are used to test the growth curve.

| Culture Medium | Abbreviation | Carbon sources concentration (mm) |
|---|---|---|
| MBM+Glucose | Glu | 2 |
| MBM+Lactate | Lac | 5 |
| MBM+DMSP | DMSP | 5 |
| MBM+MMPA | MMPA | 5 |
| MBM+Acrylate | Acr | 5 |
| MBM+Propionate | Pro | 5 |
| MBM+Acetate | Ace | 10 |
| MBM+Methylmalonate | Meth | 5 |
| MBM [1] | MBM | - |
| 1/2 YTSS | 1/2 YTSS | - |

1: The composition of basal medium (MBM) is in the supplementary material.

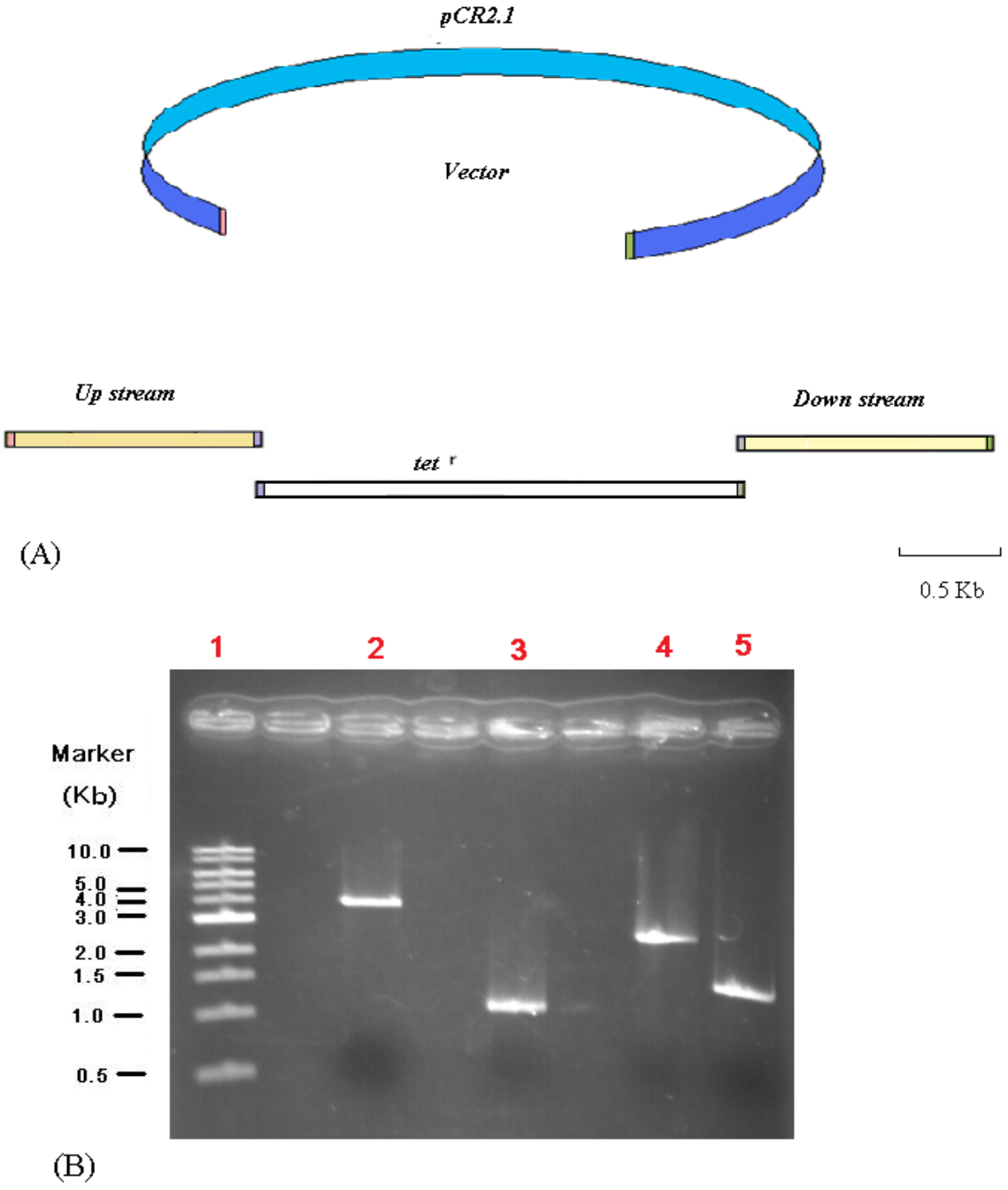
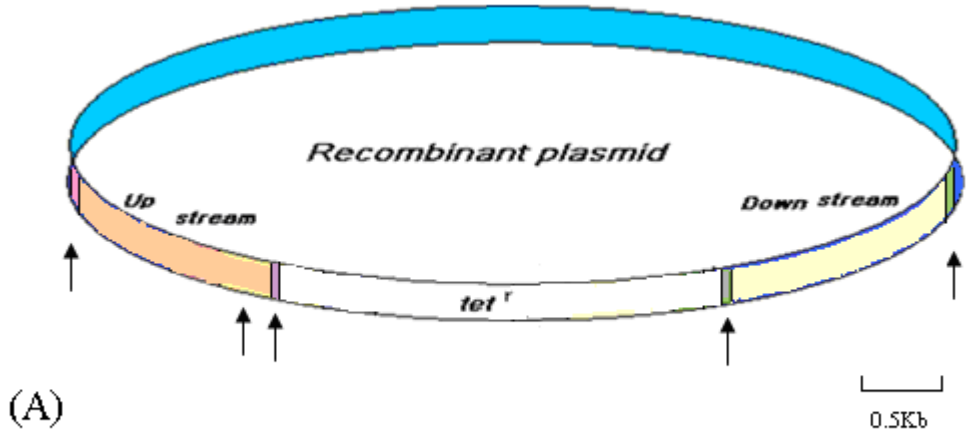
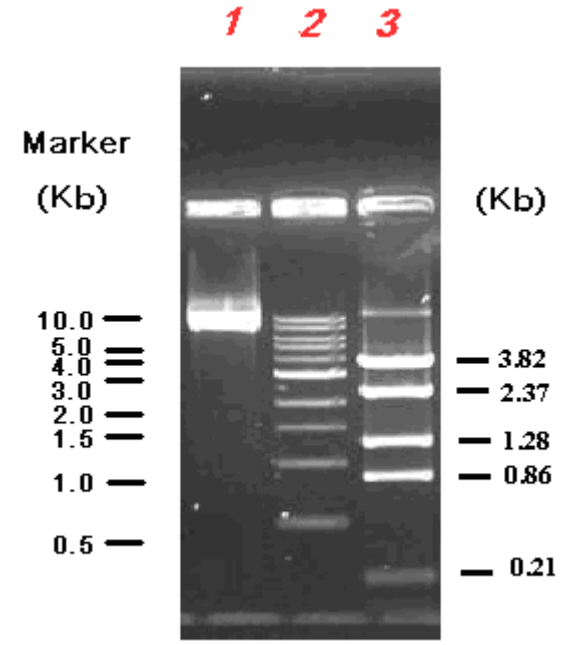



Fig 4.4 Illustration of construction of pSYW2x1703. (A) Illustration of recombinant vector construction. (B) PCR amplified and gel purified vector pCR2.1, which was digested with HidIII, XbaI, and EcoRI (lane 2), PCR products of upstream region ($m_r$ 1.1Kb, lane 3), *tet*$^r$ region ($m_r$ 2.3Kb, lane 4) and downstream region ($m_r$ 1.3 Kb, lane 5).




