Generation of large-insert metagenomic libraries from subtropical hypersaline microbial mats and their screening for antibiotic resistance

by

Irimar Torres-Zapata

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
in
BIOLOGY

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS
2012

Approved by:

_____	_____
Carlos J. Santos-Flores, Ph.D. Member, Graduate Committee	Date
_____	_____
Juan C. Martínez-Cruzado, Ph.D. Member, Graduate Committee	Date
_____	_____
Carlos Ríos-Velázquez, Ph.D. President, Graduate Committee	Date
_____	_____
Gayle W. Griggs, Ed.D. Representative, Office of Graduate Studies	Date
_____	_____
Nanette Diffoot Carlo, Ph.D. Director, Biology Department	Date

ABSTRACT

Most of the microbial diversity in an environment cannot be studied through laboratory standard media culture methods. For this, an emerging science, coined “Metagenomics”, has been developed as an effort to access the majority of microorganisms. It involves the isolation and further characterization of microbial genomes and their study from different perspectives, using sequence-based and function-based analyses. Metagenomic studies of several environments (i.e. soil, water, animals, among others) have unraveled novel activities with industrial, biotechnological, and biomedical potential. Recently (during the last decade), unique environments such as microbial mats have also been subjected to metagenomic approaches. Microbial mats are organo-sedimentary structures found in extreme environments harboring highly metabolically diverse microorganisms. While macro and microscopic analyses as well as geomicrobiological studies have been performed with the microbial mats of Cabo Rojo Salterns, non-metagenomic studies were done to these ecosystems at that point. Our study presented the generation of large-insert metagenomic libraries from two subtropical hypersaline microbial mats (benthic and ephemeral) during the dry and rainy seasons. Also, due to the increase in antimicrobial resistant clinical isolates threatening human health, the metagenomic libraries were monitored by a functional-based analysis for the search of antibiotic resistance. Using an indirect extraction method, total DNA from microbial mats samples was isolated. DNA fragments of more than 20 Kbp were cloned into fosmids, packed *in vitro* and transduced into a host strain. Four metagenomic libraries were generated with a total of 64,600 clones with inserts ranging from 20-100 Kbp. The libraries were screened for resistance to ampicillin, tetracycline, spectinomycin, gentamicin and kanamycin. Only gentamicin and kanamycin resistant clones were isolated with inserts of approximately 30 Kbp and 40 Kbp, respectively. The restriction

analysis and the retransformation of the fosmid into an isogenic strain confirmed the presence of an insert responsible for the resistance. The gentamicin resistant clone was mutagenized and further characterized by primer walking. The *in silico* analysis suggested the presence of five open reading frames from which two were related to antibiotic resistance genes. These included a 16S rRNA methyltransferase and an N-acetyltransferase most related (less than 40 % of identity with NCBI protein database) to *Chloroflexus* and uncultured prokaryote enzymes, respectively. Our data confirmed metagenomics as an emerging technology to unravel novel microbial strategies for biomedical application such as antibiotic resistance in environments as unique as microbial mats.

RESUMEN

La mayoría de la diversidad microbiana en un ambiente no puede ser estudiada a través de métodos basados en medios de cultivo de laboratorio. Por esto, una ciencia emergente nombrada “Metagenómica” ha sido desarrollada para alcanzar a estudiar la mayoría de los microorganismos. Ésta involucra el aislamiento y la subsiguiente caracterización de genomas microbianos y su estudio desde diferentes perspectivas, utilizando análisis basados en secuencia y en función. Estudios metagenómicos de varios ambientes (suelo, agua, animales, entre otros) han descubierto actividades nuevas con potencial industrial, biotecnológico y biomédico. Recientemente (menos de una década), ambientes únicos como los tapetes microbianos también han sido sometido a estudios metagenómicos. Los tapetes microbianos son estructuras laminares órgano-sedimentarias que se encuentran en ambientes extremos y albergan microorganismos con una alta diversidad metabólica. Análisis macro y microscópicos así como estudios geomicrobiológicos se han llevado a cabo con los tapetes microbianos de las salinas de Cabo Rojo en Puerto Rico mientras que ningún estudio metagenómico de estos ecosistemas se había realizado hasta este momento. Nuestro estudio presenta la generación de bibliotecas metagenómicas de alto peso molecular de dos tapetes microbianos subtropicales hipersalinos (béntico y efímero) durante la época seca y lluviosa. También, debido al aumento en el número de aislados clínicos resistentes a antibióticos que amenaza la salud humana, las bibliotecas metagenómicas fueron utilizadas en un análisis basado en función para la búsqueda de resistencia a antibiótico. ADN total de los tapetes microbianos fue aislado utilizando un método indirecto de extracción. Fragmentos mayores de 20 Kpb fueron clonados en fósmidos, empacados *in vitro* y transducidos a una cepa bacteriana huésped. Cuatro bibliotecas metagenómicas fueron generadas para un total de 64,600 clones con insertos desde 20-100 Kpb.

Las bibliotecas fueron utilizadas para buscar resistencia a ampicilina, tetraciclina, spectinomicina, gentamicina y kanamicina. Sólo clones resistentes a gentamicina y kanamicina fueron aislados con insertos de aproximadamente 40 Kpb y 30 Kpb, respectivamente. El análisis de restricción y la retransformación de de los fósmidos resistentes en una cepa isogénica confirmó la presencia de insertos responsables de los fenotipos de resistencia. Un clon resistente a gentamicina fue mutagenizado y luego caracterizado por "primer walking" y un análisis *in silico*. El análisis sugirió la presencia de cinco marcos de lectura abiertos de los cuales dos fueron relacionados con genes de resistencia a antibiótico. Éstos incluyen una 16S rRNA meiltransferasa y una N-acetyltransferasa más relacionadas (menos de 40% bde identidad con NCBI protein database) a *Chloroflexus* y un procariota no-cultivable, respectivamente. Nuestros datos confirman la metagenómica como una tecnología emergente para descubrir estrategias microbianas nuevas con aplicación biomédica como resistencia a antibiótico en ambientes tan únicos como los tapetes microbianos.

DEDICATION

I dedicate this thesis work to my precious daughter Damairí and beloved husband Damián. Both of you have been my inspiration to complete this research as a stepping stone to our future. Your unconditional love gave me the strength to finish this journey. Although there were days full of stress in which I felt uncomfortable and many times sad, when I reached home you two turned my world into a pleasant one. I can't imagine my life without you.

Damian, I want to thank you for loving me deeply. You were the one by my side in the best moments I have ever lived during my career, always being proud of me. You were the one encouraging me every day to continue and never give up. You were the one who understood me the most; even when you couldn't understand, you always listened to what I had to say. Thank you deeply for all your help with La Goly. I never had a NO for an answer and, thanks to that, I had extra time for research and writing. You are an excellent dad and husband!! You are mine, baby. I can continue detailing many things for what I am thankful but I'm just going to say: "Bebe thanks for accepting this challenge with me".

Damairí, I'm very fortunate for having you! You have changed my life. Before you, I was the only person I thought about. I had many future plans, I wanted to earn plenty of money and I never imagined myself as a mother. When you came to my story, I understood how fake my life was. I started to be a warmer person and many of my plans were replaced with others in which you and your dad could be part of because you two are part of me now. Now my steps will follow our family's best interests. Thanks for teaching me this lesson. I love you Goly and I'm very proud of being your mom. I want to thank you so much because you and your future encouraged me to grow as a professional and thanks to that I have finished this work.

“Bebé” I’ll never forget how much you have learned about bacteria, microbial mats and clones. We had many funny scientific conversations. Also, Damairí learned to say “laboratorio” and “Colegio” as part of her first set of spoken words. Unbelievable! I love you both with all my heart.

ACKNOWLEDGEMENTS

First of all I want to thank the Lord for giving me the strength and courage to walk and enjoy this path, and for helping me to achieve my goal. I learned from God that all things are possible with Him and through our faith. Thank you God for having full control of my life because all I want is to walk by your side, in your perfect will.

Although I really don't have the words to explain my gratefulness, I want to express my deep gratitude to my mentor, chief and friend Dr. Carlos Ríos-Velazquez. Carlos, I owe to you who I became since my second year of college. You have been a professional role model to me and most important an excellent example of how to conduct an ethical scientific career. You have also been an important key to my success in my personal life. Thank you so much for your confidence in me since day one. I want to thank you for your unconditional advice, for your strict correction when needed and for your support through all these years. Thanks for being so caring with my family too, especially with Damairí. Don't ever forget she loves you and your mom very much! Carlos, I deeply value your friendship; thank you because I know it will last forever.

I want to express my deep gratitude to Lynn Williamson, Heather Allen and Jo Handelsman for sharing with me their knowledge in Metagenomics and personally having instructed me for this work. Thank you very much for being so special. Also, I want to thank the members of my thesis committee Dr. Carlos J. Santos-Flores and Dr. Juan C. Martínez-Cruzado for their guidance during this journey and the corrections of this manuscript. You are professional models to me and I admire your work and passion with all of us, your students.

I want to thank the B-266 Microbial Biotechnology and Bioprospecting Laboratory team, for being there for me always and giving me the enthusiasm to continue. Thanks to the my graduate colleagues and friends Michelle Rivera, Josué Malavé, Kristina Soto, Ricardo Burgos,

Vanessa Cardona, Keila Flores, Juan Vega, Rosivette Santiago, Frank Ferrer, Lizahira Rodríguez and Moisés De Jesús. I especially want to thank the lab partners Adel González, Christian Castro, Nerymar Ortiz and Odemaris Narváez, for their unconditional collaboration and disposition for this research.

In addition, I would like to thank Magaly Zapata for her great friendship and for making things possible even when it seemed too difficult. Thanks Magaly for your daily smile and love! Also, I want to acknowledge Gladys Toro and Idaris De Jesús for being so special and for being an example of strength and courage.

I thank the Biology Department for their full support during all these years. Thank you very much María Méndez, Mary Jiménez and Sandra Zapata for caring about your graduate students. Thanks for your guidance and advice, and for all the opportunities I had to grow professionally. Dr. Nanette Diffot and Dr. Jaime Acosta, thanks for being the heads of such a great team and for your hard work. Dr. Diffoot I deeply thank you for your confidence in me and for your advice. Brenda Soto, Mitzy Zavala and Alicia Collazo, thank you for your work and kindness.

I want to acknowledge The National Science Foundation Support: RUI Cabo Rojo Salterns Microbial Observatory (grant #0455620). Lilliam Casillas and Pieter Visscher, thanks for giving me the opportunity to work in the Microbial Observatory.

I will always be grateful to my parents for motivating and supporting me during all my life. Papi y Mami, thanks to you I have become a good person and I have learned to overcome many challenges throughout my whole life. I love you! Last, but not least, I want to thank my husband for being very understanding and for his unconditional support and to my daughter for being my inspiration to continue my career.

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION AND LITERATURE REVIEW

1.1 INTRODUCTION

Microbial mats (MM) are aquatic ecosystems composed of microbial cells attached to extracellular polymeric material and mineralized scaffolds in visible layers at a millimeter scale (Ley et al., 2006). Due to animal grazing and competition from macrophytes, these organosedimentary structures are commonly found in extreme environments such as hypersaline bodies of water (Casillas-Martínez et al., 2005), hot springs (Hiraishi et al., 1999), dry temperate deserts and cold dry environments (Clocksin et al., 2007). Microbial mats are the modern analogs of stromatolites, which were the first microbial communities that inhabited the Earth and contributed to the initial production of O_2, H_2 and CH_4 on the planet (Villanueva, 2005).

The microbial interactions that occur within the mats create gradients in the concentrations of oxygen, soluble sulfide, hydrogen ions and other chemicals. According to the metabolic capabilities of microorganisms, microbial mats are vertically stratified into colored layers (e.g. green, pink and black) that are dominated by different types of microorganisms. The green layer is rich in cyanobacteria, eukaryotic algae and aerobic organoheterotrophic bacteria (Fenchel, 1998). Anoxygenic phototrophic and chemolithotrophic bacteria are found in the pink layer directly underneath the surface layer, and in the deeper anoxic stratum (black layer) dissimilatory sulfate and sulfur reducing bacteria are found (Casillas-Martínez et al., 2005). Because of their composition and ordered spatial arrangement, microbial mats are ideal models for studying microbial interactions, element cycling and microbe-mineral interactions. Also, comparisons of biosignatures in fossilized and modern mats seek to describe ancient biogeochemical cycles and microbial activities (Ley et al., 2006).

Initially, it was thought microbial mats had a simple poorly diverse community of microorganisms (Risatti, 1994); however recent molecular studies of 16s rDNA have

demonstrated these ecosystems harbor high biological diversity (Spear, 2003). This diversity of microorganisms and their versatility to cope with extreme conditions make microbial mats a valuable source of genes with novel activities. Thereby influencing our interest in studying two subtropical hypersaline seasonal microbial mats (benthic and epehermal) from the Cabo Rojo Salterns in Puerto Rico to unravel their potential biotechnological and biomedical activities. Since around 99% of microorganisms in some environments are not cultivable (Streit and Schmitz, 2004), we will study the environmental DNA of these microbial mats using a metagenomic approach.

Metagenomics is the culture-independent genomic analysis of microbial communities in the environment (Riesenfeld et al., 2004). It consists in the direct cloning of environmental DNA into a readily cultivable organism such as *Escherichia coli* and the screening of the collection of clones (metagenomic library) for desirable biological traits (Handelsman et al., 1998). This technology facilitates the study of the microorganisms that are not cultivable in laboratory media, whether because the media and/or the physical and chemical parameters they are subjected to are not favorable for their growth or because they were in a viable but not cultivable state at the environment. The employment of this technique has led to the discovery of new functions in microbial communities such as novel proteins [e.g. proteins with antiporter activity (Majernik et al., 2001)], new antibiotics [e.g. Turbomycin A and B (Gillespie et al., 2002)] and new enzymes [e.g. agarolytic enzymes (Voget et al., 2003)]. Also, metagenomics has increased our knowledge in microbial ecology since besides studying the genetic information of microorganisms, with this approach, their functional role within the community can be determined (Streit and Schmitz, 2004).

Different types of microbial mats have been studied by metagenomic analyses worldwide to unravel the microbial diversity (Treude et al., 2005; Ward et al., 2007) and characterize novel proteins (Nunoura et al., 2005). To date, the functional genomics of microbial mats in the Caribbean is still being unraveled. Therefore, our research will focus on the generation of high molecular weight insert metagenomic libraries (MLs) from a benthic and an ephemeral microbial mat (in dry and rainy seasons) and their screening for activities with biomedical and/or biotechnological applications.

1.2 LITERATURE REVIEW

1.2.1 FROM CULTURE-INDEPENDENT METHODS TO METAGENOMICS

Before the era of metagenomics, microorganisms were enumerated and studied by cultivation techniques and direct microscopic cell counts (Jannasch and Jones, 1959) until it was recognized that the number of microbial colonies isolated on solid laboratory media was several orders of magnitude less than the number of total cell counts from natural environments (Roszak and Colwell, 1987; Giovannoni et al., 1988). Staley and Konopka named this phenomenon "The Great Plate Count Anomaly" where only 1 % or less of the total of microorganisms from any environment is readily cultivable in laboratory media (Amann et al., 1995). Since it was discovered that the number of microorganisms characterized until then was only a small fraction of its actual representation, many investigations have been directed to unravel the significant unknown microbial diversity.

Pace et al. (1985) proposed the idea of studying microbial diversity without culturing by generating community DNA shotgun libraries and the screening of 16s rDNA sequences by hybridization for its further sequencing. This technique was employed by Schmidt et al. (1991) who generated gene libraries from a marine picoplankton community to examine its phylogenetic diversity and found unique cyanobacteria and proteobacteria sequences and one eukaryotic sequence. Although Pace's method reveals the unknown microbiological diversity of an environment, it is quite laborious. The advent of Polymerase Chain Reaction simplified diversity studies by the direct amplification of 16s rDNA from total DNA samples. Giovannoni et al. (1990) were the first in using the PCR based method for diversity studies; they amplified, cloned and sequenced 16S rDNA from Sargasso Sea total DNA and found cyanobacteria sequences and a novel microbial group (the SAR11 cluster), which turned out to be a major constituent of this

oligotrophic ecosystem. None of the retrieved sequences matched rDNA sequences from cultivated marine microorganisms from similar environments. Since, the PCR based method has been exploited to study aquatic (Gordon and Giovannoni, 1996) and terrestrial environments (Liesack and Stackebrandt, 1992; Dunbar et al., 1999), microbial diversity and novel groups of microorganisms have been discovered and characterized (Massana et al., 2000).

Culture-independent techniques were initially established to study the unknown microbial diversity and eventually were directed towards the screening of desired functions. As an example, Healy et al. (1995) generated gene libraries, which they called "zoolibraries", from a mixture of thermophilic anaerobic digesters that were in continuous operation with lignocellulosic feedstocks for over 10 years. DNA (of 2-12 kbp) from these anaerobic digesters was cloned into pUC19 vector and the resulting clones were screened for cellulase and xylosidase activities; four cellulose degrading clones were isolated and characterized. These libraries were generated with DNA from a selective microbial group. As well, culture independent approaches can be used to access the functional role of microorganisms in the environment.

The idea of accessing the collective genomes of an environment (the metagenome) for the construction of metagenomic libraries was first proposed by Handelsman et al. (1998). They were interested in cloning large fragments of soil metagenome in *E. coli* with a Bacterial Artificial Vector (BAC) and monitor the libraries for biological activities. Also, they stated that BAC libraries could be a powerful tool for studying gene expression from diverse prokaryotes. This expectation was based on the screening of *Bacillus cereus* traits on a metagenomic library generated with *B. cereus* DNA and *E. coli* as the bacterial host by Rondon et al. (1999). It was found that, although *B. cereus* and *E. coli* are phylogenetically very distant, more than half of the

traits tested were expressed in the library. Even though they preferred to use the BAC vector in order to clone large fragments of DNA (100-150 kb), metagenomic libraries have also been generated with other vectors such as fosmids (Schloss and Handelsman, 2003) and cosmids (Sebat et al., 2003) which can also hold quite large inserts (25-40 kb and 30-50 kb, respectively).

The culture-independent technique of cloning the soil metagenome in order to access the phylogenetic and metabolic diversity of the majority of microorganisms contributed to the development of an emerging science coined Metagenomics. The word **metagenomics** was used for the first time by Handelsman et al. (1998). However, the metagenomic approach has also been termed community genomics, environmental genomics (Handelsman et al., 2002), eDNA libraries, recombinant environmental libraries, ecogenomics and others. This technology has been very successful in the study of single genes, pathways, organisms and communities of organisms (Riesenfeld et al., 2004).

One of the interests of generating metagenomic libraries is to determine the diversity of cultivable and non-cultivable microorganisms by amplifying and sequencing phylogenetic markers (Kimura, 2006). Another valuable use of this type of libraries is the identification of novel genes and functional activities, especially those with potential biotechnological and/or biomedical applications. The sequence-based and function-based screenings of different environments with metagenomic libraries have revealed both, previously described and novel genes and gene products.

The sequence-based screening involves the identification of genes of interest linked to phylogenetic markers in order to relate the function sequence with specific taxa. This strategy was first implemented by Béjà et al. (2000) when they sequenced the flanking regions of a bacterial 16S rDNA clone and identified a new photorhodopsin gene, which was then proven to

have light-harvesting capability and was the first report of bacterial rhodopsins since, prior to this study, microbial photorhodopsin were only found in Archaea. More than amplifying the 16S rDNA from metagenomic libraries, they have searched for other target genes or gene families such as photosystem II, polyketide synthase, methyl coenzyme M reductase A, histidine protein kinase, and others. Furthermore, they have been subjected to random sequencing of clones, usually in projects of microbial or viral genomes reconstruction and those interested in large-scale whole community sequencing. In addition to analyzing metagenomic libraries by sequencing, many research projects to screen clones for specific functions have been conducted.

Metagenomic libraries have been characterized by functional analyses, which have led to the discovery of many activities with industrial, biotechnological and biomedical applications. These activities include new antimicrobials, which have led to the development of novel antibiotics such as terragina, turbomycin A y B and acyl tyrosines (Riesenfeld et al., 2004), biosynthetic functions, hydrolytic and degradative enzymes, antibiotic resistance enzymes, membrane proteins, pigments (by visual inspection), and vitamin biosynthesis functions. The functionality of the discovered active clones validates the use of functional screening for characterizing metagenomic libraries (Riesenfeld et al., 2004). Moreover, the use of both sequence-based and function-based screenings facilitates ecological inferences.

Once active functional clones are isolated, the insert sequence is characterized in order to determine whether it is a novel or previously described function. Two methods for the identification of the gene (s) responsible for the desired function are transposon mutagenesis and sub-cloning. The transposon mutagenesis method consists in the random transposon insertions inside the insert sequence followed by the transformation of a cultivable bacterial strain (i.e. *E. coli*) with each insertion. In the sub-cloning method, the active clones are partially digested with

restriction enzymes and subsequently cloned into a small plasmid vector (i.e. pUC19). After each method described above, the following steps are to sequence the clone insert and to perform *in silico* analyses (Donato et al., 2010).

1.2.2 METAGENOMICS OF SOIL

The soil contains the largest microbial diversity on Earth, yet it is still mostly uncharacterized due to the microbe's inability to grow under laboratory conditions (Mocali and Benedetti, 2010). Since soil represents an untapped novel genetic diversity, it was the first environment subjected to metagenomic analyses and has been exploited for the exploration of novel biomolecules (Handelsman et al., 1998).

Soil metagenomics has focused on the isolation of biomolecules with industrial, biotechnological and biomedical potential. Since the work by Handelsman et al. (1998), metagenomic libraries from different types of soil have been generated and functions of interest have been identified. Some of the genes and functions include antimicrobials (Liles et al., 2003), antibiotic resistance enzymes (Allen et al., 2009), biocatalysts (Voget et al., 2003), lipases (Henne et al., 2000), proteases (Gupta et al., 2002), amylases (Yun et al., 2004), cellulases (Voget et al., 2006), membrane proteins (Majernik et al., 2001), amidases (Gabor et al., 2004), oxidoreductases (Knietsch et al., 2003), and others. Information about these metagenomic libraries such as type of soil sample, number of clones in the library and the identified biomolecules or genes is summarized in Table 1.1 and Table 1.2.

Table 1.1 Examples of metagenomic libraries generated from different types of soil.

Table 1 | **Soil-based libraries constructed without enrichment steps before DNA isolation**

Origin	Vector type	Number of clones	Average insert size (kb)	Total DNA (Gb)	Genes of interest	Year of construction	Refs
Meadow, sugar beet field, river valley	Plasmid	~1,500,000	5–8	7.8	4-hydroxybutyrate utilization, lipolytic enzymes, antiporter	1999	35,38,51
Uncultivated soil	BAC	3,648; 24,576	27; 44.5	1.19	Antimicrobials, antibiotic resistance; 16S rRNA, various biocatalysts	2000	37,48,50,69
Soil type not specified	Not specified	Not specified	Not specified	-	Antimicrobials	2000	46
Soil type not specified	Cosmid	700,000	Not specified	24.5*	Antimicrobials	2000	36
Soil type not specified	Cosmid	Not specified	Not specified	-	Pigments	2001	64
Uncultivated soil	BAC	12,000	37	0.42	Antimicrobials	2001	47
Soil type not specified	Cosmid	Not specified	Not specified	-	Fatty acid enol esters	2002	73
Alkaline loessian soil	Plasmid	100,000	8–12	1.0	Protease	2002	41
Calcerous grassland (sandy)	Fosmid	25,278	32.5–43.5	0.90	16S rRNA genes	2002	65
Calcerous grassland (sandy)	Fosmid	55,680	32.5–43.5	2.12	Acidobacterial 16S rRNA genes	2003	66
Arable field	Cosmid	5,000	Not specified	0.18*	Polyketide synthases, various other activities	2003	49
Meadow, sugar beet field, cropland	Plasmid	583,000; 360,000; 324,000	4.4; 3.8; 3.5	4.05	Carbonyl formation	2003	43
Sandy soil, sandy soil, mixed woodland soil	Fosmid	25,344; 30,366; 19,978	33–45	3.03	Taxonomic marker genes	2004	58
Clay loam sandy type	Fosmid	100,000	30–40	3.50	Polyketide synthase	2004	67
Forest soil	Fosmid	33,700	35	1.18	Lipolytic enzymes	2004	39
Soil type not specified	Cosmid	Not specified	Not specified	-	Long-chain *N*-acyltyrosines	2004	61
Plano silt loam soil	Plasmid	200,000; 58,000; 250,000; 650,000	4.1; 2.7; 3.5; 3.5	4.2	Antibiotic resistance	2004	50
Soil (surface covered with moss)	Plasmid	30,000	3.5	0.11	Amylolytic enzymes	2004	45
Agricultural field	Plasmid	80,000	5.2	0.42	Amidases	2004	63

(Daniel, 2005)

Table 1.2. Several biomolecules isolated from the screening of soil metagenomic libraries.

Some examples of screening for novel biomolecules from soil metagenomics libraries.

Biomolecule	Library type	Average insert size (kb)	Number of clones screened	Library size (Mb)	Number of hits	Hit rate (hit per Mb)	Ref.
Esterase/lipase	Plasmid	6	286,000	1, 716	3	1/572	Henne et al., 2000
Esterase/lipase	Plasmid	6	730,000	4380	1	1/4380	Henne et al., 2000
Esterase/lipase	BAC	27	3648	100	2	1/50	Rondon et al., 2000
Oxidation of polyols	Plasmid	3	900,000	2700	15	1/180	Knietsch et al., 2003a
Oxidative coupling enzymes	Cosmid	25	10,000,000	400,000	25	1/16,000	Banik and Brady, 2008
Alcohol oxidoreductase	Plasmid	4	400,000	1600	10	1/160	Knietsch et al., 2003b
Amidase	Plasmid	5	193,000	965	7	1/138	Gabor et al., 2004a
Amylase	BAC	27	3648	100	8	1/12	Rondon et al., 2000
Biotin production	Cosmid	35	50,000	1750	7	1/250	Entcheva et al., 2001
Protease	Plasmid	10	100,000	1000	1	1/1000	Gupta et al., 2002
β-Lactamase	Plasmid	5	80,000	400	4	1/100	Gabor, 2004
Cellulase	Cosmid	22	1700	37	8	1/5	Voget et al., 2006
Antibiotic	Fosmid	35,6	100,000	3560	13	1/274	van Elsas et al., 2008
Antibiotic	BAC	63	12,000	756	4	1/189	MacNeil et al., 2001
Antibiotic	BAC	44. 5	24,546	1092	3	1/364	Gillespie et al., 2002

(Mocali and Benedetti, 2010)

1.2.3 VIRAL METAGENOMICS

The metagenomic approach has improved studies of viral diversity since it has enabled the discovery and characterization of completely novel environmental viruses (Breitbart and Rohwer, 2005). Initially, Breitbart et al. (2002) and (2003) generated viral metagenomic libraries to analyze the unknown viruses' diversity from seawater and marine sediments, respectively. In 2006, Angly et al. studied the diversity of viral metagenomes (viromes) from samples of four different oceanic locations (Sargasso Sea, British Columbia Coastal, the Gulf of Mexico and the Arctic Ocean) taken over 10 years time. This research determined a huge diversity in the oceanic viral composition where 91% of the isolated DNA sequences were different from those found in the genetic databases (Hoff, 2006). Also, a new clade of single stranded-DNA phages was discovered, which dominated the Sargasso Sea sample. Recently, marine RNA viruses have also been subjected to metagenomic analyses that have shown picorna-like viruses predominate in the RNA virome (Kristensen et al., 2009). Furthermore, the taxonomic distribution of large DNA viruses from the sea has been described by a metagenomic study where giant viruses from the

family *Mimiviridae* were shown to be a ubiquitous and diverse component of the population of large eukaryotic DNA viruses in the sea (Monier et al., 2008). Schoenfeld et al. (2008) generated viral metagenomic libraries from two alkaline hot springs to enrich the knowledge of thermophilic viruses' diversity, biology and ecological impact.

Viral metagenomics has represented an important contribution to the research of viruses associated with animals, plants and humans, as well. It has generated an improved profile of previously described and unidentified viral entities that, in some cases, are linked to their host infections or associated with zoonotic transmissions. With this purpose, the virome of equine feces (Cann et al., 2004), human feces (Breitbart et al., 2003; Zhang et al., 2006) and bat guano (Li et al., 2010), has been sequenced using metagenomic tools. Also, virome studies have been applied to infected animal tissue [i.e. sea turtles with fibropapilloma (Ng et al., 2009), infected plants (Roossinck et al., 2010), infected corals (Vega et al., 2008), Asian children infected with cardioviruses (Kapoor et al., 2008), children with diarrhea (Finkbeiner et al., 2008) and H1N1 influenza A patients (Greninger et al., 2010) in order to enhance our understanding of viral diversity, evolution and pathogenesis. Moreover, metagenomics has another important application for the benefit of human health. For example, it has being used to detect attenuated virus sequence changes to maintain the safety record of life-attenuated vaccines (Victoria et al., 2010). Also, metagenomics has been applied in the characterization of viral communities present in potable and reclaimed water (Rosario et al., 2009). Recently, viral metagenomics has been demonstrated to be a powerful tool to study unidentified viruses infecting cell cultures from clinical isolates more accurately than PCR methods (Svraka et al., 2010). As with viruses, metagenomic analyses have been conducted to associations of microorganisms such as symbiotic relationships.

1.2.4 METAGENOMICS OF SYMBIONTS

The metagenomic approach has been employed to study microorganisms from symbiotic *relationships* because they usually cannot be grown in pure cultures. Some metagenomic approaches have provided information about the evolution of symbiosis and the biochemical dependence mechanisms microorganisms have developed to live in the relationship (Handelsman et al., 2004-2). Other metagenomic studies have isolated natural products with biomedical, biotechnological and industrial applications from symbiotic microorganisms. Metagenomic analyses have been conducted with symbionts associated to insects, marine organisms, protists and non-human mammals.

1.2.4.1 Insects

The relationship between the endosymbiont *Buchnera aphidicola* and aphids has been a role model for understanding obligate symbiosis. The metagenome of the *Buchnera*-aphid association was fully sequenced and the genome of *B. aphidicola* was reconstructed. *B. aphidicola* showed a reduced number of genes, especially the non-essentials and the ones related to cell defense; these results suggest this bacterium is an obligate symbiont of aphids (Shigenobu et al., 2000; Moran and Degnan, 2006). Also, the gut metagenome of the red turpentine beetle (*Dendroctonus valens*) was used to identify symbiotic bacterial species and their role in the adaptation of *D. valens* to survive harsh and nitrogen poor environments. Bacterial genera involved with cellulose breakdown in the gut and genes related to nitrogen fixation were reported (Morales-Jiménez et al., 2009). As with beetles, the gut of termites has been subjected to diversity studies and sequence-based metagenomic analyses where bacterial genera and genes involved in metabolic functions (i.e. H_2 production, reductive acetogenesis, and nitrogen

fixation) have been determined (Hongoh, 2010). Also, functions with biomedical, biotechnological and industrial applications have been isolated from beetles and termites metagenomes. The symbiotic metagenome of an uncultured bacterium and the beetle *Paederus fuscipes* has led to the discovery of type I polyketides synthases (Piel et al., 2003) with novel catalytic domains that produce pederin, a potential anti-tumor treatment, naturally used by this beetle for chemical defense (Piel et al., 2004a,b; Piel et al., 2005). The metagenome of termites have a large set of genes involved in cellulose and xylane hydrolysis that have been expressed *in vivo* and *in vitro* (Warnecke et al., 2007); these functions could be very useful in biomass utilization and industrial processes (Matsui et al., 2009).

1.2.4.2 Marine Invertebrates

Marine invertebrates such as annelids, bryozoans and sponges have been studied by metagenomic approaches. Woyke et al. (2006) performed a shotgun metagenomic analysis of the marine gutless worm *Olavius algarvensis* to study its symbiotic microbial community. A metabolic pathway reconstruction explained how symbionts meet the energy and waste management needs of this marine invertebrate. The symbiont metagenomic dataset of *O. algarvensis* also showed a large number of proteins rich in selenocysteine and pyrrolysine aminoacids, a feature associated to living under anaerobic conditions (Zhang and Gladyshev, 2007). Another annelid subjected to metagenomics is the hydrothermal tube worm *Riftia pachyptila.* Hughes et al. (1997) generated a *R. pachyptila* symbiont population DNA fosmid library to elucidate the signal transduction pathway for host-symbiont communication by probing the library with labeled histidine kinase PCR products. Recently, the uncultured *R. pachyptila* symbiont (named *Endoriftia persephone*) metagenome was sequenced and showed the presence of genes involved in carbon fixation plus all the genes necessary for heterotrophic metabolism,

suggesting the symbiont could live mixotrophically (Robidart et al., 2008; Bright and Bulgheresi, 2010). The metagenome of the symbiotic relationship between the marine bryozoan *Bugula neritina* and the bacteria *Endobugula sertula* lead to the isolation of bryostatins, a group of polyketides with anticancer activity (Davidson et al., 2001; Hildebrand et al., 2004).

Environmental genomics has also enhanced the knowledge of marine sponges' symbiotic microbial diversity and has contributed to the isolation of products with different potential applications. Marine sponges host a large number of uncultivated diverse bacteria that are important contributors to their chemistry (Hochmuth et al., 2010). These animals have been extensively subjected to sequence-based metagenomic approaches in order to study their diversity, evolution of symbiotic features and the presence of biomedically relevant gene sequences. Sponge associated microbial prokaryotes have been studied by 16S rDNA approaches (Grozdanov and Hentsche, 2007); however, metagenomic data has provided information about specific groups that live in symbiotic association with sponges, as well. For example, the genome sequence of novel bacteria was detected in a recombinant fosmid from a metagenomic library of the sponge *Aplysina aerophoba* (Fieseler et al., 2004). Also, the 16S rDNA of the uncultured archaeon *Cenarchaeum symbiosum* was detected in a large-insert metagenomic library from the sponge *Axinella mexicana* and sequenced (Schleper et al., 1998). Comparative metagenomic analyses of the bacterial microbiota in sponges have provided new insight into the evolution of their symbiotic relationships (Thomas et al., 2010). Hallam et al. (2006) assembled the genome of *C. symbiosum* by overlapping a set of clones from the sponge's metagenomic data. Furthermore, the metagenome of sponges has been used to compare the community composition of certain groups of microbes in sponges with high anthropogenic impact versus sponges with less impact (Turque et al., 2010).

Sponges are considered the most important marine source of biologically active compounds (Taylor et al., 2007; Hochmuth and Piel, 2009). Piel et al. (2004) found a group of genes that encodes polyketides which are structurally similar to pederin in the metagenome of the sponge *Theonellaswinhoei*, this being the first experimental proof that marine sponge derived compounds are produced by bacteria (Grozdanov and Hentschel, 2007). Metagenomic libraries from *Discodermia dissoluta* also were screened for polyketide (PKS) synthase gene clusters and they discovered a novel group of sponge-specific PKS ketosynthases (Schirmer et al., 2005). *Pseudoceratina clavata* metagenome was screened for PKS genes as well, and sponge-specific gene clusters related to *T.swinhoei* and *D. dissolute* were found (Kim and Fuerst, 2006). Recently a diverse group of ketosynthases were isolated from the metagenome of the marine sponge *Haliclona simulans* in Irish waters samples (Kennedy et al., 2008). All these studies show it is possible to isolate biologically relevant genes and gene products from marine sponges using metagenomic approaches, and that they can be applied to the search of other bioactive molecules. Interesting metagenomic studies have also been performed on other eukaryotic organisms (or to specimens of them) such as protists, birds, and non-human mammals.

1.2.5 METAGENOMICS OF PROTISTS AND HIGHER EUKARYOTES

1.2.5.1 Protists

Free-living protists genomics have relied mostly on the single cell sequencing approaches of cultivable protists and the sequence determination of amplified 18S rDNA (Dawson and Fritz-Laylin, 2009). A small number of metagenomic approaches have been applied to the study of protists. For example, the generation of a metagenomic library from a picoplanktonic community led to the first report of an uncultured marine alveolate (Massana et al., 2008). Also, protists genomics has relied on the expressed sequence tags (EST's) sequencing projects. Dawson and

Fritz-Laylin (2009) explained the advantages of applying three metagenomic approaches for protists research: sequencing environmental samples with high numbers of protists, sequencing protist/bacteria consortia and the transcriptomics of eukaryotic cDNA (expressed sequence tags). These applications will enhance the knowledge of protistan diversity and their ecological role.