(A)




(B)

Fig 4.5 Verification of construction of the recombinant plasmid pSYW2x1703. (A) The plasmid was verified by digested with restriction enzymes: HindIII, XhoI, XbaI, and SpeI. "↑" shows the restriction digestion sites. (B) Gel photo of the recombinant plasmid ($m_r$ 8.6Kb, lane 1) constructed through SLIC, marker (lane 2), and fragments after digestion.

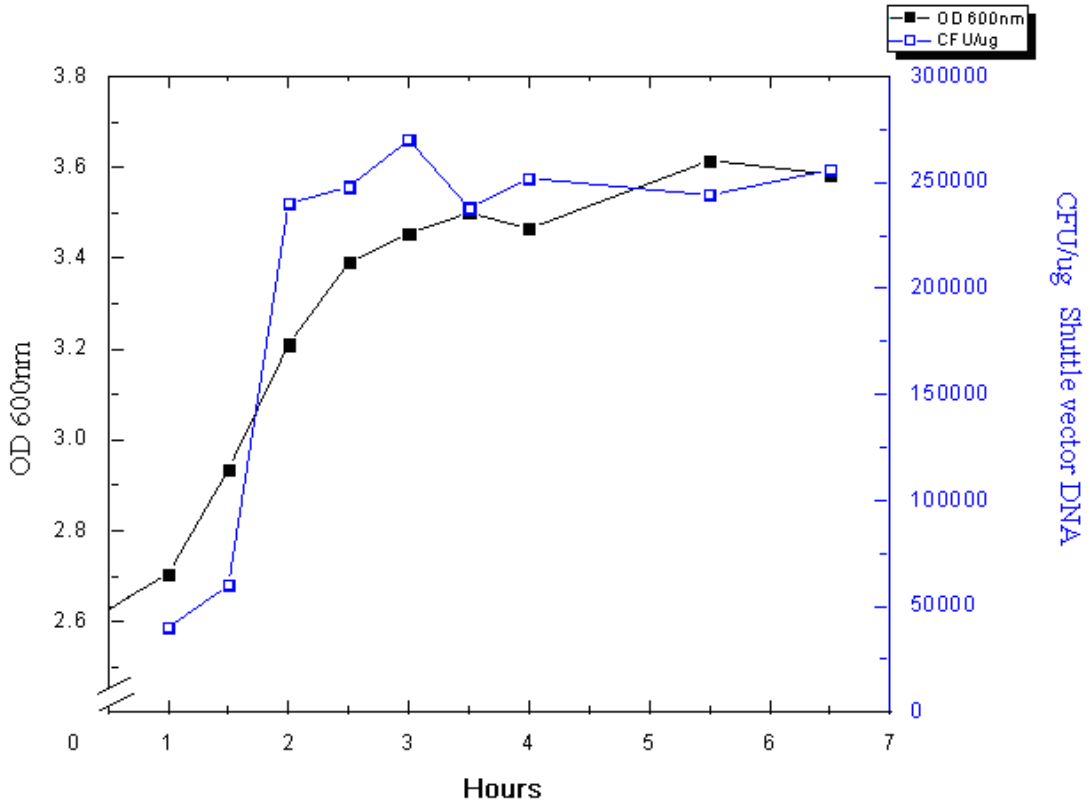



Fig 4.6 Shuttle vector pRK415 was used to test recovery after electroporation. The relationship between growth (increase in absorbance at 600 nm; ■) and *tet*$^r$ CFU/ug of shuttle vector pRK415 (□). Absorbance at 600 nm was measured after 10-fold dilution in 1/2 YTSS.

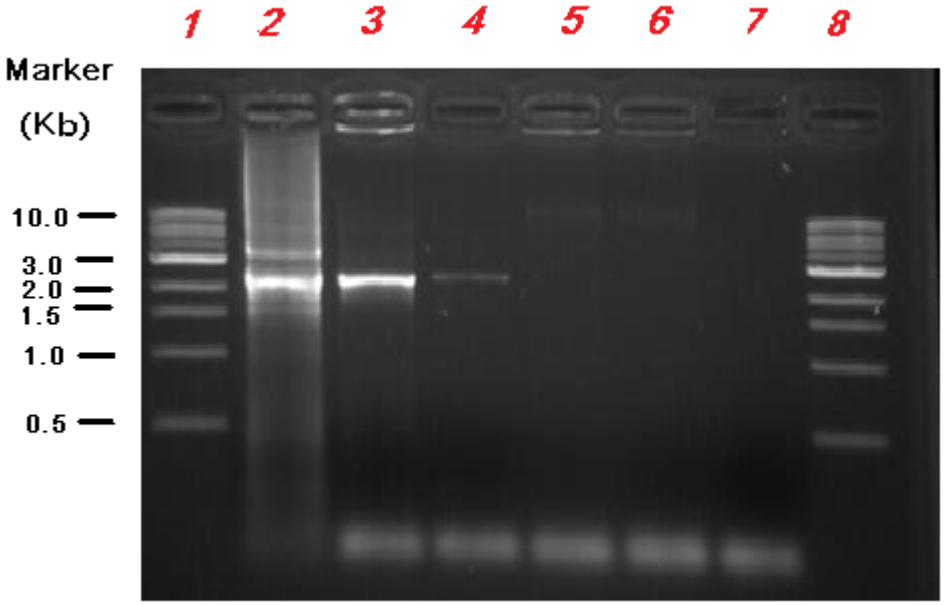



Fig 4.7 Screening of strain ΔSPO1703 by amplifying the *tet*$^{r}$ gene of colonies which grew on tetracycline plates. Lane 1 and 8 are markers. Lane 2 is the positive control of the *tet*$^{r}$ gene. Lanes 3~6 are amplifications of the *tet*$^{r}$ gene from four colonies which grew on the plates. Lane 7 is the negative control