1.2.5.2 Birds

Different types of birds (i.e. chicken, turkey) have been subjected to metagenomic analyses. Lu et al. (2007) determined metagenomic regions in the chicken fecal microbiome that can be used as potential genetic markers for fecal pollution tracking. Lu et al. (2008) revealed the microbial diversity in the feces of turkeys by analyzing 16s rDNA and metagenomic data. Recently, the viral metagenome of the gut of a turkey experiencing an enteric disease was used to determine the presence of previously described viruses and possible uncultured viruses that affect poultry production (Day et al., 2010).

1.2.5.3 Non-human Mammals

Metagenomics has also contributed to non-human mammal research. Mice, as animal models for human diseases, have been extensively studied by metagenomic approaches especially for obesity research. The obese phenotype has been related to the mouse's gut microbiota using metagenomic tools. Metagenomic and biochemical analyses have associated obesity with changes in the abundance of bacteria from the divisions *Bacteriodetes* and *Firmicutes* (Turnbaugh et al., 2006; Murphy et al., 2010). Furthermore, variations in mice gut's microbiota before and after the intake of modifiers (i.e. probiotics or antibiotics) have been determined and linked to obesity, using metagenomic data (Raoult, 2008). Also, metagenomic analyses have shown that the nutritional status of mice (low-fat or high-fat diet) changes their

gut's microbiome composition and gene expression (Turnbaugh et al., 2009). Mice's microbiome has led to the isolation of activities with potential applications as well. For example, β-glucanase gene products were isolated from mice's large-bowel metagenomic libraries (Walter et al., 2005).

As with mice, ruminant's microbiome metagenomic analyses have contributed to the isolation of several gene products with potential applications for biomedical, biotechnological and industrial procedures (Singh et al., 2008). Antibiotic resistance genes from cattle feces (Durso et al., 2010) and the fecal flora of feedlot steers (Harvey et al., 2009) have been detected by metagenomic approaches. Other enzymes isolated from ruminant metagenomes are cyclodextrinase (Ferrer et al., 2007), lipase (Liu et al., 2009-2), glycoside hydrolase (Brulc et al., 2009), and cellulase (Duan et al., 2009). Cellulases and tetracycline resistance genes have been detected in rabbits' (Feng et al., 2007) and pigs' (Kazimierczak et al., 2009) metagenomes, respectively.

Metagenomics has also been applied in evolutional studies of higher eukaryotes evolution studies. Noonan et al. (2005) generated a metagenomic library with DNA extracted from skeletal remains of 40,000 year-old cave bears to compare these sequences with the sequences of modern cave bears and reveal their evolutionary relationship. Also, 28 million base pairs of a Siberian mammoth were sequenced by Poinar et al. (2006), using a metagenomic approach. Shizuya et al. (1992) constructed a library from human total DNA using an F-factor-based vector and *E. coli* as the bacterial host. They obtained stable clones from about 300 kbp. Kim et al. (1996) also cloned human DNA in a BAC vector with inserts of approximately 140 kbp. Furthermore, Noonan et al. (2006) suggested and implemented the idea of studying the Neanderthal genome by metagenomic analysis. They developed a Neanderthal metagenomic library and determined by

high-throughput sequencing that Neanderthal and human genomes are at least 99.5% identical. This initial analysis suggests that Neanderthal genomics could advance the understanding of the evolutionary relationship of *Homo sapiens* and *Homo neanderthalensis*.

1.2.6 METAGENOMIC STUDIES OF MICROBIAL MATS

The first culture-independent analyses of microbial mats were focused on determining the diversity and community structure of microorganisms. Moyer et al. (1994) generated a 16S rDNA library from a microbial mat at an active hydrothermal vent to perform a Restriction Fragment Length Polymorphism analysis. Afterward, the diversity of microbial mats from different habitats was extensively studied by 16S rDNA clone libraries (Hiraishi et al., 1999; Taton et al., 2003; Summers et al., 2004; Sørensen et al., 2005; Ley et al., 2006; Meisinger et al., 2007). Also phylogenetic analyses of microbial mats using different DNA markers such as the coding gene for the large sub-unit of I ribulose 1,5-bisphophate carboxylase/oxygenase (RubisCO) (*rbc*L) have been conducted (Nanba et al., 2004).

The first metagenomic library of a microbial mat was generated by Elshahed et al. (2005). They were interested in studying the phylogenetic diversity of the microbial community of the mesophilic spring Zodletone at Oklahoma. By metagenomic analyses of a microbial mat formation in the subsurface of a hot water stream in Japan, Nunoura et al. (2005) have found genes encoding for proteins with industrial and biomedical uses such as a homologue of the carbon monoxide dehydrogenase and the V-type ATPase, respectively. A microbial mat from the Black Sea shelf was also studied by a metagenomic approach. Kube et al. (2005) found within this mat all the genes required for the complete pathway of anaerobic benzoate degradation. Another metagenomic analysis of microbial mats was done by Bryant and Frigaard (2006). They

studied the microbial mats of alkaline hot springs in Yellowstone National Park and revealed the presence of a distinctive bacteriochlorophyll synthesizing phototrophic bacterium.

Microbial mats metagenomic data has been used to compare environmental sequences with microorganisms' genome sequences. For example, Bhaya et al. (2007) compared the genomes of two cyanobacterial isolates (*Synechococcus* sp.) from microbial mats of Yellowstone hot springs with metagenomic libraries sequences from these mats at different temperature spaces. Differences in microbial diversity were revealed where lower temperature spaces showed higher diversity than those at high temperatures. Metagenomic sequences from Yellowstone hot springs have also been used to study phylogenetic relationships between *Roseiflexus* isolated strains and *in situ* populations inhabiting the same mats. The strains were closely related to various metagenomic sequences (Klatt et al., 2007)

Recently, the insertion sequence distribution of two termophilic microbial mat *Synechococcus* isolates was compared to the distribution present in the microbial mats' metagenomic data. Both distributions were equivalent indicating the cultured isolates are appropriate models for insertion sequences studies in this environment (Nelson et al., 2011). Moreover, the metagenomes of Arctic and Antarctic ice shelves have been compared to explore their similarities and differences in terms of major phyla and subphyla and their potential responses to environmental stress (Varin et al., 2011). Metagenomic shotgun sequencing of microbial mat communities from effluent channels of Yellowstone Mushroom and Octopus springs revealed the microbial predominant populations and their functional potential (Klatt et al., 2011). Furthermore, a novel crenarchaeotic group-genome present in a gold mine sub-surface microbial mat was re-constructed using metagenomic data (Nunoura et al., 2011).

1.2.7 ANTIBIOTIC RESISTANCE

1.2.7.1 Origins of Antibiotics and Antibiotic Resistance

In 1928, Alexander Fleming discovered the first natural antimicrobial named penicillin (Raper et al., 1944). Penicillin was used to treat bacterial infections during World War II and it was introduced as an antibiotic into clinical settings to treat human infections by *Staphylococcus* and *Streptococci* bacterial strains in 1946, having a vast impact on public health. Several years before the global use of penicillin, a penicillinase was identified in *Escherichia coli* (reported in 1940) and in *Staphylococcus aureus* (reported in 1944). After the widespread use of penicillin, resistant strains became prevalent and difficult to treat. The emergence of penicillin resistant genes before its wide application on clinical patients, and right after the introduction of the antibiotic, suggested that resistance genes are present in natural microbial populations without the need of being exposed to any antibiotic (Gaze et al., 2008).

While Fleming was working on penicillin, the sulfonamides were generated synthetically by Gerhard Domagk in 1937. The sulfonamides or sulfa drugs were effective synthetic antibiotics developed for the treatment of urinary tract infections, pneumonia and other conditions. After the use of sulfonamides, in the late 1930s resistant strains were reported with mechanisms that still prevail after 70 years (Davies and Davies, 2010). Several other antibiotics were discovered and introduced after these first antimicrobials.

The late 1940s and the early 1950s were recognized as the age of antibiotic chemotherapy. By this time, the antibiotic methicillin, a semi-synthetic version of penicillin, was developed in response to penicillin resistance. Also, broad spectrum antibiotics such as streptomycin, chloramphenicol and tetracycline were introduced against bacterial pathogens, intracellular parasites and the tuberculosis bacillus. Although these antibiotics were effective,

resistant strains began to appear. In 1953, a multi-drug resistant (MDR) strain of the dysentery bacilli was isolated during an outbreak of *Shigella* sp. in Japan (Bisht et al., 2009). Afterwards, in the 1970s a penicillin-resistant *Neisseria gonorrhea* strain and a ß-lactamase producing *Haemophilus influenzae*strain were isolated. By the late 1970s and 1980s a methicillin resistant *S. aureus* (MRSA) strain and re-surgence of the MDR *Mycobacterium tuberculosis* strain appeared in hospitals and became very difficult to treat. In the 1990s, strains of *Shigella* sp., *Salmonella* sp., *Vibrio cholerae*, *E. coli*, *Klebsiella pneumoniae*, *Acinetobacter baumanii* and *Pseudomonas aeruginosa* became antibiotic resistant. Recently, antibiotic resistant *Streptococcus* strains and more virulent strains of MRSA (due to the production of toxins) have been spreading to the community. Moreover, *S. aureus* and Enterococci strains are becoming resistant to vancomycin, the antibiotic designed to control multi-drug resistance (Alanis, 2005). Although antibiotic resistance has been extensively studied at the clinical level, resistance genes have been reported present in different types of environmental samples.

1.2.7.2 Environmental Antibiotic Resistance: The Resistome

Beneviste and Davies (1973) demonstrated that antibiotic resistance is not limited to clinical pathogens because it is widespread in environmental bacteria. They found aminoglycoside-inactivating enzymes in actinomycetes with identical properties to those found in antibiotic resistant bacterial clinical isolates. The most prominent resistant bacteria have an environmental origin, in habitats with low or no antibiotic selective pressure (Fajardo et al., 2008). Environmental non-pathogenic bacteria have shown to be potential reservoirs of resistance genes and have become a focus of research (Wright, 2010).

Several approaches have been directed to uncover the extent of the ***antibiotic resistome***. This concept refers to the collection of all the antibiotic resistant genes present in non-pathogenic

and pathogenic bacteria (Cantón, 2009). D'Costa et al. (2006) presented the first attempt to study the extent of the soil's resistome. A total of 480 spore forming bacterial strains were isolated and screened for resistance against 21 different types of antibiotics (natural products, semi-synthetic derivatives and synthetic compounds). All strains showed to be multi-drug resistant and there were identified resistance mechanisms known in pathogenic bacteria as well as novel ones. Dantas et al. (2008) also contributed with the unraveling of the soil antibiotic resistome. They isolated bacterial strains from 11 different soils that subsist on antibiotics as their sole carbon source. These bacteria were screened for resistance to high antibiotic concentrations (20 mg/L-1g/L) and they showed multi-drug resistance. The chemical modification and degradation capacity of antibiotic subsisting bacteria represent a reservoir for novel resistance strategies. Another important investigation was done by Demaneche et al. (2008) to determine the effect of transgenic plants on the amount of ß-lactamase resistance in soil. After isolating ß-lactamase resistant bacteria from soils with transgenic plants and from control soils (soils growing non-transgenic plants and non-agricultural soils), results concluded ß-lactamase resistance was prevalent in all the soils under study, transgenic plants having no effect on the resistome. These investigations were based on culture dependent methods which give access to only 1% or less of the microbial diversity in the environment. However, the advent of culture-independent methods has also uncovered novel mechanisms of resistance (Riesenfeld et al., 2004).

A classic culture-independent method for studying the resistome of an environment is the PCR amplification of antibiotic resistance genes, yet this technique allows only the amplification of known resistance genes (Wright, 2010). Functional metagenomics, pioneered by the Handelsman's Group (section 1.21), has offered a better alternative to access novel resistance strategies. This approach has led to the discovery of novel resistance mechanisms such as

aminoglycoside acetyl-transferases (Riesenfeld et al., 2004), inactivators of the anti-cancer antibiotic bleomycin (Mori et al., 2008), ß-lactamases (Allen et al., 2009) and efflux proteins (Lang et al., 2010). In addition to pathogenic bacteria and soil, antibiotic resistance has also been explored in aquatic environments, sub-surface settings and animals.

Antibiotic resistant bacteria have been isolated from natural aquatic environments such as fresh water (Sharma et al., 2009), mineral water (Messi et al., 2005) and marine water (Dang et al., 2008). These bacteria include indigenous species as well as exogenous strains that are occasionally present due to anthropogenic activities (Baquero et al., 2008). Resistant bacteria have also been isolated from wastewater treatment plants (Xi et al., 2009), hospitals wastewater, and animal production and aquaculture areas. Moreover, using different molecular methods (PCR, quantitative PCR, DNA hybridization and DNA microarray), antibiotic resistance genes have been detected in aquatic environments (Zhang et al., 2009). Antibiotic resistance goes beyond the surface of earth. Multi-drug resistant bacterial strains have been found between 173 m and 259 m beyond land surface at the Department of Energy Savannah River Site in South Carolina and Hanford Site in Washington. This high resistance is present without any selective pressure, which makes this environment one of great research interest (Brown and Balkwill, 2009). The animals, as components of the environment, have been associated with resistant bacteria as well. For example, several studies have shown that resistant bacteria can be isolated from birds and insects (Wright, 2010). Allen et al. (2009) applied functional genomics to the gut of a gypsy moth and identified several antibiotic resistant genes. Also, resistant bacterial strains have appeared in mammals such as rodents, wild boars and pig's gut (Kazimierczak et al., 2009). Furthermore, the human gut (Sommer et al., 2009) and oral microbiota resistance reservoir of healthy individuals has been explored. Most of the found resistance genes were novel and a small

subset was identical to the known resistance strategies of human pathogens (Díaz-Torres et al., 2006).

1.2.7.3 Aminoglycosides Discovery, Mode of Action and Resistance

The aminoglycosides are one of the oldest classes of antimicrobials and were the first therapeutic agents produced by bacterial fermentation. The first reports of aminoglycosides dates back when streptomycin and neomycin were isolated from *Streptomyces* spp. by the Waksman's laboratory team in 1944 and 1949, respectively. Afterward, Umezawa (1957) discovered kanamycin (produced by *Streptomyces kanamyceticus*) and the Weinstein's Group isolated gentamicin from *Micromonospora* spp. in 1963. Subsequently, in 1968 tobramycin was isolated from *Streptomyces tenebrarius* (Davies, 2006). These were naturally occurring substances produced by actinomycetes; however, after 1970 several derivatives of the natural products were developed as semi-synthetic aminoglycosides. For example, amikacin, bekanamycin, arbekacin and dibekacin are related to kanamycin. Similarly, netilcimin, sisomicin and isepamicin are related to gentamicin (Zembower et al., 1998). In general, all the aminoglycosides share a common antibacterial action.

The aminoglycosides exhibit antimicrobial activity because they bind to the 30S subunits of the prokaryotic ribosomes interfering with messenger RNA translation. Nuclear magnetic resonance spectroscopic analyses (Ogle and V. Ramakrishnan, 2005) and X-ray crystallographic structure studies of paramomycin have confirmed aminoglycosides bind to the 30S ribosome subunit in the region where codon-anticodon interaction occurs called the aminoacyl-acceptor site (A-site). This causes a decrease in translational fidelity because the ribosomes adopt a conformation that increases the selection of incorrect tRNAs to the A-site (Lynch et al., 2003). Although aminoglycosides have been effective to treat infections caused by

gram-negative and gram-positive bacteria, mycobacteria and protozoa, there are resistance mechanisms that counteract their function. Three basic mechanisms of resistance to aminoglycosides have been described: the decreased accumulation of the antibiotic inside the cell, the enzymatic modification of aminoglycosides and the methylation of ribosomal RNA.

The first mechanism is the adaptation of bacteria to reduce the entrance of the antibiotic to the cell and to facilitate the excretion of the antibacterial by the alteration of the anionic lipopolysaccharide cell surface and the use of transporters. Alterations in the transmembrane potential of aminoglycoside resistant *Staphylococcus aureus* strains have been evident. Also, *Pseudomonas aeruginosa* and *E. coli* decreased aminoglycoside uptake has been observed due to modifications in the cell's lipopolysaccharide structure (Zembower et al., 1998). Another resistance mechanism relies on the action of aminoglycoside modifying enzymes (AMEs) that change the antibiotic structure by suppressing their ability to bind to the target.

There are three classes of AMEs: *N*-acetyltransferases (AACs), *O*-nucleotidyltransferases (ANTs) and *O*-phosphotransferases (APHs). The AAC's catalyze the N-acetylation of aminoglycoside amino groups. The three-dimensional structures of four members of this class have been reported to have homologies with GCN5 superfamily of acyltransferases. The ANTSs add an AMP moiety to aminoglycoside hydroxyl groups. The encoding gene of ANT (2″) has been found widely distributed among pathogenic bacteria (Wright, 2005). This enzyme has been purified and its mechanisms of specificity have been reported. The *S. aureus* ANT (4′,4″) three-dimensional structure has been studied and reported. The APHs are responsible for the hydroxyl *O*-phosphorylation of aminoglycosides. Genes encoding these enzymes are distributed among multi-drug resistant bacteria. APH (3′)-IIIa has been extensively studied since it is the best example to understand this enzyme class mechanism of action. Three-dimensional structure

studies of this enzyme show its resemblance to Ser, Thr, and Tyr protein kinases, their mechanism and sensitivity to inhibitors (Wright, 2005). The latest described mechanism of aminoglycoside resistance is the methylation of ribosomal RNA. The action of RNA methyltransferases produces a steric barrier that impedes the binding of the aminoglycoside to the 30S ribosome (Fisher and Mobashery, 2010). This mechanism was first identified in *Citrobacter freundi* isolates from Poland in 2002, then in a *P. aeruginosa* isolated from Japan and in *Klebsiella pneumoniae* isolates from France in 2003. Ribosomal methyltransferases have also been found in other *Enterobacteriaceae* members such as *Serratia marcescens*, *Enterobacter aerogenes*, *Proteus mirabilis*, *E. coli*, *Klebsiella oxytoca*, *Shigella flexneri*, *Salmonella* spp. and *Acinetobacter* spp. The production of aminoglycosides methylases is associated with several gene families (including *arm*A, *rmt*A, *rmt*B, *rmt*C and *rmt*D) originated from actinobacteria: *Spreptomyces* and *Micromonospora* spp. Also, RNA methyltransferases have been linked to insertion sequences such as ISCR1, ISEcp1 and IS26 (Cantón, 2009).

CHAPTER 2
GENERATION OF LARGE-INSERT METAGENOMIC LIBRARIES FROM SUBTROPICAL HYPERSALINE BENTHIC AND EPHEMERAL MICROBIAL MATS

2.1 INTRODUCTION

There are methods for total environmental DNA extraction that can be classified as direct and indirect. These methods have special characteristics that must be taken into consideration before being used for metagenomic library (ML) generation and analysis of any environment. Such considerations include the amount of total DNA that can be extracted, its molecular weight and purity, the diversity of inserts sizes to be cloned, and the chance of isolating eukaryotic DNA. Direct DNA extraction methods involve the lysis of cells *in situ* within the sample using each or a combination of chemical and enzymatic solutions or mechanical forces such as bead beating, vortex, sonication and freezing and thawing, among others. Direct methods are known to allow the extraction of large amounts of DNA with usually low molecular weight and a higher diversity of inserts (Roose-Amsaleg et al., 2000). On the other hand, indirect extraction methods are based on the isolation of microbial cells and their subsequent enzymatic and chemical lysis away from the sample leading mostly to the isolation of high molecular weight DNA with a greater purity than with direct methods. However, in many cases lower amounts of DNA are recovered and a lower diversity of inserts are obtained when applying indirect methods because the extraction of cells may leave some groups of microorganisms underrepresented, depending on the type of sample. Despite these common disadvantages, previous studies suggest that higher amounts of DNA do not always give higher diversity. A comparison of a direct versus an indirect DNA isolation method applied to environmental samples with very different characteristics showed that, for some samples, the cell extraction-based method reduced the co-extraction of eukaryotic DNA and, for some samples, it yielded inserts with a higher biological diversity and complexity (Gabor et al., 2003).

An indirect DNA extraction method was employed in this study to obtain high molecular weight microbial mat (MM) DNA and generate large-insert metagenomic libraries. These libraries represent a collection of DNA inserts harboring more than one gene, possibly large contiguous genomic sequences that contain complete biosynthetic pathways (Liles et al., 2008). For this, large-insert metagenomic libraries are an excellent tool for the search of bioactive molecules with biotechnological, biomedical and industrial interest.

This chapter describes an indirect DNA extraction method used to isolate high-molecular weight DNA from benthic and ephemeral tropical hypersaline microbial mats. Also, the generation of large-insert metagenomic libraries from these ecosystems is presented.

2.2 METHODOLOGY

2.2.1 Sample collection and physical-chemical parameters determination

Microbial mat (MM) samples were taken from two saltern lagoons, Candelaria (benthic mat) and Fraternidad (ephemeral mat), at the Microbial Observatory in Cabo Rojo, Puerto Rico. One sample from each lagoon was taken at two different hydroperiods, the dry season (January-April) and the rainy season (August-December) [Casillas-Martinez et al., 2005]. The samples consisted of rectangular portions (15 cm x 15 cm x 10 cm) of microbial mat including all its characteristic layers (green, pink and black). The coordinates of the sampled sites were determined using a Global Position System (GPS). Salinity and temperature parameters were measured to the surface water on the mat at the sampling site with a refractometer and an alcohol thermometer (Fisherbrand, model No.14-997), respectively. Mats samples were transported to the laboratory at room temperature before an hour and the pH values were measured to a sample of microbial mats surface water using a pH meter (Orion, model No. 420) before processing the samples.

2.2.2 Indirect DNA extraction

2.2.2 a Extraction of microbial mats cells

A modification of the Liles et al. (2008) cells extraction method was employed as follows (Figure 2.1). Thirty grams of MM were homogenized in a blender (Warning Commercial, model 51BL31) with 150 ml of deionized distilled sterile water at low speed for 1 minute. After completing the blending step, the sample was incubated at 4 ºC for five minutes. These steps were repeated twice. The 150 ml of homogenized sample was distributed into four 50 ml sterile Falcon tubes and centrifuged (Eppendorf centrifuge, 5804R) (1000 x g) at 4 ºC for 15 minutes. Supernatants were transferred to a sterile Erlenmeyer flask and incubated at 4 ºC. Pellets of

microbial mat sample were resuspended in 50 ml of deionized distilled sterile water with vortex and the total volume of 100 ml was homogenized and centrifuged as described before. Supernatants were combined with the first batch and microbial mat sample pellets were discarded. Combined supernatants (200 ml total) were filtered through 8 ply sterile cheesecloth, poured into four 50 ml sterile Falcon tubes and centrifuged (15000 x g) at 4 ºC for 30 minutes. Pellets of microbial mat cells in each tube were gently resuspended in 50 ml of NaP Buffer (2% sodium hexametaphosphate, pH 8.5 adjusted with 1% sodium carbonate) and transferred to a clean Warning blender bucket. The 200 ml of Nap buffer with cells were homogenized three times at low speed with 5 minutes of incubation at 4 ºC between each run. Homogenized cells were poured into four 50 ml Falcon tubes and centrifuged (10000 x g) at 4 ºC for 30 minutes. The supernatants were discarded and cells in each tube were gently resuspended in 50 ml of Crombach buffer (0.33 M Tris-HCl, 0.001M EDTA pH 8.0) followed by centrifugation (10000 x g) at 4 ºC for 30 minutes. Cells were resuspended in 1.5 ml of Crombach buffer (0.375 ml in each tube), transferred to a tissue homogenizer (Kontes Glass Co. DUALL No. 23) and homogenized until cells clumps were not visible. Homogenized cells were poured into a sterile 50 ml Falcon tube and incubated at 4 ºC. Another thirty grams of microbial mat sample were processed as described above and combined with the first 1.5 ml of cells for an approximate final volume of 3 mls. To generate agarose plugs (named “worms” by the Handelsman’s group), 0.5 ml of 1.4 % low melting point (LPM) agarose (SeaPlaque® GTG® Agarose, FMC Bioproducts) (melted at 45 ºC) were mixed with 0.5 ml of microbial mats cells with vortex and transferred to a 1 ml syringe. Agarose plugs were cooled on ice for 5 minutes and stored at 4 ºC.

2.2.2 b High molecular Weight DNA Isolation

High molecular weight MM metagenomic DNA was isolated following the Liles et al. (2008) method (Figure 2.1). Worms containing microbial cells were incubated in a lysis buffer (0.01M Tris, 0.05 M NaCl, 0.20 M EDTA pH 8.0, 1% sarkosyl, 1% sodium deoxycholate and lysozyme 1mg/ml) for 3h at 37 ˚C and exposed to proteinase K in ESP buffer (1% sarkosyl, 0.5M EDTA pH 8.0 and proteinase K [1mg/ml]) for 24h at 55 ˚C. To inactivate proteins and enzymes after cells lysis, DNA "worms" were subjected to a 1mM phenylmethylsulfonyl fluoride (PMSF) treatment for 1h at room temperature. PMSF was discarded and fresh PMSF was added for an additional hour. To wash the PMSF, worms were submerged in TE 1X buffer (0.10 M Tris-HCl, 0.001 M EDTA pH 8.0) for 10 minutes at room temperature. TE 1X was discarded and the previews two steps were repeated twice. DNA was isolated by an electrophoresis where worms were integrated in a 1% agarose gel. Agarose fragments containing the DNA (named as "noodles" by the Handelsman's group) were cut out from the gel and treated with formamide for 24 h at 14˚C. To remove the formamide, noodles were dialyzed with gentle stirring in 1L of TE 1X for 48 h at 4 ˚C, adding fresh TE 1X every 24 h. The DNA was isolated from the noodles by an electroelution for 3h, 80V at 4 ˚C. At the end, a reverse current (80V) was applied for 1 minute and DNA was collected from membranes (Spectra/Por Dialysis MWCO: 12-14,000) and its concentration was determined spectrophotometrically using a biophotometer (Eppendorf Biophotometer No. 6131).

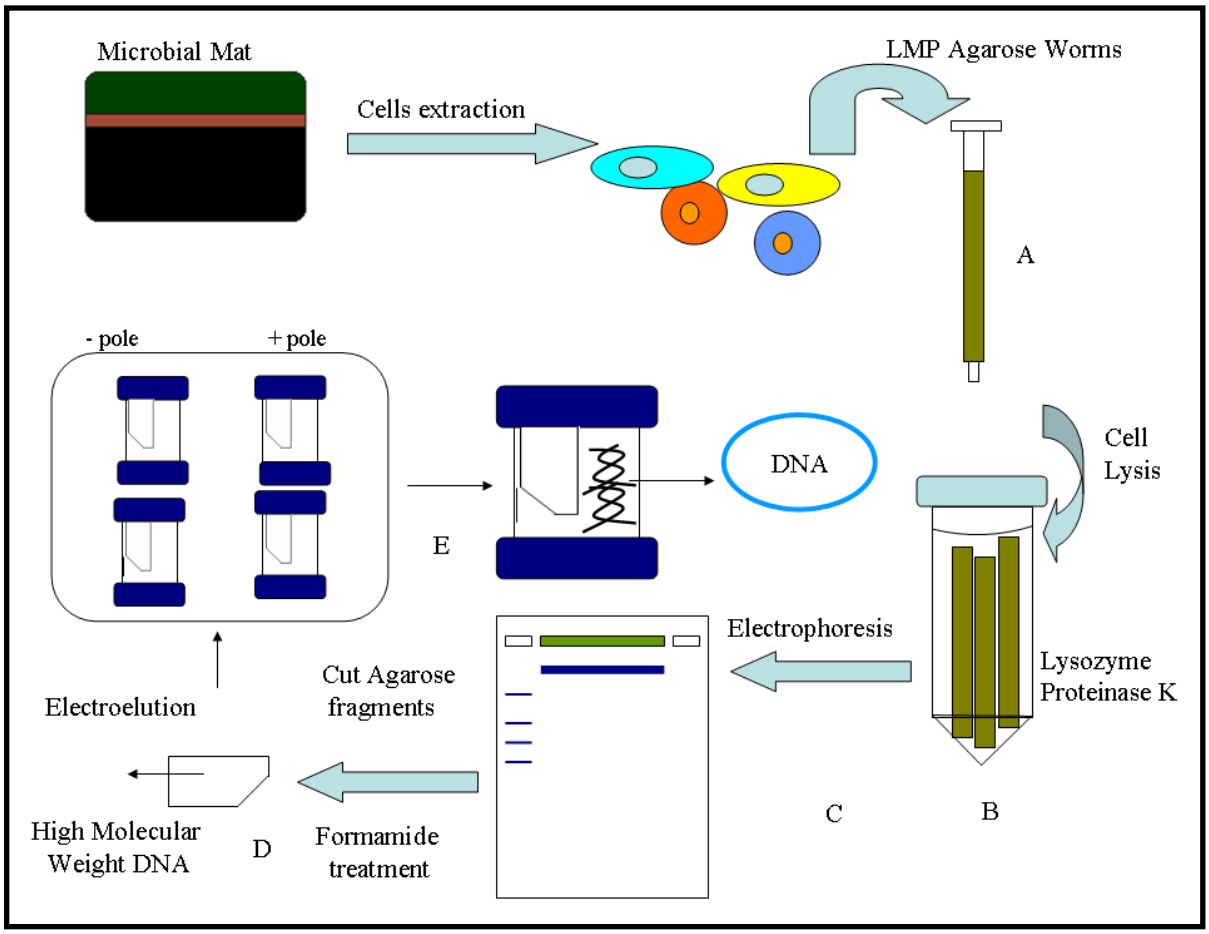

Figure 2.1 Cells-extraction based method for subtropical hypersaline microbial mat high molecular weight DNA extraction. Mats cells were embedded into agarose plugs **(**A), and chemically- enzimatically lysed within the agarose matrix (B). Total DNA was extracted by an electrophoresis of lysed "worms" (C) and noodles were subjected to denaturation of nucleases and proteins (D). High molecular weight DNA was isolated by an electroelution of "noodles" (E).

2.2.3 Generation of large-insert metagenomic libraries

High molecular weight DNA (20-40 Kbp) was cloned into a pCC1FOS (Epicentre) fosmid with the Copy Control Fosmid Library Production Kit (Epicentre) following the manufacturer's specifications (Figure 2.2). Briefly, total DNA was end-repaired (with an exo-endo nuclease) to generate blunt-ended inserts. End-repaired inserts of more than 20 kbp were selected by an overnight run electrophoresis in 1% LMP agarose at 35V, 4 ºC. Selected DNA

was excised in agarose fragments and purified by a GELase enzyme reaction followed by DNA precipitation, ethanol washes and resuspention. Insert DNA was cloned to the fosmid pCC1FOS in a blunt ligation reaction at room temperature for 2h. Recombinant fosmids were packed in vitro by lambda phage particles and transduced into the host strain EPI300 Phage T1-resistant. Clones were selected in Luria Bertani (LB) agar media with chloramphenicol (15 µg/ml). The clones were combined and collected as a mass of cells (by scrapping the colonies present in each Petri plate) to generate pools of clones in microtubes. Pools of clones were stored and preserved in LB media (20% glycerol) at -80˚C (Allen et al., 2009).

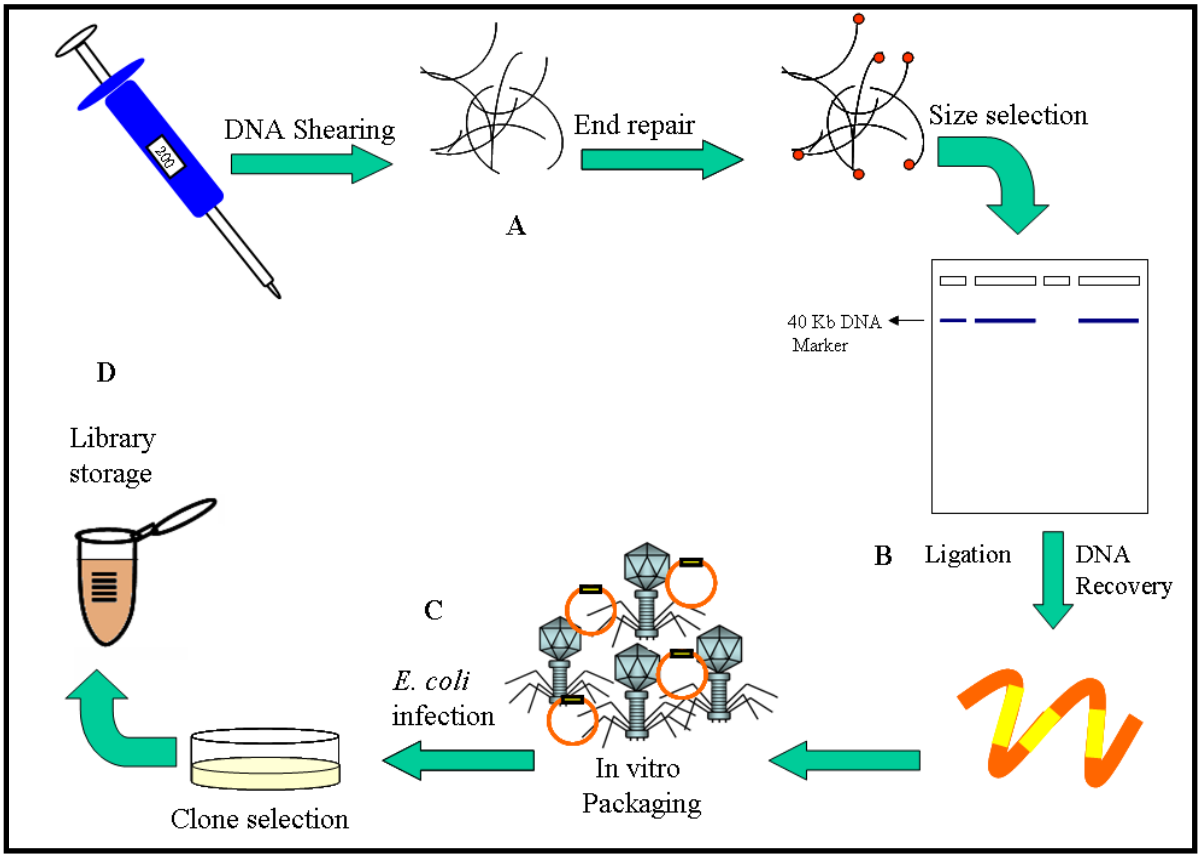

Figure 2.2 Generation of large-insert metagenomic libraries. Total DNA ends were modified to blunt-ends (A) and after size selection of inserts, a blunt ligation was performed. Recombinant DNA was packed in vitro and transduced to *E. coli* EPI300 Phage T1-resistant strain. Clones were selected and stored as pools (D).

2.2.4 Characterization of metagenomic libraries and clones

Metagenomic libraries were characterized by enumerating the total of clones and determining the presence of fosmids with inserts and the range of insert sizes. After enumerating the total clones, ten clones from each library were randomly picked for their molecular characterization. Clones were grown in LB media with chloramphenicol (15 ug/ml) and fosmid purification was performed using the QIAGEN® Plasmid Purification Kit. Fosmid DNA concentration was determined using a biophotometer (Eppendorf Biophotometer No. 6131-24631). Fosmids were subjected to a restriction analysis using the endonuclease *Not* I (Promega). The presence of inserts was determined by a 1% agarose pulse field electrophoresis in the CHEF Mapper System XA (BioRad). Non-vector DNA bands were identified as environmental inserts and their size was determined using the DNA markers 1 Kb ladder (Promega) and the MidRange Pulse Field Gel II (New England Biolabs).

2.3 RESULTS

2.3.1 Sampling sites and physical-chemical parameters determination

The sampling site's coordinates and physical-chemical parameters of the samples are documented in Table 2.1. Figure 2.3 illustrates the location of the subtropical hypersaline lagoons from where MM samples were collected. The characteristic colored layers of benthic and ephemeral MM were observed in the samples and are shown in Figures 2.4-2.5.

Table 2.1. Physical-chemical parameters of microbial mats at different seasons and the Global Positioning System coordinates of sampling sites.