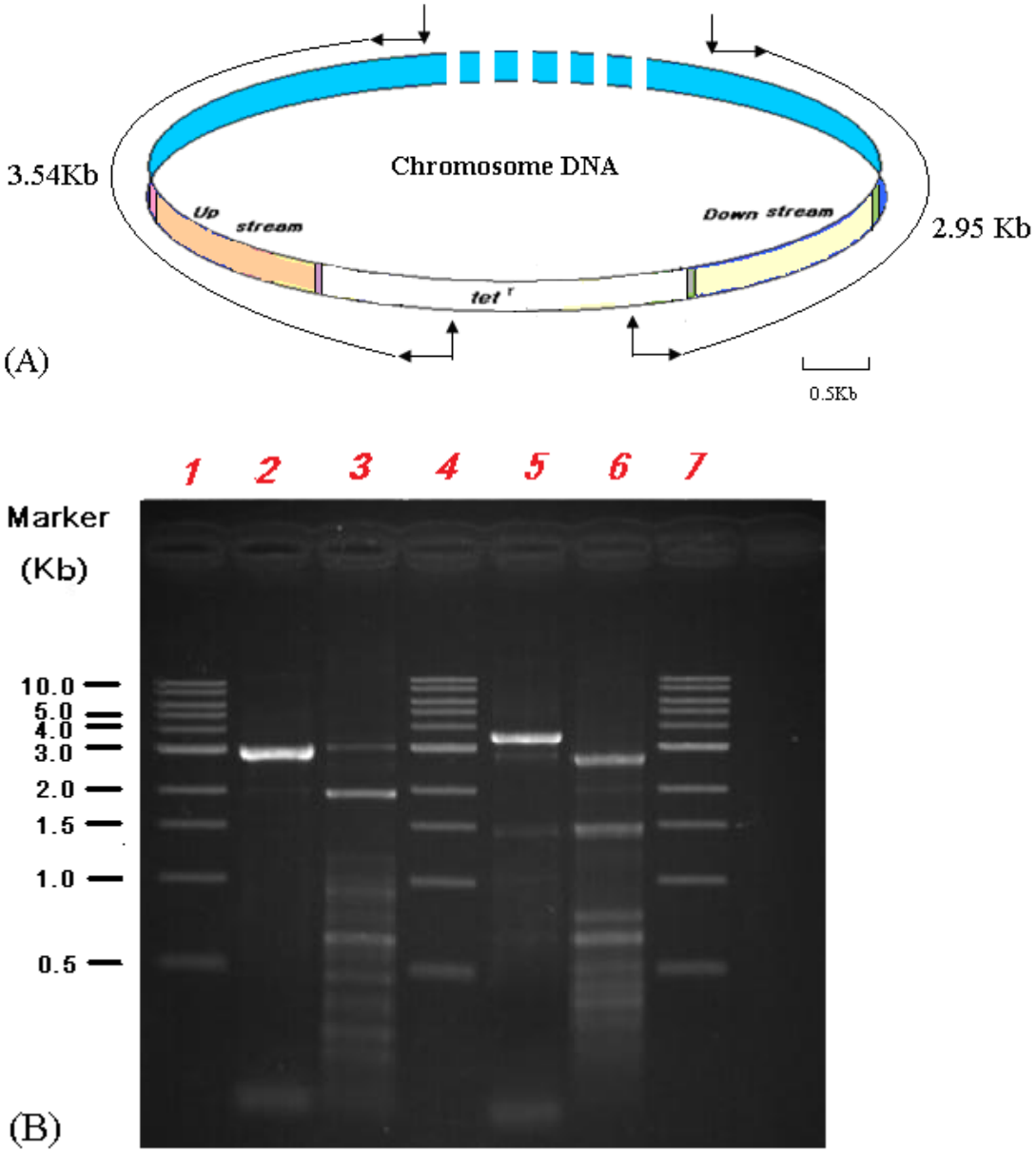



Fig 4.8 Distinguishing the double recombinant mutant ΔSPO1703. (A) Illustration of the regions amplified. (B) Gel picture showing the result of verification. Lanes 1, 4, and 7 are markers. Lane 2 is amplified from the middle of *tet*$^r$ region to the 1Kb down-stream region. Lane 5 is amplified from the 1Kb up-stream region to the middle of the *tet*$^r$ region. Lanes 3 and 6 are negative controls of two pairs of primers.

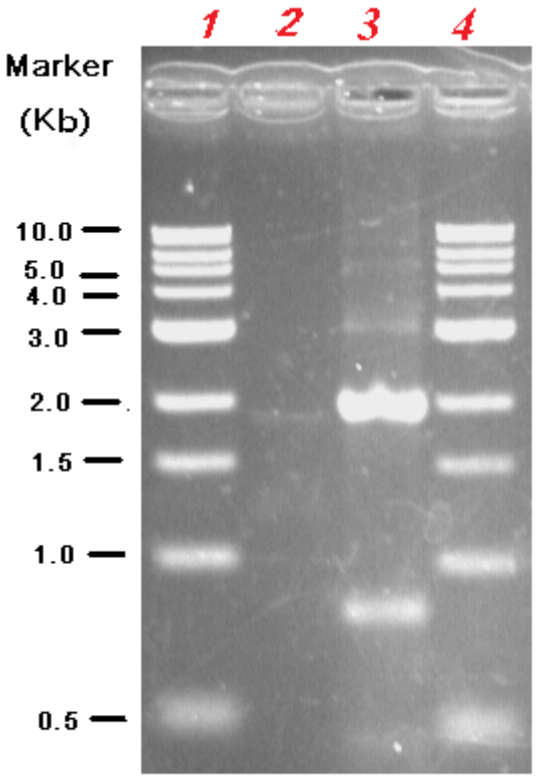



Fig 4.9 Verification of mutant ΔSPO1703. Primers were designed to amplify the internal region of gene SPO1703. Lane 1 and 4 are markers. Lane 2 is the amplification of SPO1703 gene in the mutant. Lane 3 is the positive control of wild type.

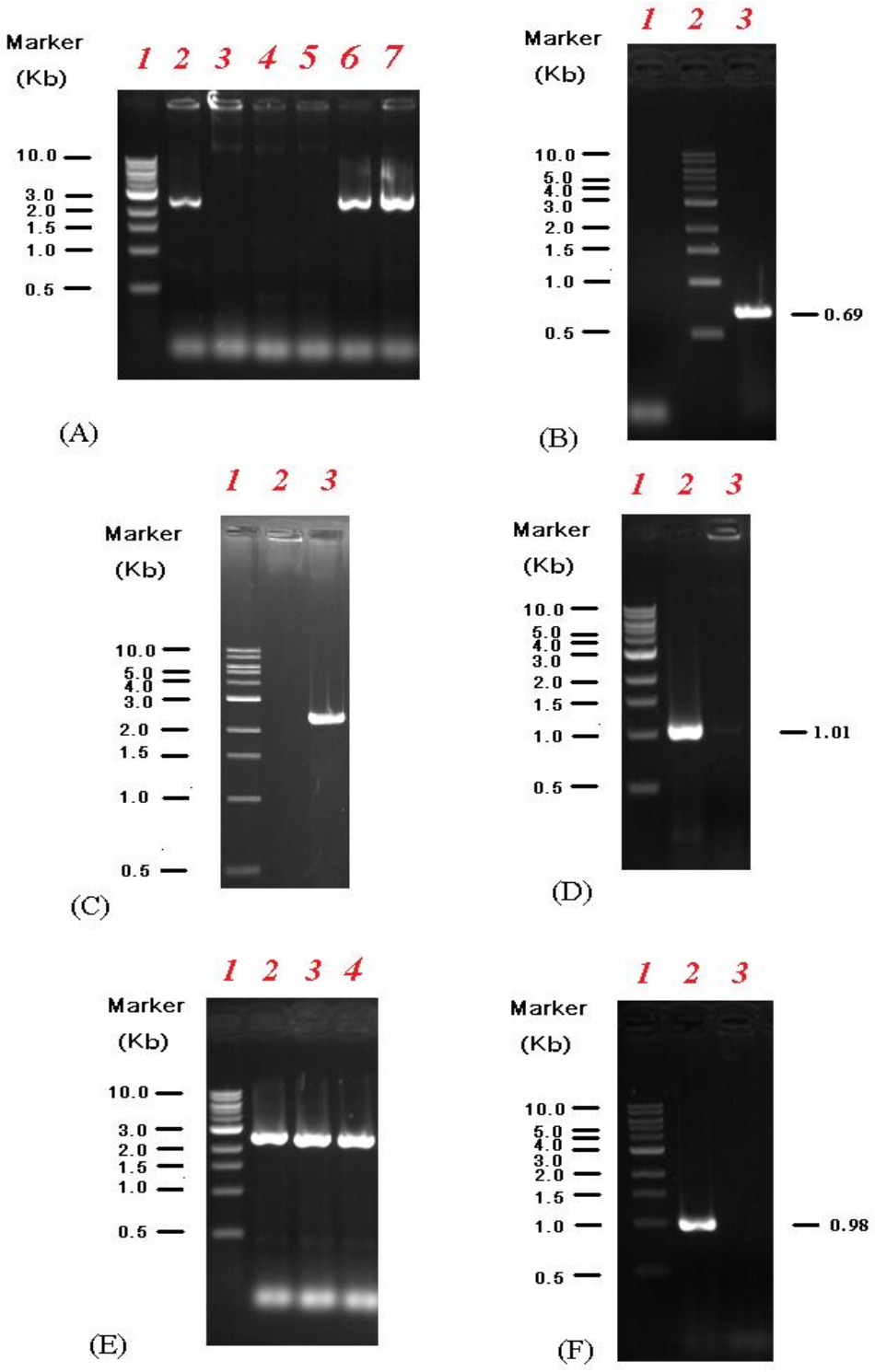
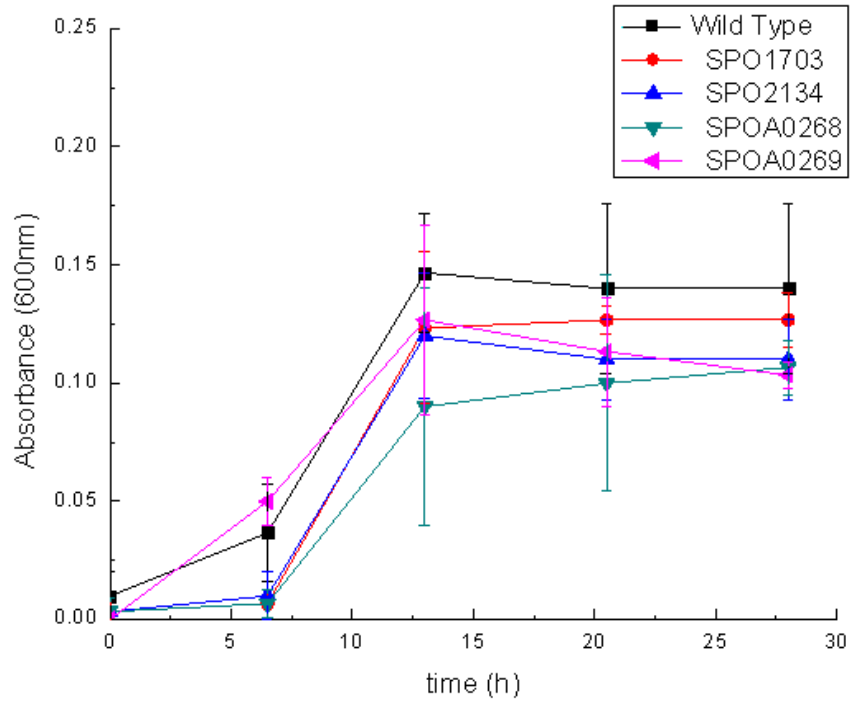



Fig 4.10 Verification of mutant ΔSPOA0268 (A) and (B), ΔSPOA0269 (C) and (D), and ΔSPO2134 (E) and (F). Primers were designed to amplify the *tet*$^r$ region (A, C, E), and internal regions of genes (B, D, F, respectively; positive controls are labeled with a size marker in the tight margin).

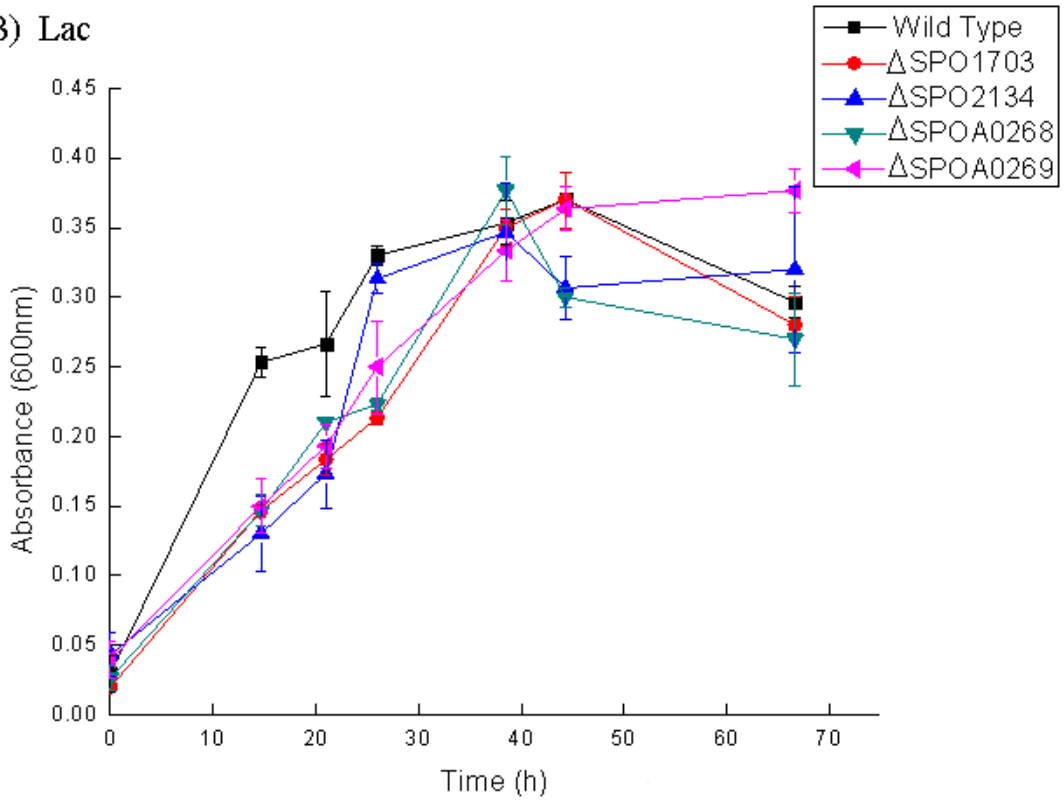
*4.4.4 Growth testing*

Both the wild type and mutants were grown to mid-log growth stage (0.15–0.20 $OD_{600nm}$) in ½ YTSS. The cells were then diluted 1:1000 with Basal medium (MBM), transferred into test tubes, and cultured at 30 °C. Different carbon sources were added to the MBM during the incubation as the substrate (Table 4.3). From the growth curves obtained, we can find that all the strains grow well in glucose, lactate, and acetate (Fig 4.11 A, B, and G). Besides, all the strains grew poorly in acrylate, methylmalonate, and no substrate (Fig 4.11 E, H, and I). More importantly, we can also find that compared to the wild type, all the mutants grew weakly in DMSP (Fig 4.11 C), to supporting the idea that these candidate genes may play a role in the DMSP degradation pathway. Finally, the mutant ΔSPOA0268 acted differently compared to both wild type and other mutants (Fig 4.11 D and F), since it grew slowly and weakly with MMPA but quickly with propionate. Although gene SPOA0269 and SPOA0268 are in the same operon, the phenotypes of their mutants are very different. SPOA0268 is annotated as an IclR family transcriptional regulator, which is a DNA binding transcription regulator, and SPOA0269 as a hypothetical protein. Therefore, it is possible that this operon plays a role not only in DMSP degradation but also MMPA uptake, MMPA degradation, or downstream of these pathways. SPOA0268 is also possible a regulator for uptake of propionate, or degradation of propionate.