Microbial mat	Season	Temperature (℃)	Salinity (%)	pH	GPS positioning coordinates
Benthic	Dry	34	25	7.09	17.95° N, 67.29°W
Ephemeral	Dry	36	25	7.21	17.98°N, 67.21°W
Benthic	Rainy	33	6	7.50	17.95° N, 67.29°W
Ephemeral	Rainy	35	5.5	7.16	17.98°N, 67.21° W

Figure 2.3 Geographic location of microbial mats sampling sites (http://maps.google.com/). A, B-Subtropical hypersaline lagoons. A. Candelaria lagoon, sampling site of benthic microbial mat. B. Fraternidad lagoon, sampling site of ephemeral microbial mat.

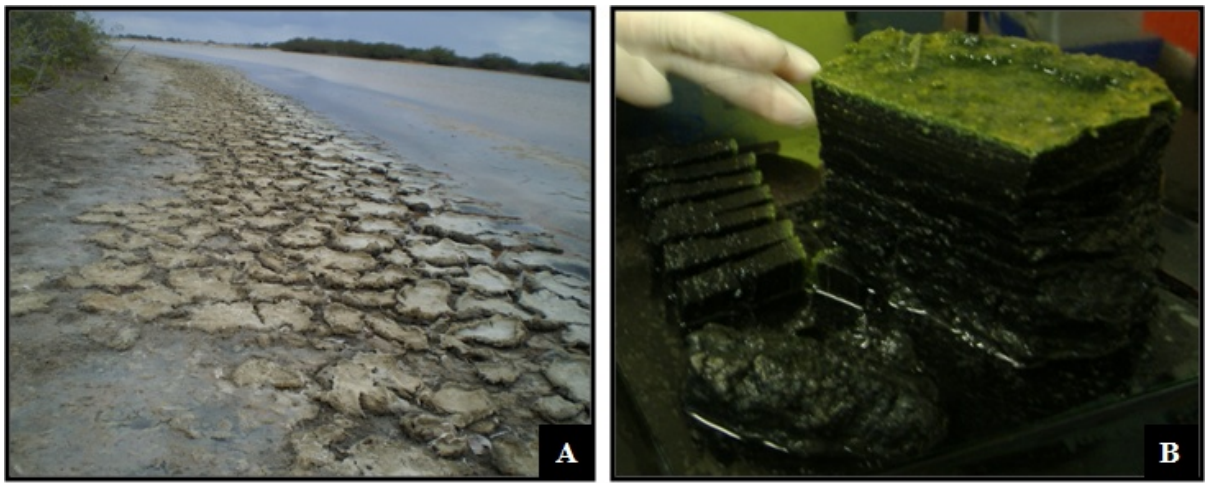

Figure 2.4 Sub-tropical hypersaline benthic microbial mats from Cabo Rojo Salterns in Puerto Rico (A). This mat has well defined layers and a compact structure (B).

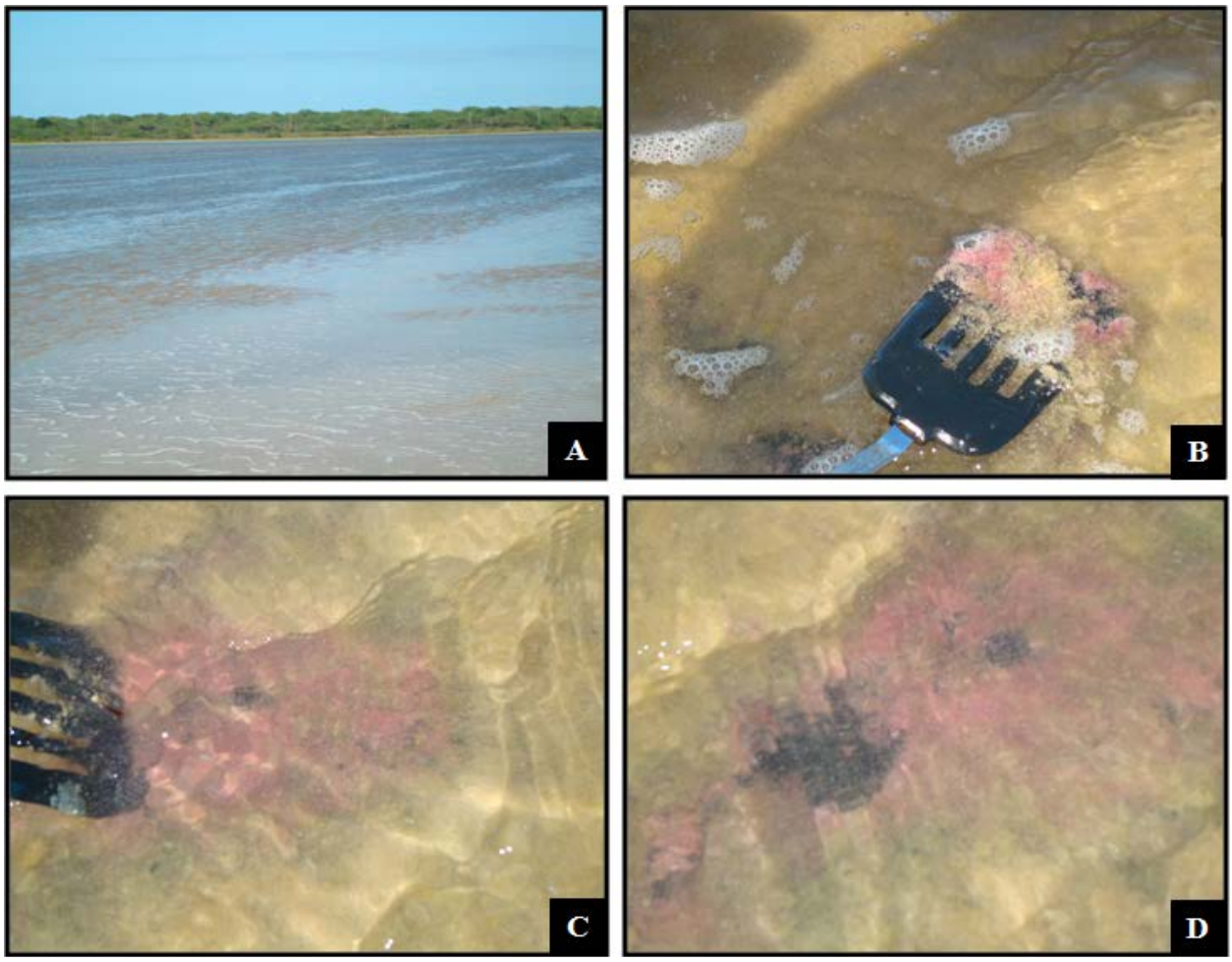

Figure 2.5 Tropical hypersaline ephemeral microbial mat from Cabo Rojo Salterns in Puerto Rico (A). This mat has a less compact structure than the benthic microbial mat; however, the layers are visible and have the characteristic colors: green, pink and black (B, C and D, respectively).

2.3.2 Indirect DNA extraction

2.3.2 a. Cells extraction and high molecular DNA isolation

Microbial cells were extracted from benthic and ephemeral mat samples, washed and embedded into agarose plugs ("worms") (Figure 2.6). Worms were lysed and high molecular weight DNA (DNA inserts of more than 20 Kbp) was obtained (Figure 2.7).

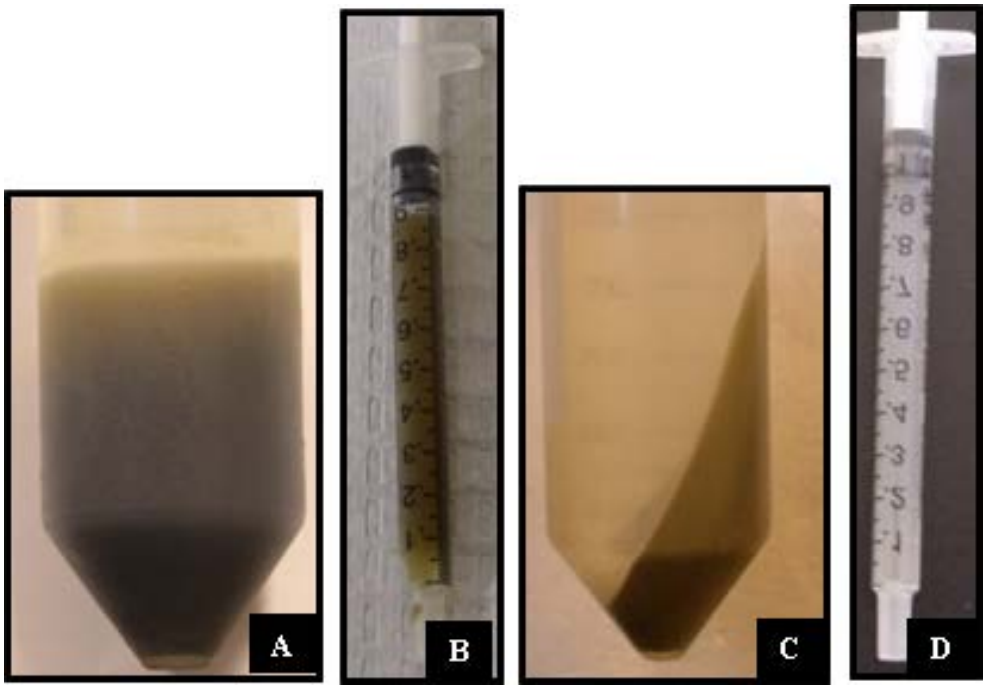

Figure 2.6 Samples of sediments and the generation of "worms". **A, C**- Sediments after homogenization and cells extraction from Benthic Microbial Mat and Ephemeral Microbial Mat, respectively; **B,D**- Cells embedded into agarose plugs from Benthic Microbial Mat and Ephemeral Microbial Mat samples, respectively.

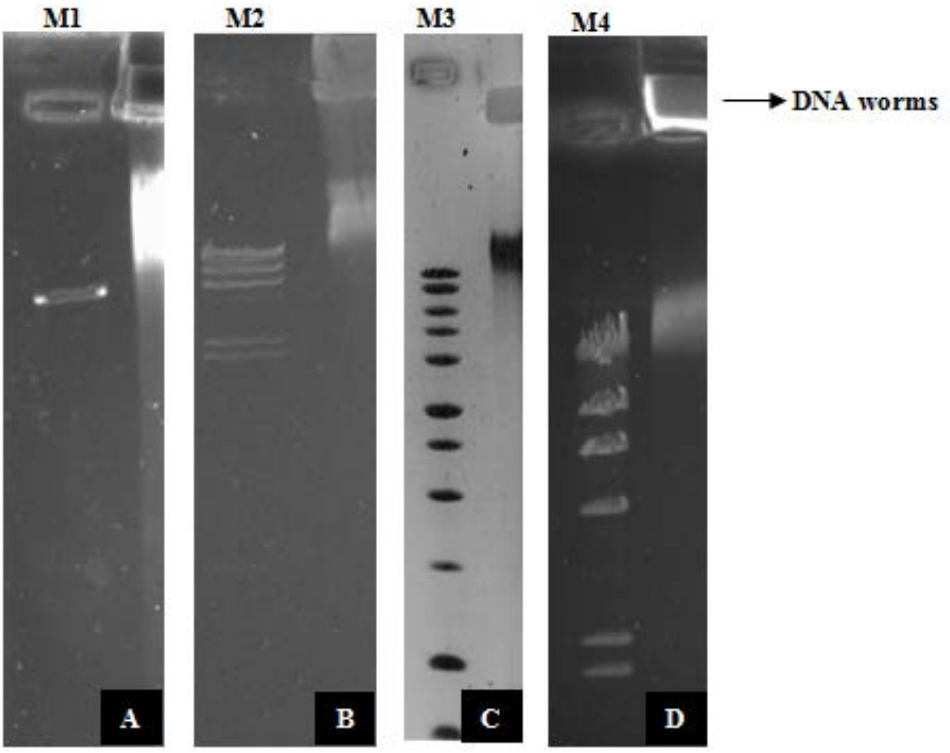

Figure 2.7 Verification of cells lysis and aproximation of extracted DNA molecular weight in a regular agarose gel for the four Microbial Mat samples under study. **A**- Benthic Microbial Mat Dry Season, **B**-Benthic Microbial Mat Rainy Season, **C**-Ephemeral Microbial Mat Dry Season and **D**-Ephemeral Microbial Mat Rainy Season. Molecular markers: M**1**- 36Kbp Lambda DNA, M**2,M4**- Lambda *Hind*III, M**3**- 1 Kb Ladder (Promega).

2.3.3 Generation and characterization of large-insert metagenomic libraries

Four large-insert metagenomic libraries were generated, one from each MM type during the dry and rainy seasons. These libraries were characterized by determining the number of clones and the insert size ranges. Metagenomic libraries from dry season samples consisted of 32,000 clones (benthic MM) and 30,000 clones (ephemeral MM). The molecular analysis of the randomly selected clones showed that 100 % of the tested clones have an insert with a size range from 20-100 Kbp (Figure 2.8, 2.9). Libraries from rainy season samples consisted of 1,200 clones (benthic MM) and 1,400 clones (ephemeral MM). The molecular analysis showed 100% of the tested clones have inserts within the 20-80 Kbp size range (Figure 2.10, 2.11).

Pulsed Field Electrophoresis of Dry Season benthic and ephemeral metagenomic libraries clones

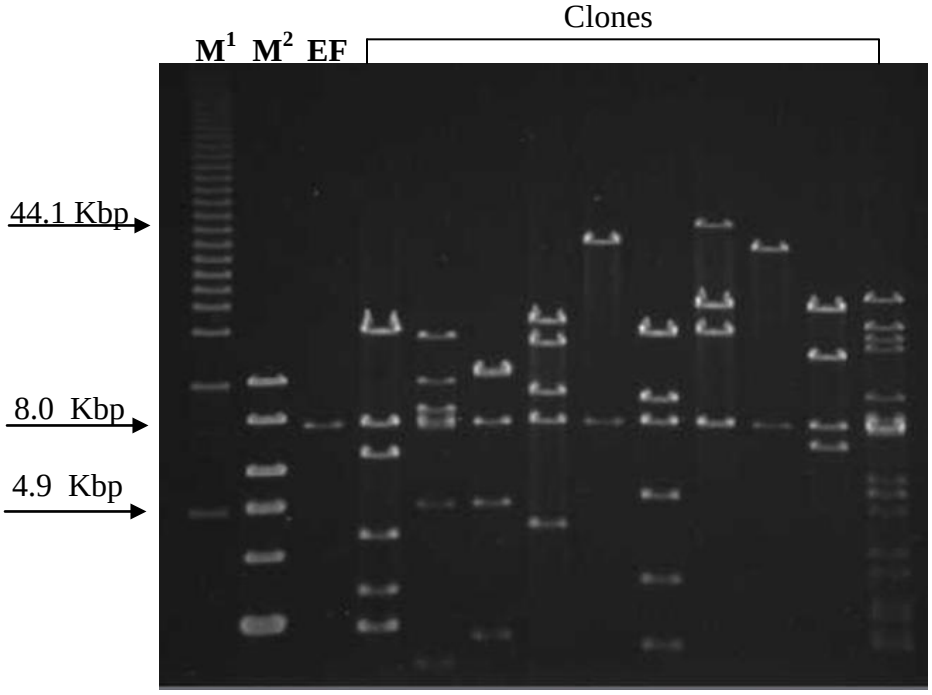

Figure 2.8. Pulsed Field Electrophoresis of Dry Season Benthic Microbial Mat clones *Not*I (Promega) restriction analysis. **M**[1]-5Kb Ladder (Bio-Rad); **M**[2]-1Kb Ladder (New England Biolabs); **EF**-Empty Fosmid.

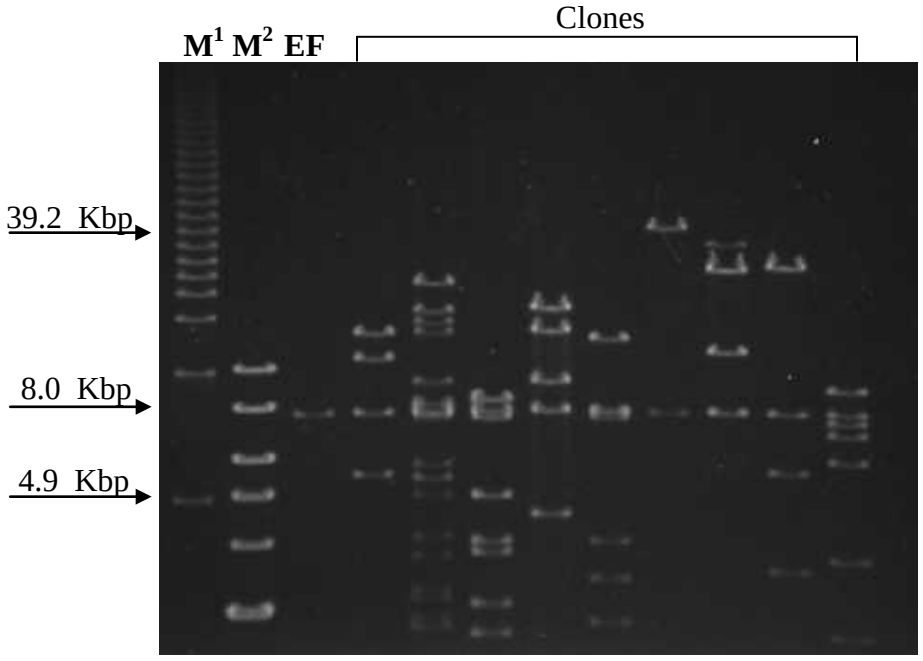

Figure 2.9. Pulsed Field Electrophoresis of Dry Season Ephemeral Microbial Mat clones *Not*I (Promega) restriction analysis. **M**[1]-5Kb Ladder (Bio-Rad); **M**[2]-1Kb Ladder (New England Biolabs); **EF**- Empty Fosmid.