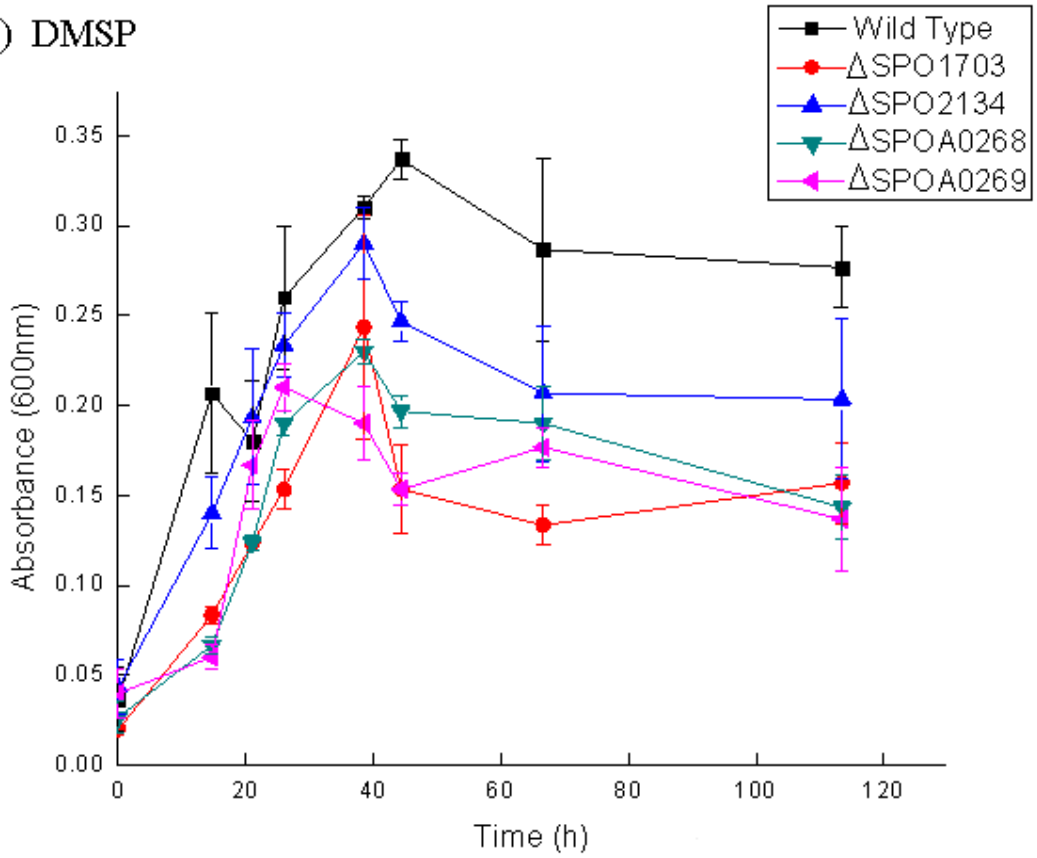
(A) GLU


Wild Type
SPO1703
SPO2134
SPOA0268
SPOA0269
Absorbance (600nm)
time (h)
0.00
0.05
0.10
0.15
0.20
0.25
0
5
10
15
20
25
30

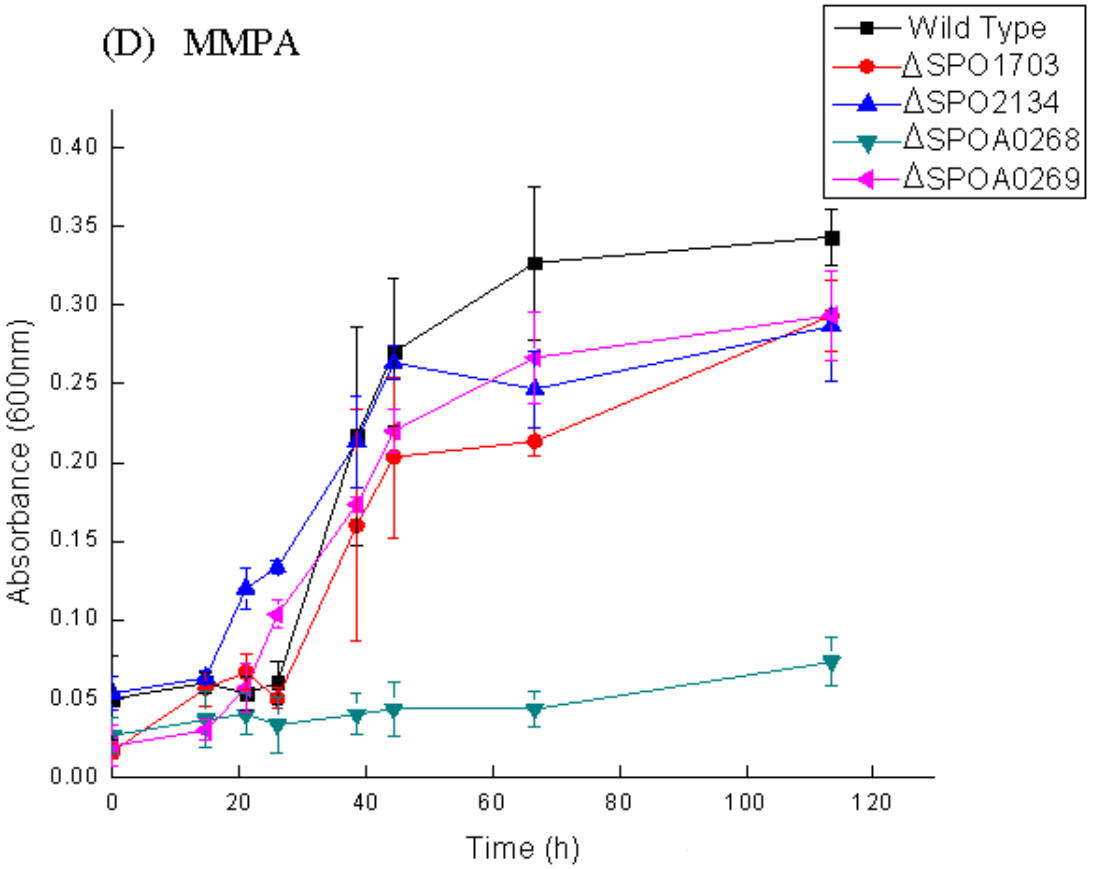
(B) Lac


Wild Type
ΔSPO1703
ΔSPO2134
ΔSPOA0268
ΔSPOA0269
Absorbance (600nm)
Time (h)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0
10
20
30
40
50
60
70

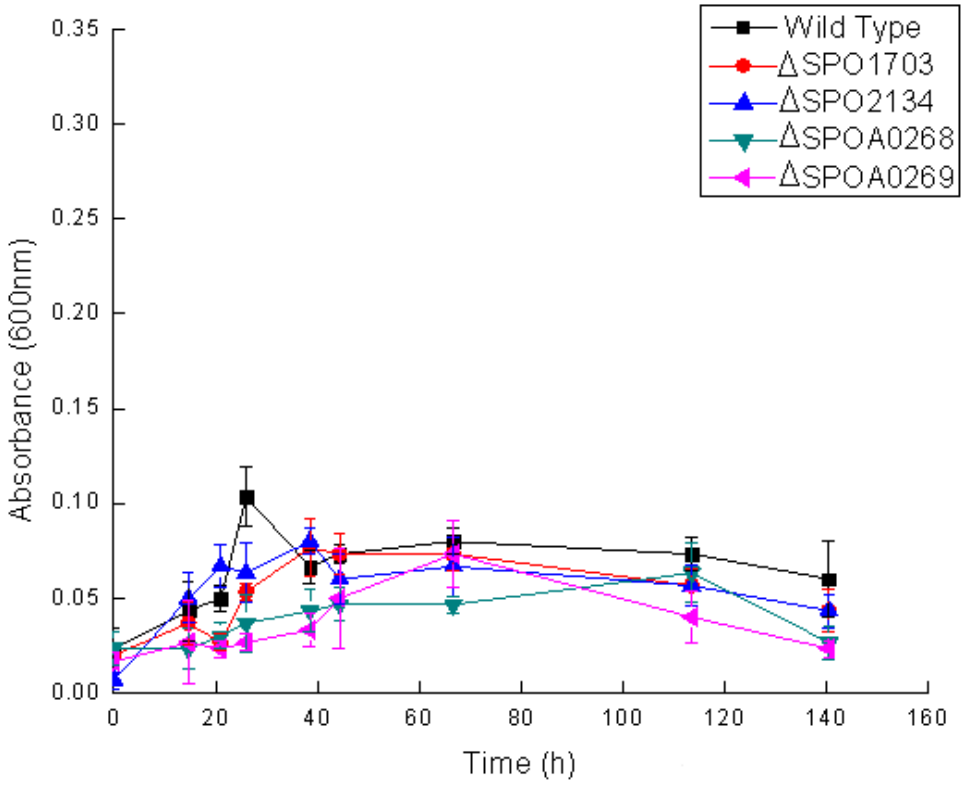
(C) DMSP


Wild Type
ΔSPO1703
ΔSPO2134
ΔSPOA0268
ΔSPOA0269
Absorbance (600nm)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
20
40
60
80
100
120
Time (h)

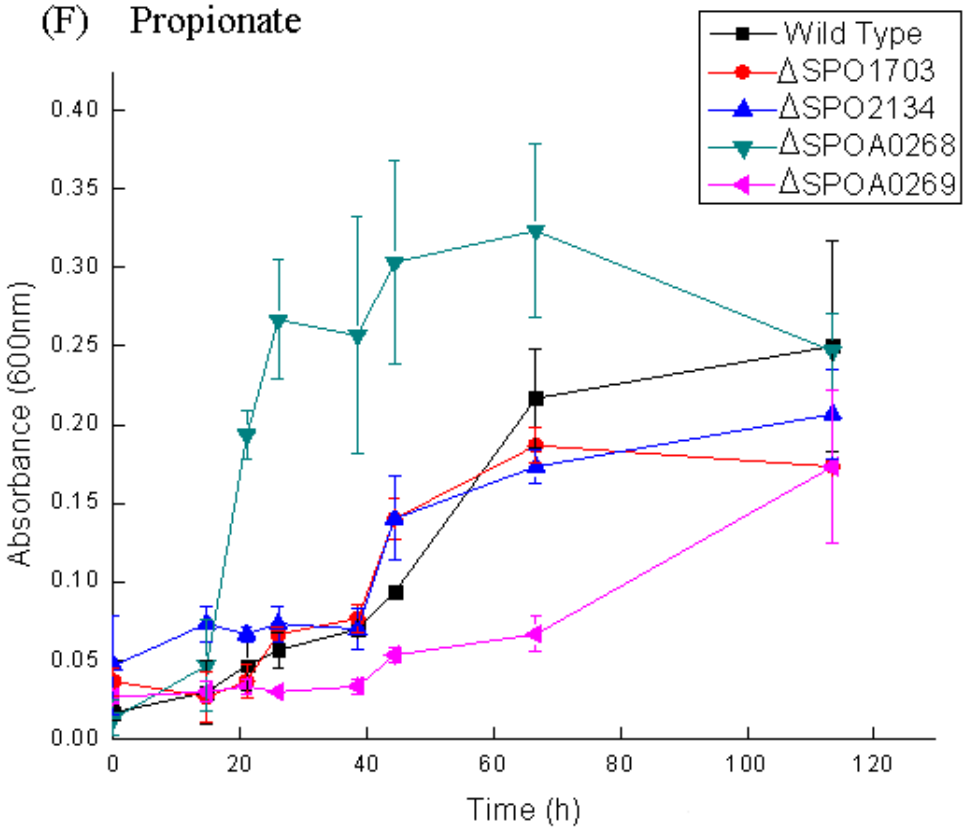
(D) MMPA

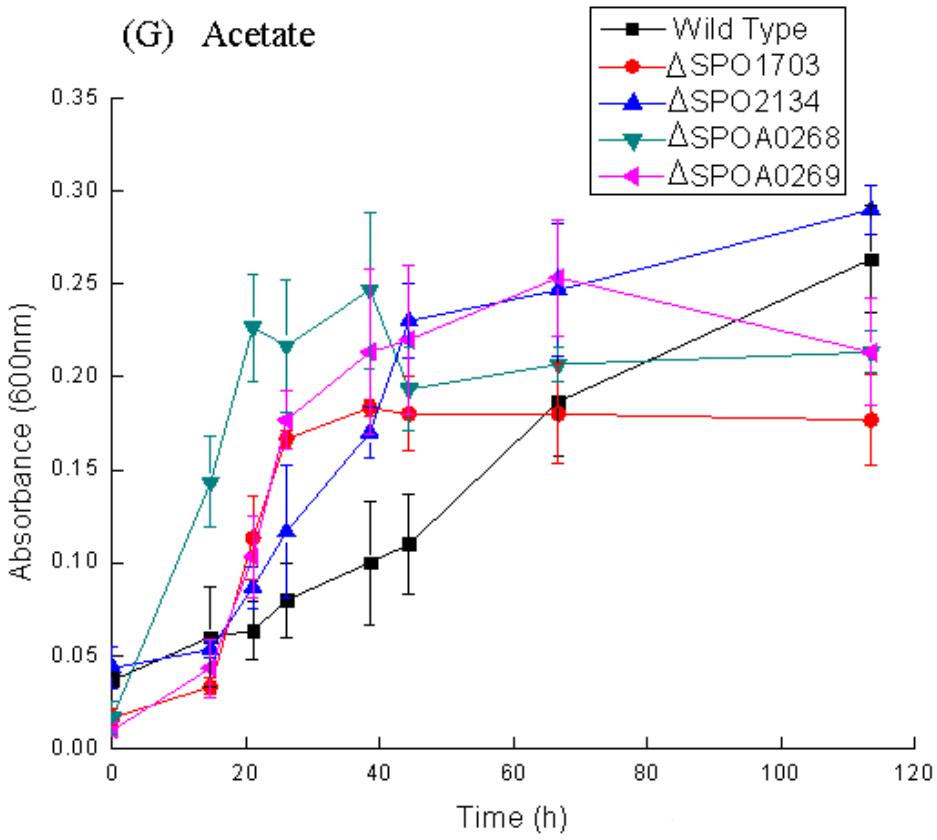

Wild Type
ΔSPO1703
ΔSPO2134
ΔSPOA0268
ΔSPOA0269
Absorbance (600nm)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0
20
40
60
80
100
120
Time (h)