Pulsed Field Electrophoresis of Rainy Season benthic and ephemeral metagenomic libraries clones

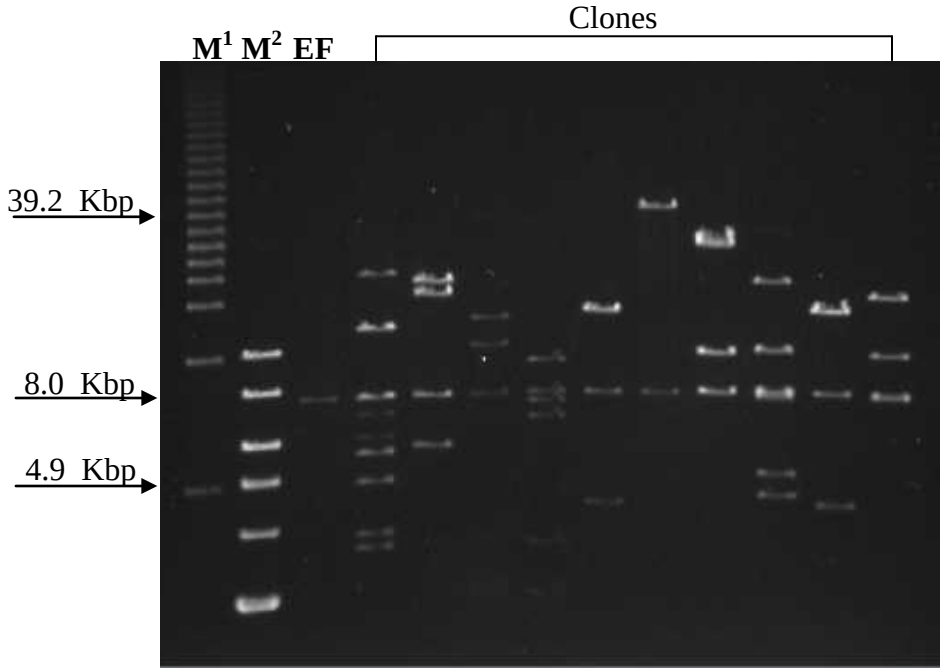

Figure 2.10. Pulsed Field Electrophoresis of Rainy Season Benthic Microbial Mats clones *Not*I (Promega) restriction analysis. M1-5Kb Ladder (Bio-Rad); M2-1Kb Ladder (New England Biolabs).

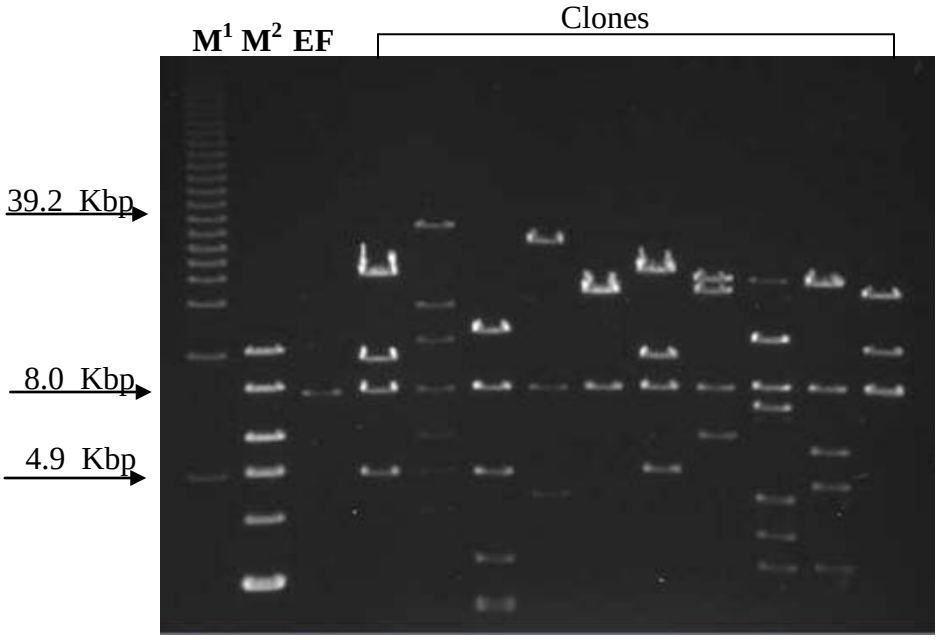

Figure 2.11. Pulsed Field Electrophoresis of Rainy Season Ephemeral Microbial Mats clones *Not*I (Promega) restriction analysis. M1-5Kb Ladder (Bio-Rad); M2-1Kb Ladder (New England Biolabs).

2.4 DISCUSSION OF RESULTS

This study consisted of the extraction of high molecular weight DNA from subtropical hypersaline microbial mats to clone large inserts and generate metagenomic libraries. An indirect extraction method was performed and DNA fragments of more than 20Kbp were isolated. The electrophoretic analysis of lysed "worms'' showed a discrete band of DNA (not visible smear) where clearly degradation and disruption was avoided at the most (Figure 2.7). Chemical and enzymatic lysis of cells embedded in agarose plugs allowed the DNA to remain intact in the matrix, minimizing its rupture and contamination with sediment impurities. Also, degradation of DNA by nucleases was controlled by protease, PMSF and formamide treatments.

DNA extraction from microbial mats is a challenging task because this ecosystem is rich in extracellular polymeric substances (EPS), salts and a high concentration of microorganisms with complex cell walls and membranes such as cyanobacteria (Decho, 2000). Therefore, the most implemented extraction methods in microbial mats molecular studies have relied on the direct extraction of DNA within the sample. Since direct methods yield low molecular weight DNA fragments, improved methodologies have recently been proposed for extracting high molecular weight metagenomic DNA from microbial mats. Bey et al. (2010) applied 4 different indirect methods of DNA extraction to hypersaline microbial mats samples. All consisted in the isolation of microbial cells and each one differed in the cell lysis strategy. The lysis mechanisms were beat beating (BB); freeze and thaw and the addition of β-mercaptoethanol and polyethylene glycol (PEG); freeze and thaw and a combination of SDS, lysozyme and proteinase K (LYS/FT); and rounds of freeze and thaw only (FT). Their conclusion was that the PEG method was the best mechanism to extract large metagenomic DNA fragments from hypersaline microbial mats based on the DNA size, concentration and purity, and the biological diversity indices obtained with 16S

rDNA analyses (such as denaturing gel electrophoresis and pyrosequencing of 16S rDNA amplicons). Their electrophoretic analysis showed high quality DNA bands from 20 Kbp to more than 48 Kbp. The indirect extraction method used in our study did not include the mechanic effect of freezing and thawing the cells; instead, the cells were embedded in an agarose matrix and lysed with a chemical/enzymatic approach. Similar to Bey et al. (2010), high molecular weight and quality DNA bands were obtained (Figure 2.7). This demonstrates the extraction method described here (chemical/enzymatic) is another alternative to overpass the challenge of obtaining large metagenomic DNA from hypersaline microbial mats. Unfortunately, the inserts size range after the DNA extraction obtained by Bey et al. (2010) cannot be compared with the results in this study because our extraction samples were not analyzed in a pulsed-field gel electrophoresis system. Also, the DNA extraction data presented here was not sufficient to make a parallel comparison between methods and conclude which would be the best option. Some suggestions would be to run total DNA extractions with BMM and EMM samples using the PEG and the chemical/enzymatic methods followed by a pulsed field electrophoresis of the DNA, the verification of its purity index and a 16S rDNA diversity study. The diversity study could include the use of “miniprimers” for the amplification of the 16S rDNA instead of longer primers applied in standard PCR methods. Isenbarger et al. (2008) applied this method to soil and microbial mat (benthic microbial mat under study) samples and demonstrated that “miniprimer ” PCR revealed novel 16S rDNA sequences that could not be detected with longer primers.

This study has generated the first metagenomic libraries from Cabo Rojo Salterns benthic and ephemeral microbial mats and, to the best of our knowledge, the first large-insert libraries from ephemeral hypersaline microbial mats in general. The restriction analysis of clones showed the libraries consist of diverse high molecular weight insert fragments from 20-100 Kbp (inserts

diversity was determined based on restriction pattern differences) regardless of the sampling season (Figures 2.8-2.11). However, rainy season libraries had a lower number of clones than dry season libraries. A possible explanation for this difference in the number of clones [i.e. 1,400 clones (rainy season EMM) versus 30,000 clones (dry season EMM) to] is that mats during rainy season could have a more elevated concentration of EPS than during dry season. The EPS has adhesive properties for which some of the functions that have been attributed to it are the sequestration of dissolved organic matter and contaminants, and it is suspected to increase the transfer of contaminants through the food webs (Hirst et al., 2003). Therefore, if present in DNA samples, the EPS could act as an inhibitor of the cloning process limiting the number of clones in the metagenomic libraries. A study of the EPS distribution in littoral sediments have found the EPS content is directly related to chlorophyll **a** and sediment water content (Cyr and Morton, 2006). Although littoral sediments are not comparable to marine environments, some correlations can be made. For example, a previous study of the benthic and ephemeral mats under study took transmission electron micrographs of each of its layers and observed copious amounts of EPS in dry and rainy season samples, yet it was not quantified. Rainy season mats had a thicker green layer stratum indicating a larger cyanobacterial population, which is one of the major producers of EPS. Also, it was found that rainy season mats had the greatest binding and trapping potential. Moreover, chlorophyll **a** quantifications were higher in rainy season samples than in dry season samples (Casillas-Martínez et al., 2005). An EPS extraction experiment should be performed to quantify the EPS concentration from mats at both seasons to sustain or neglect the hypothesis that rainy season samples have higher amounts of EPS.

CHAPTER 3
SCREENING METAGENOMIC LIBRARIES FOR ANTIBIOTIC RESISTANCE AND CHARACTERIZATION OF RESISTANT CLONES

3.1 INTRODUCTION

Metagenomic libraries (MLs) are studied by two major methods: sequence-based and function-based analyses. Both approaches are designed to unravel the potential biological and functional diversity present in the collection of microbial genomes of certain environments. Each analysis has characteristic goals and together they provide an overview of how microorganisms interact with each other and with the environment. For instance, the ecological role of microbes can be presumed.

The sequence-based screening of MLs involves the determination of the presence of specific genes in the clones' insert sequence by the use of hybridization probes or the sequencing of inserts. One major goal of this approach is the sequencing of clones whose inserts include phylogenetic markers. These studies have been useful to link groups of microorganisms with specific functions (Schloss and Handelsman, 2003). The random sequencing of metagenomic clones has also led to important genetic discoveries, although some researchers disagree about the utility of this effort. Moreover, undirected large-scale sequencing projects have contributed to the expansion of sequence databases and have enriched our understanding of uncultured microorganisms (Handelsman, 2004).

On the other hand, ML functional screening relies on the heterologous expression of the genes of interest by the host cell and the subsequent characterization of active clones. This approach has identified numerous biomolecules with potential applications in medicine, biotechnology and the industry. However, functional analyses have several limiting factors. For example, the host cell is required to be able to express the encoding gene or set of genes of the desired function. Function-driven studies also depend on the availability of assays for the functions of interest and, because usually the frequency of active clones is low, there is a need

for generating effective screening methods (Schloss and Handelsman, 2003). Two basic types of screening assays have been developed: non-selective and selective screenings.

Non-selective assays allow the identification of a physiological property usually by the observation of qualitative characteristics. In this type of assay, all library clones are able to grow on the media and have the same opportunity of showing their metabolic capabilities. Some examples of non-selective methods are the screening for antimicrobial substances production (Rondon et al., 2000), pigment production (by visual inspection) (Gillespie et al., 2002) and the expression of proteases (Waschkowitz et al., 2009). In contrast, selective methods are designed to allow only the growth of clones that express a function in particular and restrict the growth of the rest of the clones. Due to the low frequency with which biological activities could be expressed in a metagenomic library, selective methods facilitate the isolation of desired functions (Handelsman et al., 2002). Some selective assays include the growth of clones in extreme conditions such as pH, temperature, salinity and toxicity. Others are based on the addition of substrates (i.e. metals and antibiotics) to the growth media that the library host strain is not capable of resisting or degrading. Also, several selective screenings have been performed using minimal media and the addition of substrates of interest as the whole source of energy. Recently, efforts are being directed to design new selective assays [such as the one generated to identify novel antiporters (Majernik et al., 2001)] and highly sensitive screens to detect low levels of activity (Schloss and Handelsman, 2003).

Our research implemented a selective antibiotic resistance screening to the generated metagenomic libraries. This assay consisted in the inoculation of metagenomic clones on media with antibiotics at concentrations where the host cell is incapable of growing. Clones that showed to be resistant were further characterized to confirm the insert sequence as the encoding

unit for this phenotype. For this, retransformation experiments were performed followed by transposon mutagenesis using the Genome Priming System and sequencing.

Once an active clone is isolated, its insert sequence is characterized by an analysis *in silico* using genetic databases to determine whether it is a novel or previously described function. Sequencing large-insert metagenomic clones is quite laborious. To simplify the process, the clone of interest is mutagenized with a transposon carrying a selectable gene marker (i.e. antibiotic resistance gene) and primer sequences at the 5' and the 3' ends. The clone DNA and the transposon are exposed to a transposase to promote the transposon random insertion (Genome Priming System instruction manual, http://www.neb.com/nebecomm/manualFiles/manualE7100.pdf). *E. coli* is then transformed with the mutagenized clone DNA and mutated clones are collected. Subsequently, clones that are deficient expressing the function of interest are identified and isolated. The loss of function indicates the transposon insertion occurred inside the genetic environment responsible for its expression. Mutated clones are sequenced and the first set of sequences is used as template for further primer walking.

The purpose of this chapter is to present the screening of large-insert metagenomic libraries from benthic and ephemeral mats' metagenomic libraries and the characterization of resistant clones.

3.2 METHODOLOGY

3.2.1 Metagenomic libraries screening for antibiotic resistance

3.2.1a Antibiotic susceptibility test

In order to screen the libraries for antibiotic resistance, pools of clones were grown in LB media with chloramphenicol (15 μg/mL) until reaching an optical density (OD_{600nm}) of 0.2. Bacterial clones were washed three times with physiological saline solution (0.85 % NaCl), and inoculated in LB agar media containing inhibitory concentrations of kanamycin (20 μg/ml), gentamicin (50 μg/ml), ampicillin (100 μg/ml), tetracycline (10 μg/ml) and spectinomycin (100 μg/ml) for 24 h at 37 ˚C. From the resistant candidates, ten clones were randomly picked and preserved in LB medium (20% glycerol) at -80 ˚C. As a negative control for the antibiotic resistance screening, a clone with an empty pCC1FOS (no insert) was subjected to the same conditions as the metagenomic libraries clones. All the antibiotics were purchased from Sigma-Aldrich.

3.2.1b Minimum inhibitory concentration determination test

The minimum inhibitory concentrations (MIC's) of kanamycin, gentamicin, spectinomycin, tetracycline and ampicillin were determined for the control strain used in the antibiotic resistance screening (section 3.1a). Following a modification of the protocol described in Andrews (2001), the control strain's cells were washed three times with physiological saline solution (NaCl 0.85%) to remove the LB media with chloramphenicol (15 μg/mL) where it was previously inoculated. LB broth media with the antibiotics at concentrations from 0.25-1024 μg/mL were inoculated with 10^4 of washed colony forming units of the control strain. Also, the MIC test was performed to the resistant clones to verify up to what extent they could show resistance.

3.2.1c Characterization of resistant clones

The presence of inserts in resistant clones was confirmed by fosmid isolation and its restriction analysis with the enzyme *Not* I (Sigma-Aldrich), as described in section 2.4. Recombinant fosmids with unique restriction fragment length polymorphisms (RFLP) patterns were retransformed into the Electrocompetent isogenic strain TransforMax EPI300. To confirm the resistance phenotype was encoded by the environmental insert sequence, transformants were subjected to an antibiotic resistance screening with the same inhibitory concentrations of antibiotics as described in section 3.1a (Figure 3.1).

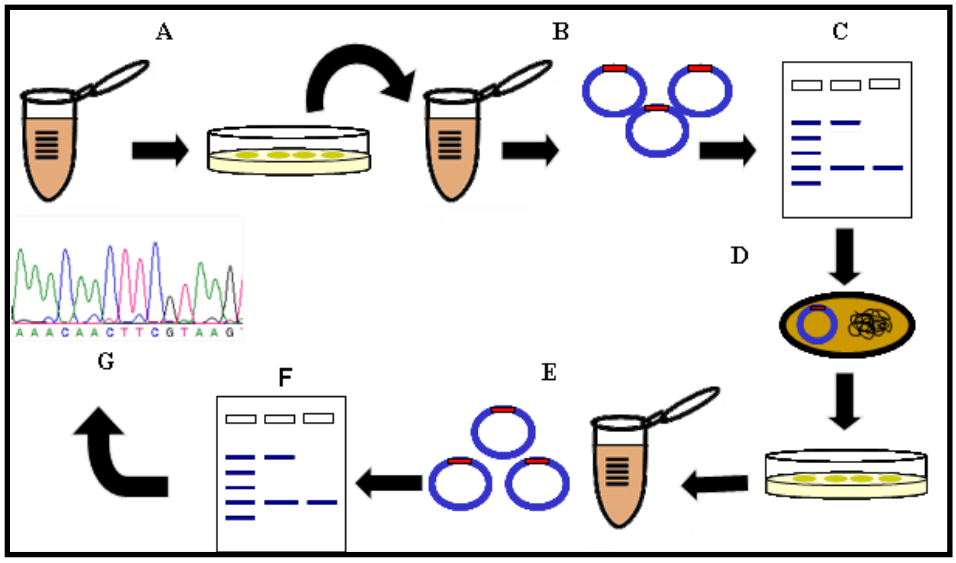

Figure 3.1 Metagenomic library antibiotic resistant screening. A- Growth of the library in medium supplemented with the antibiotic of interest. B-Selection of resistant clones. C-Restriction analysis of recombinant fosmids. D-Retransformation of an isogenic strain. E-Selection of resistant clones and extraction of fosmids. F-Restriction analysis. G- Inserts i*n silico* analysis.