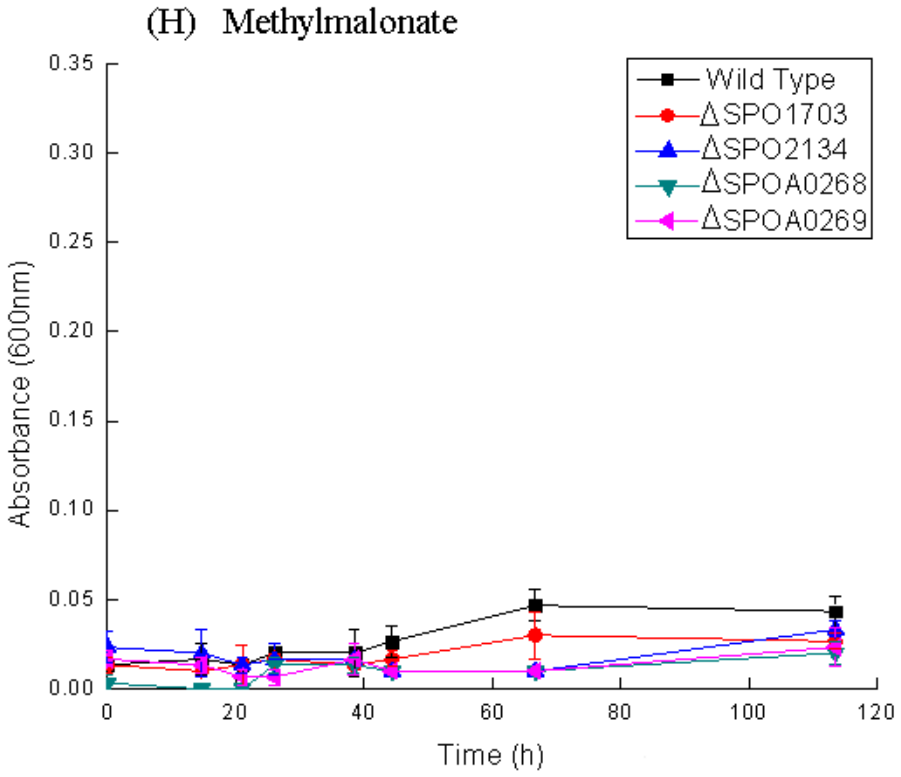
(E) Acrylate




(F) Propionate




(G) Acetate
Wild Type
ΔSPO1703
ΔSPO2134
ΔSPOA0268
ΔSPOA0269
Absorbance (600nm)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
20
40
60
80
100
120
Time (h)


(H) Methylmalonate
Wild Type
ΔSPO1703
ΔSPO2134
ΔSPOA0268
ΔSPOA0269
Absorbance (600nm)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
20
40
60
80
100
120
Time (h)

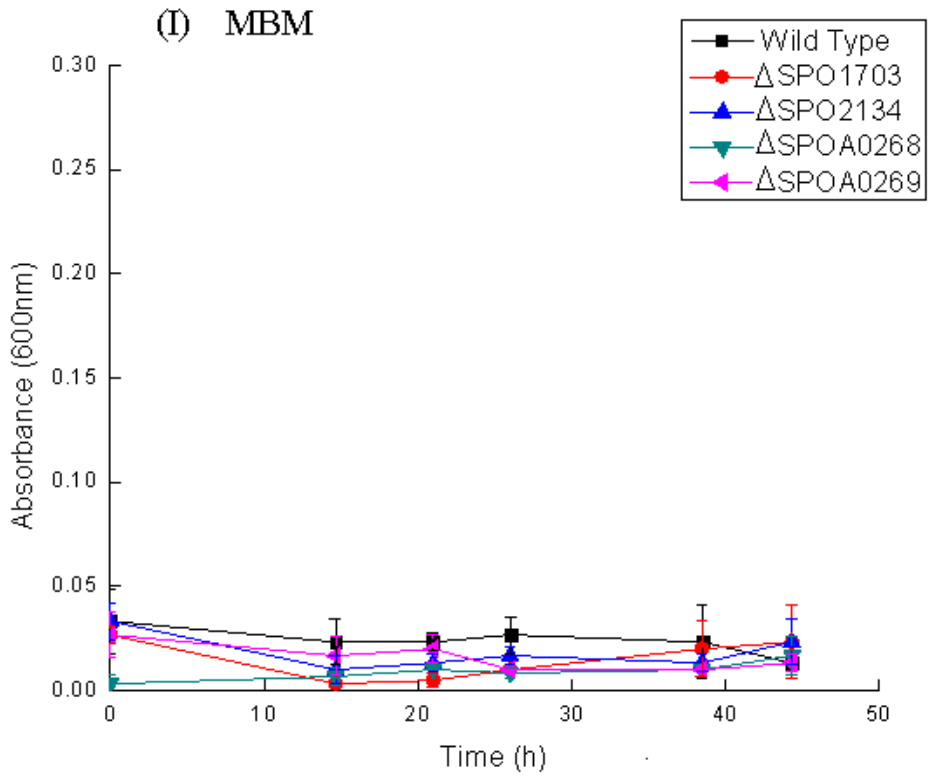



Fig 4.9 Growth curve under different carbon sources.

## 4.5 REFERENCES

Bates, T. S., R. J. Charlson, and R. H. Gammon (1987). "Evidence fof the climatic role of marine biogenic sulphur." Nature 329: 319 - 321.

Elledge, M. Z. L. S. J. (2007). "Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC." NATURE METHODS 4.

Ivanova, E. P., Gorshkova, N. M., Sawabe, T., Zhukova, N. V., Hayashi, K., Kurilenko, V. V., Alexeeva, Y., Buljan, V., Nicolau, D. V., other authors (2004). "Sulfitobacter delicatus sp. nov. and Sulfitobacter dubius sp. nov., respectively from a starfish (Stellaster equestris) and sea grass (Zostera marina)." Int J Syst Evol Microbiol 54: 475 - 480.

Kiene, R. P., L. J. Linn, J. González, M. A. Moran, and J. A. Bruton (1999). "Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton." Applied and Environmental Microbiology 65: 4549 - 4558.

Kiene, R. P. a. L. J. L. (2000). "Distribution and turnover of dissolved DMSP and its relationship with bacterial production and dimethylsulfide in the Gulf of Mexico." Limnology and Oceanography 45: 849 - 861.

Macián, M. C., Ludwig, W., Aznar, R., Grimont, P.D.A., Schleifer, K.H., Garay, E. and Pujalte, M.J. (2001). "Vibrio lentus sp. nov., isolated from Mediterranean oysters." Int. J. Syst. Evol. Microbiol 51: 1449 - 1456.

Matthias Labrenz, B. J. T., Paul A. Lawson, Matthew D. Collins, Peter Schumann and Peter Hirsch (2000). "Staleya guttiformis gen. nov., sp. nov. and Sulfitobacter brevis sp. nov.,

a-3-Proteobacteria from hypersaline, heliothermal and meromictic antarctic Ekho Lake." Int J Syst Bacteriol 49: 137 - 147.

Moran MA, B. A., González JM, Heidelberg JF, Whitman WB, Kiene RP, Henriksen JR, King GM, Belas R, Fuqua C, Brinkac L, Lewis M, Johri S, Weaver B, Pai G, Eisen JA, Rahe E, Sheldon WM, Ye W, Miller TR, Carlton J, Rasko DA, Paulsen IT, Ren Q, Daugherty SC, Deboy RT, Dodson RJ, Durkin AS, Madupu R, Nelson WC, Sullivan SA, Rosovitz MJ, Haft DH, Selengut J, Ward N (2004). "Genome sequence of Silicibacter pomeroyi reveals adaptations to the marine environment." Nature 432(7019): 910 - 913.

Pukall, R., Buntefuß, D., Fru¨ hling, A., Rohde, M., Kroppenstedt, R. M., Burghardt, J., Lebaron, P., Bernard, L., Stackebrandt, E. (1999). "Sulfitobacter mediterraneus sp. nov., a new sulfite-oxidizing member of the a-Proteobacteria." 49 Int J Syst Bacteriol: 513 - 519.

Rippka R, C. T., Hess W, Lichtlé C, Scanlan D J, Palinska K A, Iteman I, Partensky F, Houmard J, Herdman M (2000). "Prochlorococcus marinus Chisholm et al. 1992 subsp. pastoris subsp. nov. strain PCC 9511, the first axenic chlorophyll a2/b2-containing cyanobacterium (Oxyphotobacteria)." International journal of systematic and evolutionary microbiology 50: 1833- 47.

Sambrook, J., E. F. Fritsch, and T. Maniatis (2001). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory, New York, 3rd edition.

SIMO (2001). "Production of atmospheric sulfur by oceanic plankton: biogeochemical, ecological and evolutionary links." Trends Ecol Evol 16: 287 - 294.

Sorokin, D. Y. (1995). "Sulfitobacter pontiacus gen. nov., sp. nov. a new heterotrophic bacterium from the Black Sea, specialised on sulfite oxidation." Microbiology (English translation of Mikrobiologiya) 64: 295 - 305.

Suzuki, M. T., Béjà, O., Taylor, L. T. & DeLong, E. F (2001). "Phylogenetic analysis of ribosomal RNA operons from uncultivated coastal marine bacterioplankton." Environ. Microbiol 3: 323 - 331.

Thompson, F. L., Thompson, C. C., Li, Y., Gomez-Gil, B., Vandenberghe, J., Hoste, B., Swings, J. (2003). "Vibrio kanaloae sp. nov, Vibrio pomeroyi sp. nov. and Vibrio chagasii sp. nov., from sea water and marine animals." Int J Syst Evol Microbiol 53: 753 - 759.

Todd, J. D., R. Rogers, Y. G. Li, M. Wexler, P. L. Bond, L. Sun, A. R. J. Curson, G. Malin, M. Steinke, and A. W. B. Johnston (2007). "Structural and regulatory genes required to make the gas dimethyl sulfide in bacteria." Science 315: 666 - 669.

Uchino, Y., Hirata, A., Yokota, A., Sugiyama, J. (1998). "Reclassification of marine Agrobacterium species: proposals of Stappia stellulata gen. nov., comb. nov., Stappia aggregata sp. nov., nom. rev., Ruegeria atlantica gen. nov., comb. nov., Ruegeria gelatinovora comb. nov., Ruegeria algicola comb. nov., and Ahrensia kieliense gen. nov., sp. nov., nom. rev." J Gen Appl Microbiol 44: 201 - 210.

Wagner-Döbler, I., Rheims, H., Felske, A., Pukall, R. & Tindall, B. J (2003). "Jannaschia helgolandensis gen. nov., sp. nov., a novel abundant member of the marine Roseobacter clade from the North Sea." International Journal of Systematic and Evolutionary Microbiology 53: 731 - 738.

# APPENDIX A

## CHAPTER 3 SUPPLENMENTARY MATERIAL

R Code for MDS, HCLUSTER, and HEATMAP generation

```
#### Install Biocondustor package ###########

source("http://bioconductor.org/biocLite.R")
biocLite()


#### Update source  ########################

source("http://bioconductor.org/biocLite.R")
update.packages(repos=biocinstallRepos(), ask=FALSE)


######## Install additional packages ########

source("http://bioconductor.org/biocLite.R")
biocLite("EBImage")
biocLite(c("pkg1", "pkg2"))


################################################
#Upload data into a variable called simo_data#
################################################

################Code Start#####################

####################  HCluster ####################
rawdata<-read.table("cluster.txt")
dim(rawdata)
 probes<-rawdata$probe.set
alldata<-rawdata[,2:7]
```

```
samplename<-colnames(alldata)
traindata<-rawdata[,2:6]
testdata<-rawdata[,7]
distance<-dist(t(alldata))
hc<-hclust(distance,method="average")
plot(hc,main="hclust")



#################### MDS #########################

dist<-dist(t(alldata),diag=T,upper=T)
SampleMDS<-cmdscale(dist,k=2)
plot(SampleMDS,main="MDS")
text(SampleMDS,samplename,col=c(rep("black",3),rep("red",2),rep("blu
e",1)))


################### Heat map ####################

data<-rawdata[,-1]
data<-rawdata[,-1]
data<-as.matrix(data)
expdata<-log(data)
expdata

r1<-expdata[,1]
r2<-expdata[,2]
r3<-expdata[,3]
r=(r1+r2+r3)/3
expdata<-expdata-r


color_time <- rep(c(2,3,4,5,6,7),1)
map_color <- colorRampPalette (c("Red", "Green")) (32)
heatmap (expdata, cexRow=0.5, cexCol=1.5, col=map_color,xlab="Samples",
ylab="Genes",ColSideColors=as.character(color_time))
```

# APPENDIX B

## CHAPTER 4 SUPPLENMENTARY MATERIAL

### Marine Basal Media

| Final Volume | 1 Liter | 500 ml | 100 ml |
|---|---|---|---|
| Sea Salts | 20 g | 10 g | 2 g |
| Basal Media | 250 ml | 125 ml | 25 ml |
| $DH_2O$ | 700 ml | 350 ml | 70 ml |
| FeEDTA Stock | 50 ml | 25 ml | 5 ml |
| Vitamin Supplement | 1 ml | 0.5 ml | 0.1 ml |

To prepare, mix $DH_2O$, Basal Media, and Sea Salts thoroughly. Most of the Sea Salt should be dissolved, before autoclaving. Allow to cool before adding FeEDTA, Vitamin Supplement, and Carbon Source. This medium recipe was from Christopher Reisch, Dr. Whitman's lab.

Basal Media

160 ml 1M HEPES, pH 6.8
0.16 g $K_1HPO_4$
1.52 g $NH_4Cl$
340 ml $DH_2O$
Autoclave and store at room temperature.

FeEDTA Stock

0.050 g FeEDTA
100 ml $DH_2O$

Autoclave and store at $4^oC$

Vitamin supplement (per 100 ml)

2 mg biotin
2 mg folic acid
10 mg pyridoxine-HCl
5 mg riboflavin
5 mg thiamine
5 mg nicotinc acid
5 mg pantothenic acid
0.1 mg cyanocobalamin
5 mg p-aminobenzoic acid
Filter sterilize (0.2 μm) and store at -20°C