3.2.2 Sequence determination of inserts conferring antibiotic resistance

3.2.2a Transposon mutagenesis of environmental inserts

Since sequencing large inserts is a challenging task, a transposon mutagenesis of recombinant fosmids was performed using the GPS-1® Genome Priming System (New England Bio-labs) (Figure 3.2). The transposons 5' and 3' sequence ends of this system were used as primers to determine the sequence of the gene or genes encoding antibiotic resistance. Following the manufacturer's instructions, recombinant fosmids were mutated with a transprimer kanamycin resistant gene. TransforMax EPI300 Electrocompetent cells were transformed with carrying a mutated fosmids. Mutations inside the insert sequence were selected by patching the clones in three different media in sequential order: an LBA plate with chloramphenicol (15 μg/ml) (patch #1), an LBA plate with chloramphenicol (15 μg/ml) and the antibiotic to which clones have shown to be resistant (patch #2), and an LBA plate with chloramphenicol (15 μg/ml) (patch #3). Mutants susceptible to the selectable marker in patch #2 were chosen for sequencing. Selected mutants fosmids were isolated and their concentrations were determined, as described in section 2.4. Sequencing primers, forward (Primer N 5'ACTTTATTGTCATAGTTTAGATCTATTTTG3') and reverse (Primer S 5'ATAATCCTTAAAAACTCCATTTCCACCCCT 3') with a concentration of 3.4 pmol/μL (supplied with the Genome Priming System Kit) and DNA samples with a concentration of greater than 200 ng/μL were sent to Macrogen USA Sequencing Center at Maryland, USA (http://www.macrogenusa.net/)

Instead of performing the GPS system to the kanamycin resistant clone, it was sent to be sequenced along with fosmid-ends specific primers: pEpiFOS Foward Sequencing Primer

5'GGATGTGCTGCAAGGCGATTAAGTTGG3' and pEpiFOS Reverse Sequencing Primer 5'CTCGTATGTTGTGTGGAATTGTGAGC3' (explained in section of results 3.2a).

3.2.2b Primer walking of inserts sequences

After performing the methodology described in section 3.2.2 a., sequences were used as templates to design primer oligos in order to extend the forward and reverse sequences (primer walking). Primers were designed with the web program Primer3 Input Version 0.4.0 (http://frodo.wi.mit.edu/primer3/) and each oligo's secondary structure was verified with the Oligo Analyzer 3.1 web program tool at Integrated DNA Technologies website (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/). Designed primers (10 pmol/µL) and DNA samples (200ng/µL) were sent to MacrogenUSA sequencing center at Maryland, USA. Each sequence was extended using the multiple alignment program Clustal X 2.0.11. After three "primer walking" rounds (Table 3.1, 3.2 and 3.3), the reverse and forward extensions of each mutated clone were overlapped into a complete sequence. For this, the reverse complements of the reverse sequences were determined using a DNA manipulation tool at the Sequence Manipulation Suite website (www.bioinformatics.org/sms/index.html).

3.2.2c Open reading frames identification and *in silico* analysis of insert sequence

The open reading frames (ORF) of the sequenced samples were identified using the ORF Finder tool of the online Basic Local Alignment Search Tool (BLAST) on the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/). The sequences' identities were analyzed using the amino acid sequence database (blastp) on NCBI.

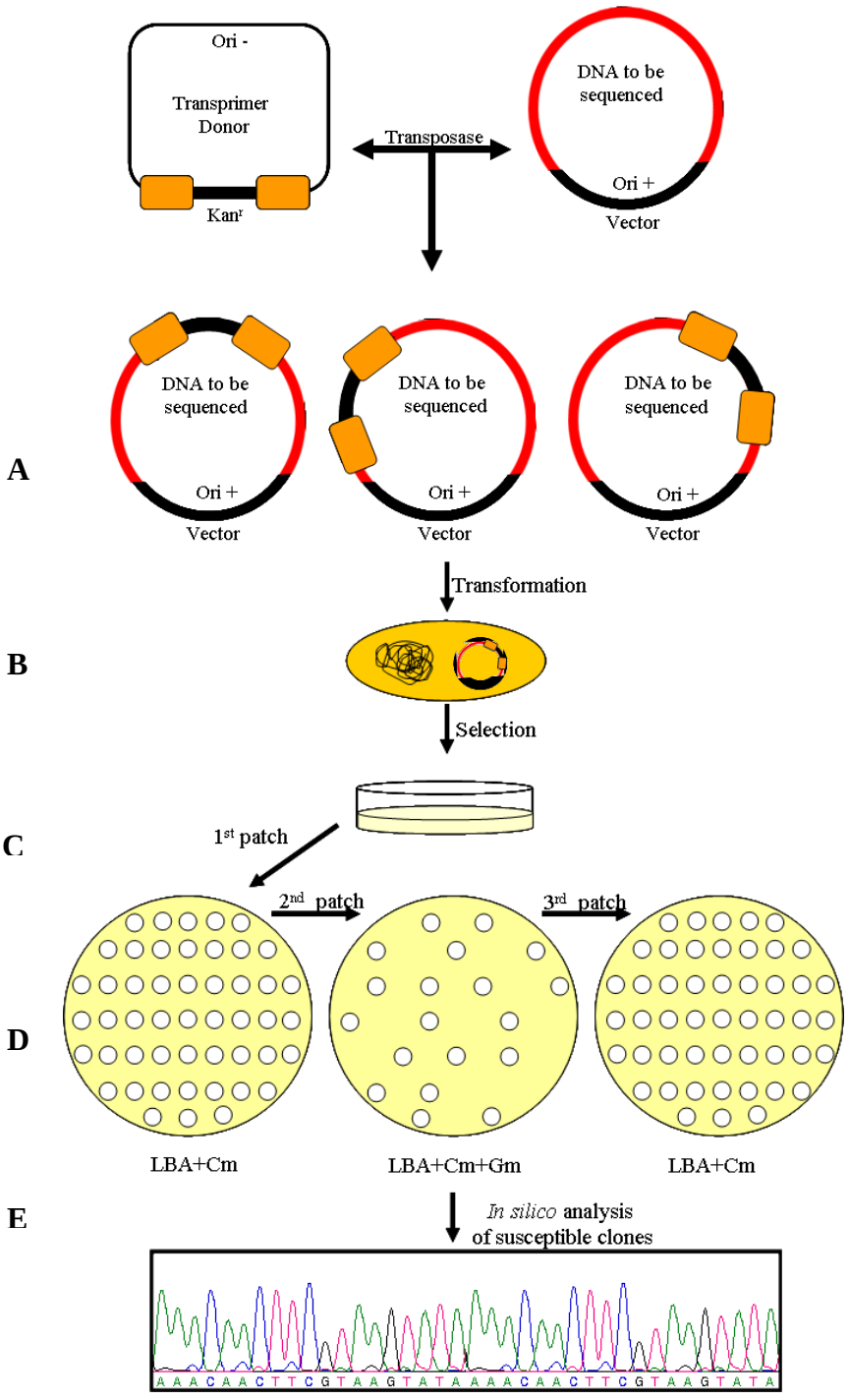

Figure 3.2. Sequence determination of antibiotic resistance large-insert by transposon mutagenesis and further sequencing of mutated clones. **A-** Transposon mutagenesis with a transposon carrying sequencing primers (forward and reverse) and a kanamycin resistance gene. **B,C**- Transformation of an isogenic strain with mutated clones and selection of transformants, respectively. **D-** Isolation of clones susceptible to the experimental antibiotic. **E-** Sequencing and *in silico* analysis of selected clones.

3.3 RESULTS

3.3.1 Metagenomic libraries screening for antibiotic resistance

3.3.1a Resistance screening test and characterization of resistant clones

From all the tested antibiotics, there were only clones resistant to gentamicin and kanamycin. The screening test was repeated for ampicillin, tetracycline and spectinomycin and no resistant clones were obtained from any of the libraries.

Gentamicin resistance

Gentamicin resistant clones (GRCs) were isolated from BMM dry season metagenomic library (Figure 3.3). The restriction analysis performed on ten randomly selected clones to verify the presence of inserts and its molecular weight showed a unique restriction pattern with an insert of approximately 44 Kbp in all the isolates, which resulted to be identical (Figure 3.4). The TransforMax EPI300 Electrocompetent strain transformed with the fosmid extracted from gentamicin resistant clones became resistant, confirming the phenotype was due to the insert. The negative control did not grow in LBA with gentamicin (50 µg/mL) (Figure 3.5).

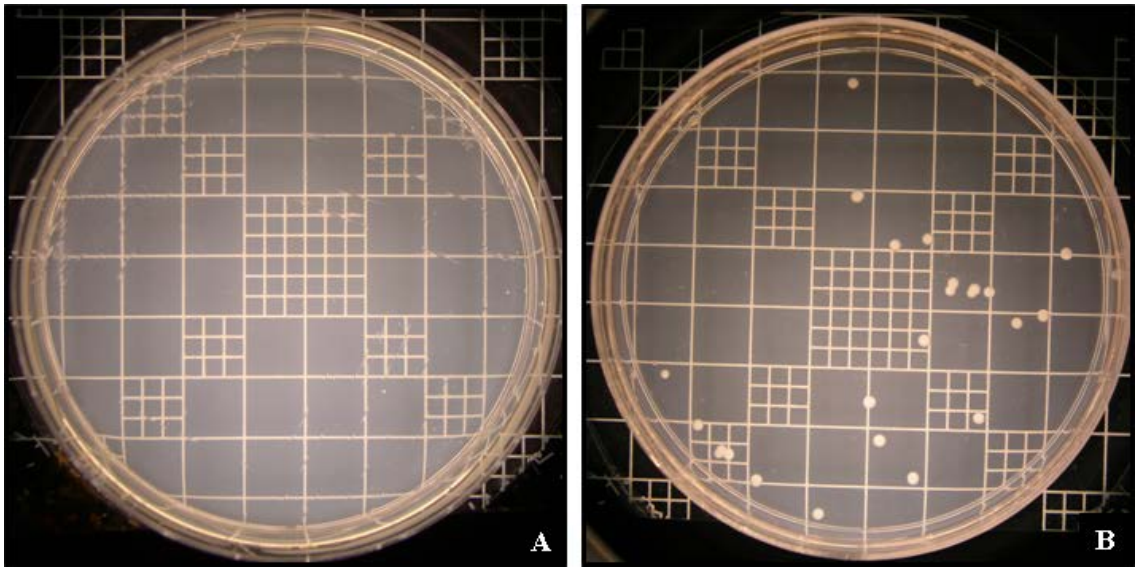

Figure 3.3 Gentamicin resistant clones (GRC) were isolated from benthic microbial mat rainy season metagenomic library. **A**- Negative control inoculated in Luria Bertani agar with gentamicin (50 µg/mL). **B**-Benthic microbial mat metagenomic library clones growing in Luria Bertani agar with gentamicin (50 µg/mL).

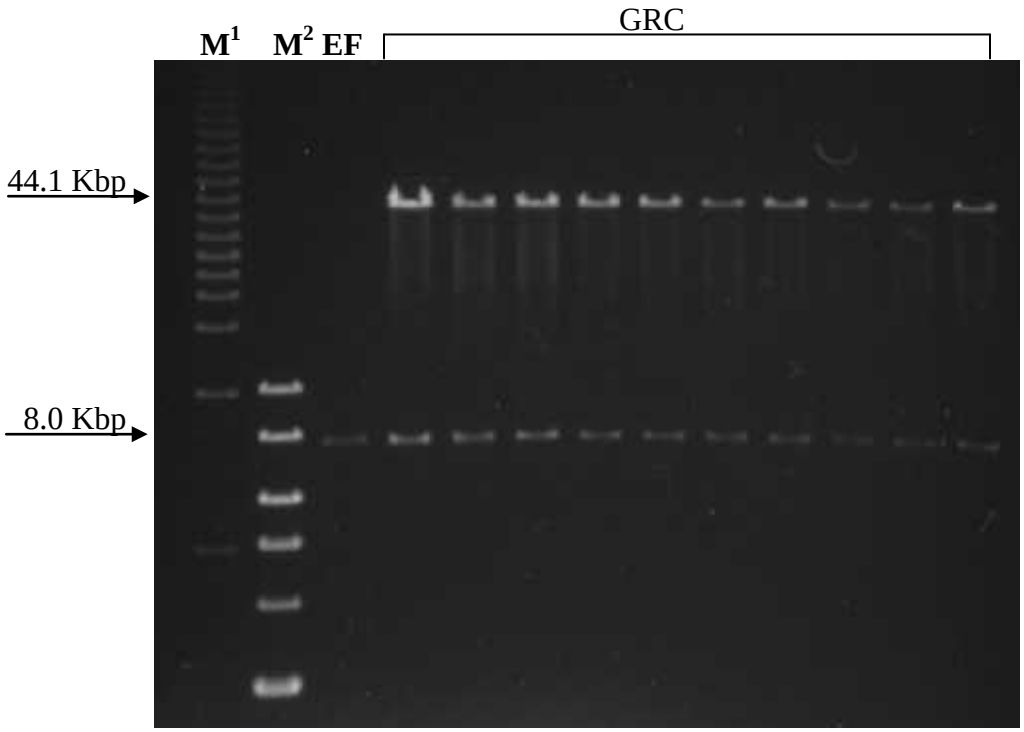

Figure 3.4 Pulsed field gel electrophoresis of the restriction analysis of tested gentamicin resistant clones (GRC) with *Not* I. **M**^{1}-5Kb ladder (Bio-Rad); **M**^{2}- 1Kb ladder (New England Biolabs); **EF**-empty fosmid.

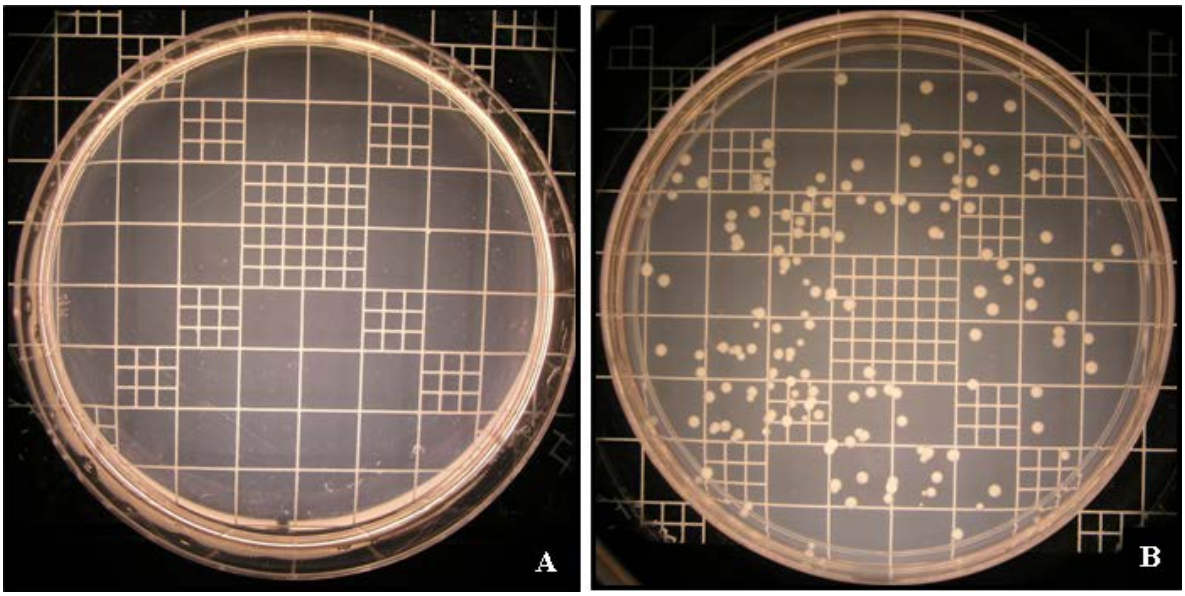

Figure 3.5. Transformation of the TransforMax EPI300 Electrocompetent strain with the fosmid extracted from gentamicin resistant clone. **A**-Negative control inoculated in Luria Bertani agar with gentamicin (50 µg/mL). **B**-Transformants growing in Luria Bertani agar with gentamicin (50 µg/mL).

The MIC value of the GRC for gentamicin was determined as described in section 3.2.1b.

It showed to have high-level resistance to gentamicin (≥ 512µg/mL) (Figure 3.6).

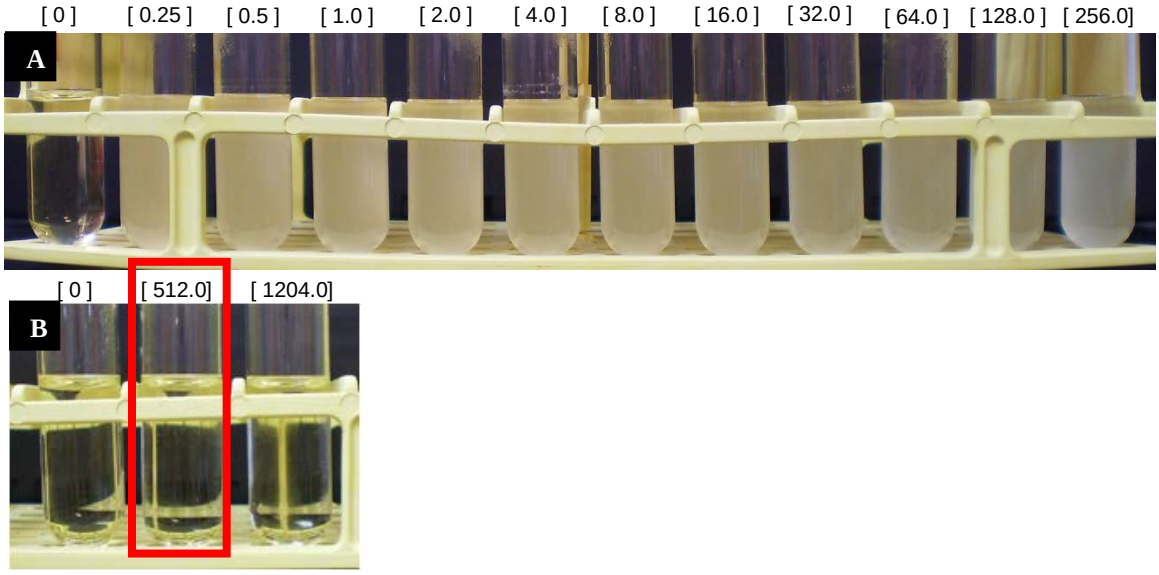

Figure 3.6. Minimum inhibitory concentration (MIC) of gentamicin for the Gentamicin Resistant Clone (GRC). **A**- The GRC inoculated in Luria Bertani Media with gentamicin from 0.25-256 µg/mL. **B**- The GRC inoculated in Luria Bertani Media with gentamicin from 512-1024 µg/mL. **Tube [0]**- had no gentamicin added and was not inoculated with the GRC. The GRC showed to have a high level resistance to gentamicin (MIC > 512 µg/mL).

Kanamycin resistance

Kanamycin resistant clones (KRCs) were isolated from all the tested metagenomic libraries (Figure 3.7). The restriction analysis performed on all the chosen clones showed a unique restriction pattern with an insert of approximately 39 Kbp (Figure 3.8). The TransforMax EPI300 Electrocompetent strain transformed with the fosmid extracted from kanamycin resistant clones became resistant as well, confirming the phenotype is due to a gene(s) within the clone insert. The negative control did not grow in LBA with kanamycin (Figure 3.9). The KRC showed to have high-level resistance to kanamycin with a MIC value > 1024 µg/mL (Figure 3.10).

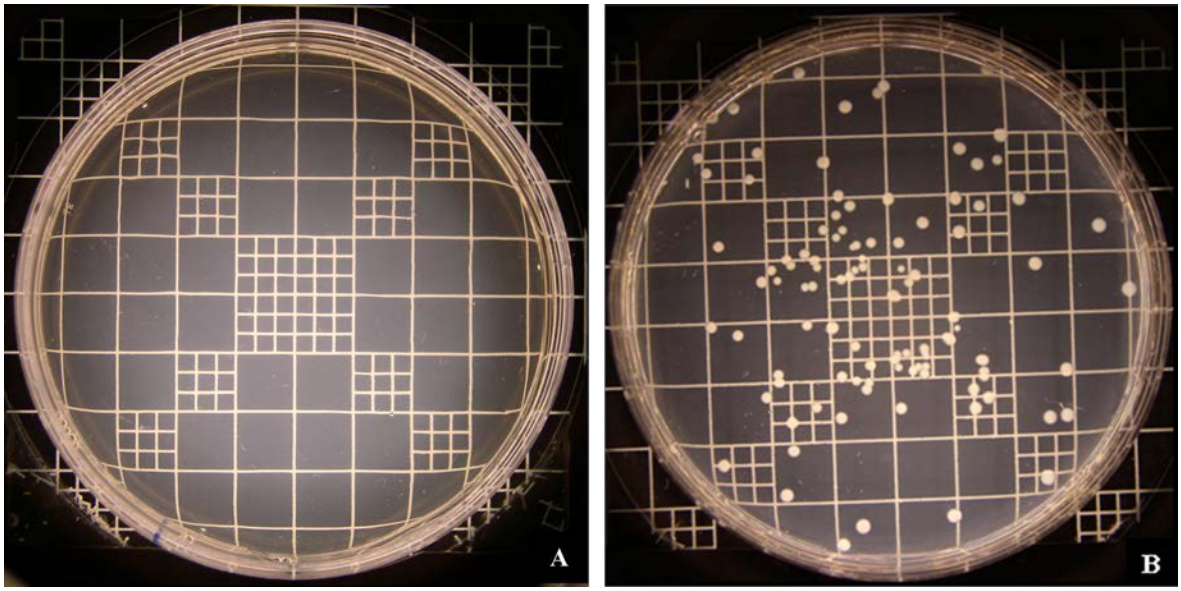

Figure 3.7 Kanamycin resistant clones (KRC) were isolated from the four tested metagenomic libraries. **A**-Negative control inoculated in Luria Bertani agar with kanamycin (20 µg/mL). **B**-KRC from the ephemeral microbial mat metagenomic library inoculated in Luria Bertani Agar with kanamycin (20 µg/mL).

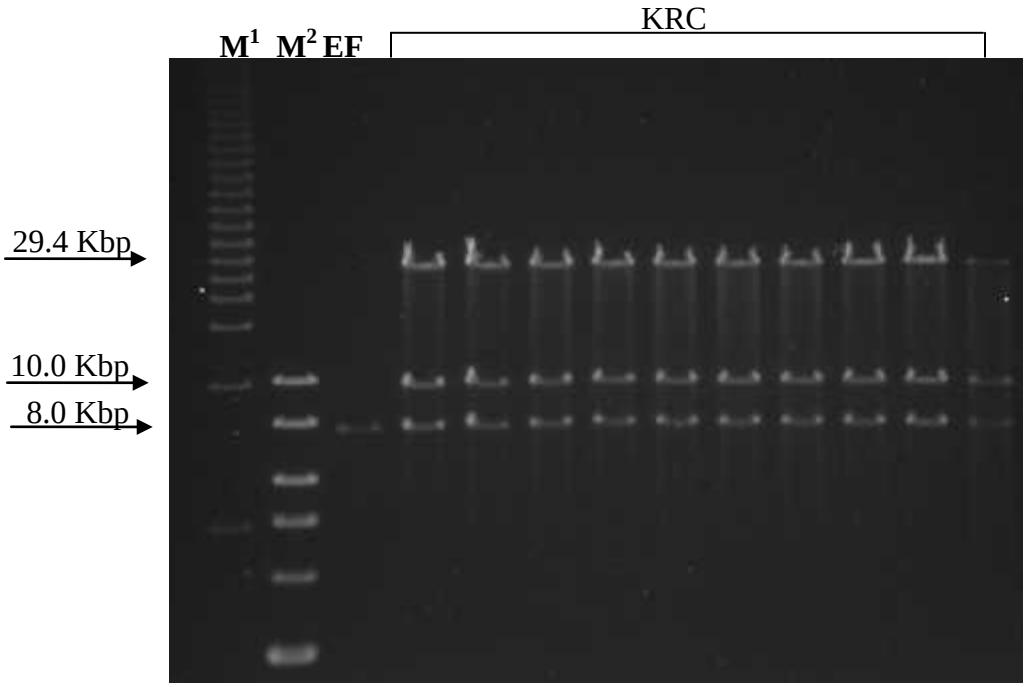

Figure 3.8 Pulse field gel electrophoresis of the restriction analysis of tested kanamycin resistant clones (KRC) with *Not* I. **M^1**-5Kb ladder (Bio-Rad); **M^2-** 1Kb Ladder; **EF**- Empty fosmid.

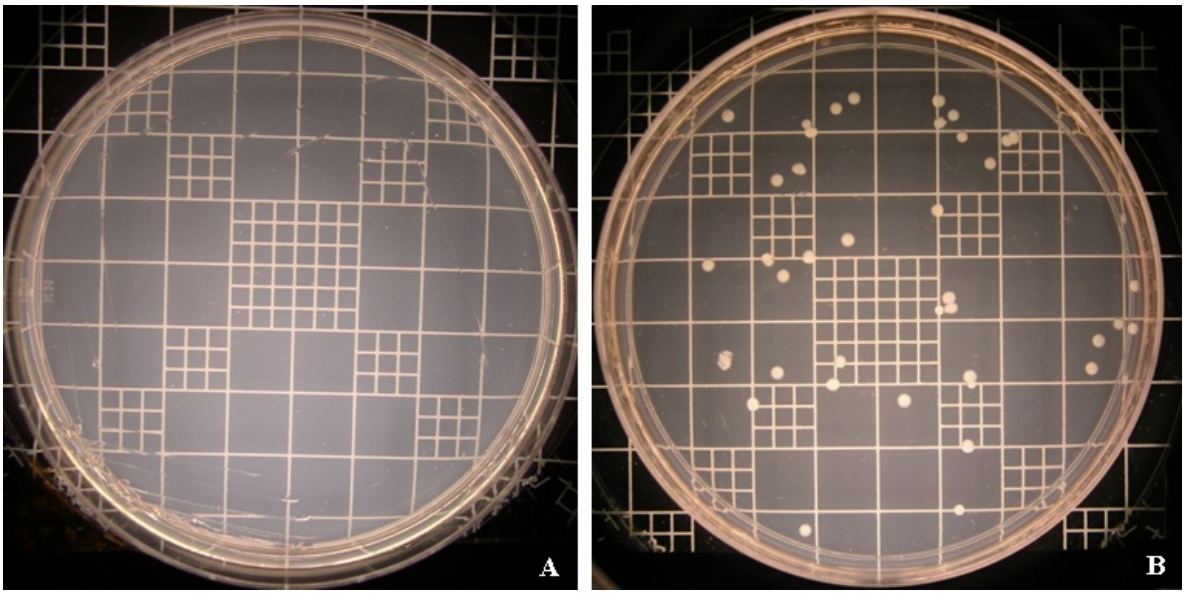

Figure 3.9 Transformation of the TransforMax EPI300 Electrocompetent strain with the fosmid extracted from the kanamycin resistant clone. **A**-Negative control inoculated in Luria Bertani agar with kanamycin (20 µg/mL). **B**-Transformants growing in Luria Bertani agar with kanamycin (20 µg/mL).

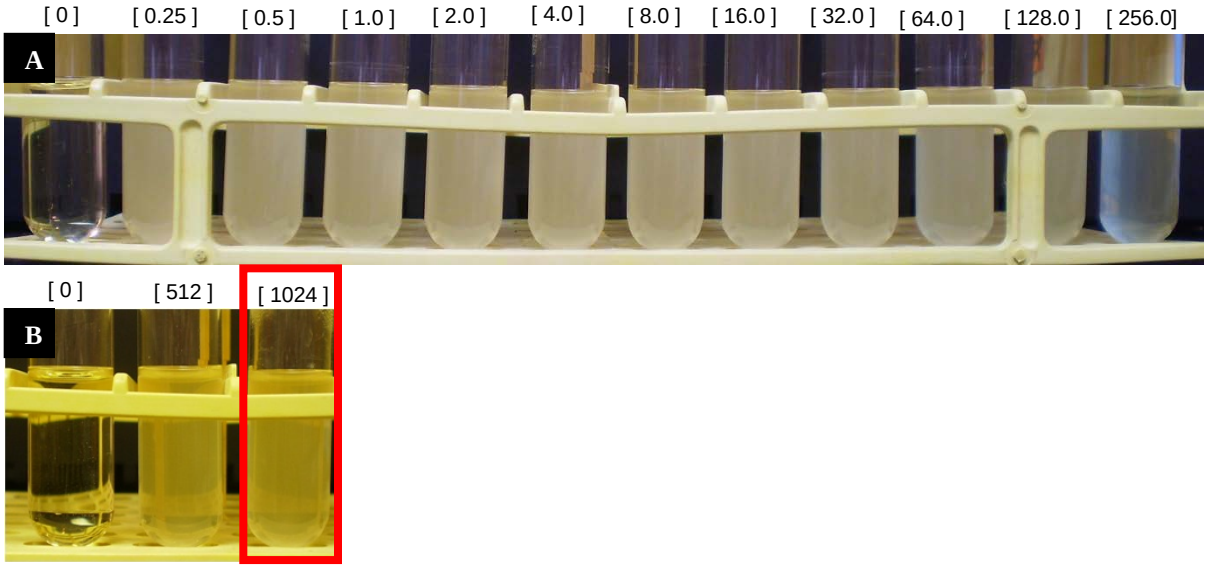

Figure 3.10. Minimum inhibitory concentration (MIC) of kanamycin for the Kanamycin Resistant Clone (KRC). **A**- The KRC inoculated in Luria Bertani Broth with kanamycin from 0.25-256 µg/mL. **B**- The KRC inoculated in Luria Bertani Broth with kanamycin from 512-1024 µg/mL. **Tube [0**]- had no gentamicin added and was not inoculated with the GRC. The KRC showed to have a high level resistance to gentamicin (MIC > 1024 µg/mL).

Aminoglycoside cross-resistance verification

Since the isolated clones were resistant to gentamicin and kanamycin, both aminoglycosides, a cross-resistance test was performed. The GRC and the KRC were inoculated on LB broth media with kanamycin (4.0 µg/mL) and gentamicin (2.0 µg/mL), respectively (MIC's for the control strain). A cross-resistant phenotype was observed with the GRC that showed to be resistant to kanamycin as well. The KRC was not able to grow on gentamicin. The MIC value for the GRC in kanamycin was determined and it showed to have high-level resistance (MIC > 1024 µg/mL) towards this aminoglycoside (Figure 3.11).

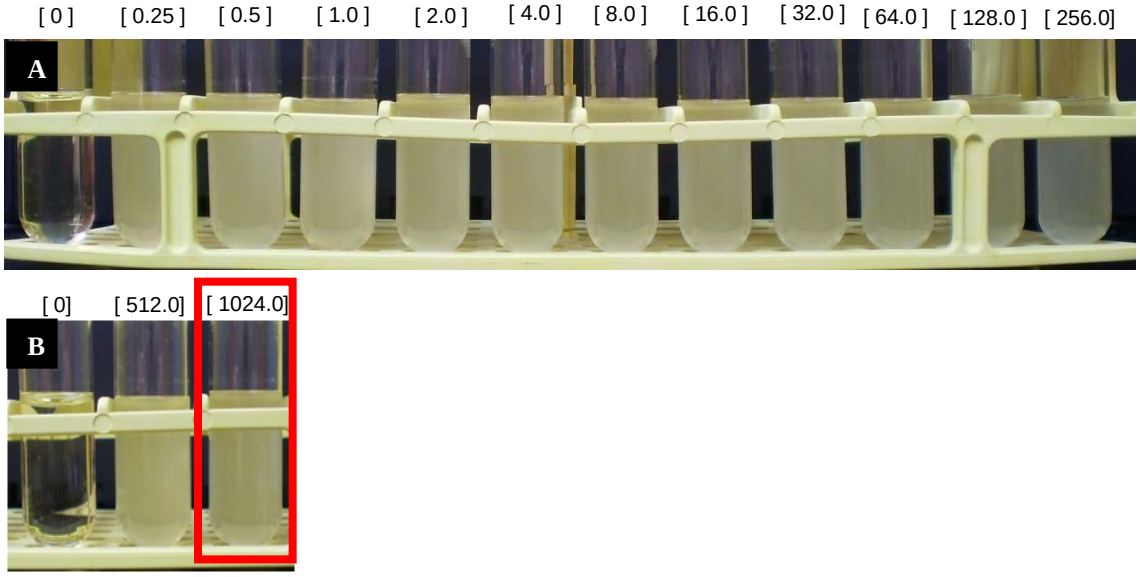

Figure 3.11. Minimum inhibitory concentration (MIC) of kanamycin for the Gentamicin Resistant Clone (GRC). **A**- The GRC inoculated in Luria Bertani Broth with kanamycin from 0.25-256 µg/mL. **B**- The GRC inoculated in Luria Bertani Broth with kanamycin from 512-1024 µg/mL. **Tube [0**]- had no kanamycin added and was not inoculated with the GRC. The GRC showed to have a high level resistance to kanamycin (MIC > 1024 µg/mL).

3.3.1b Minimum inhibitory concentration determination

The minimum inhibitory concentrations of gentamicin and kanamycin were determined for the control strain Epi300 Phage-T1 Resistant with an empty fosmid. For gentamicin, the MIC was 2.0µg/mL and for kanamycin 4.0µg/mL (Figures 3.12 and 3.13, respectively).

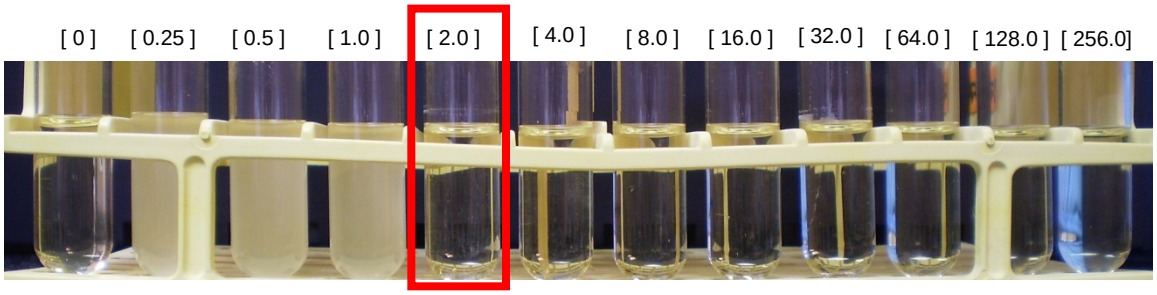

Figure 3.12 . Minimum inhibitory concentration (MIC) of gentamicin for Epi300 Phage-T1 resistant strain with an empty fosmid. This EPI300 strain showed to have a MIC of 2.0µg/mL for gentamicin.

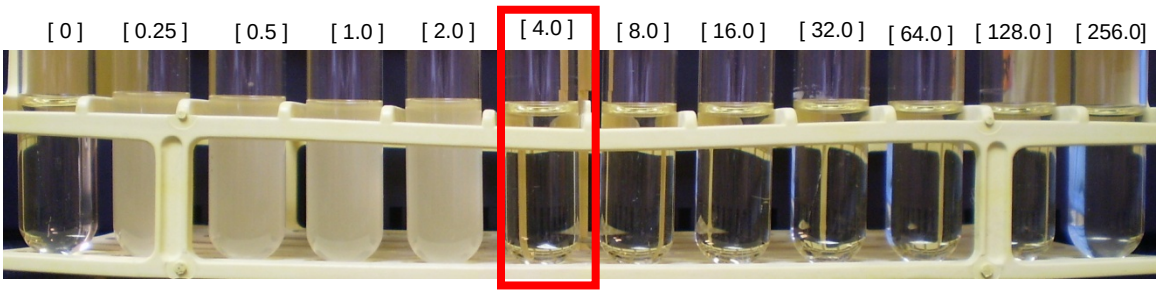

Figure 3.13 . Minimum inhibitory concentration (MIC) of kanamycin for Epi300 Phage-T1 resistant strain with an empty fosmid. The EPI300 strain showed to have a MIC of 4.0µg/mL for kanamycin.

3.3.2 Sequence determination of inserts conferring antibiotic resistance

3.3.2a Transposon mutagenesis of environmental inserts

The gentamicin resistant insert was mutagenized with the transprimer-2 (with a kanamycin resistant gene) of the Genome Priming System. After the transformation of TransforMax EPI300 Electrocompetent strain with mutated recombinant fosmids, a total of 164 clones were obtained and subjected to the patching test. Only 6 clones showed to be susceptible to gentamicin, indicating the transposon insertion took place inside the insert sequence (Figure 3.14). These clones were isolated for further sequencing and were denominated GPS4, GPS6, GPS8, GPS15, GPS21, and GPS27.

The Genome Priming System experiment could not be performed on the kanamycin resistance insert because the transprimer's selectable marker alternatives included a kanamycin resistance gene and a chloramphenicol resistance gene. Since the recombinant fosmid carries both of these genes, the selection of the mutations would have been impractical. Instead, the kanamycin resistant insert was sent for sequencing using the pEpiFOS forward and reverse sequencing primers. These primers allow sequencing from the fosmid ends into the cloned fragment. Unfortunately, after multiple sequencing trials the sequence of the kanamycin resistance insert could not be determined.

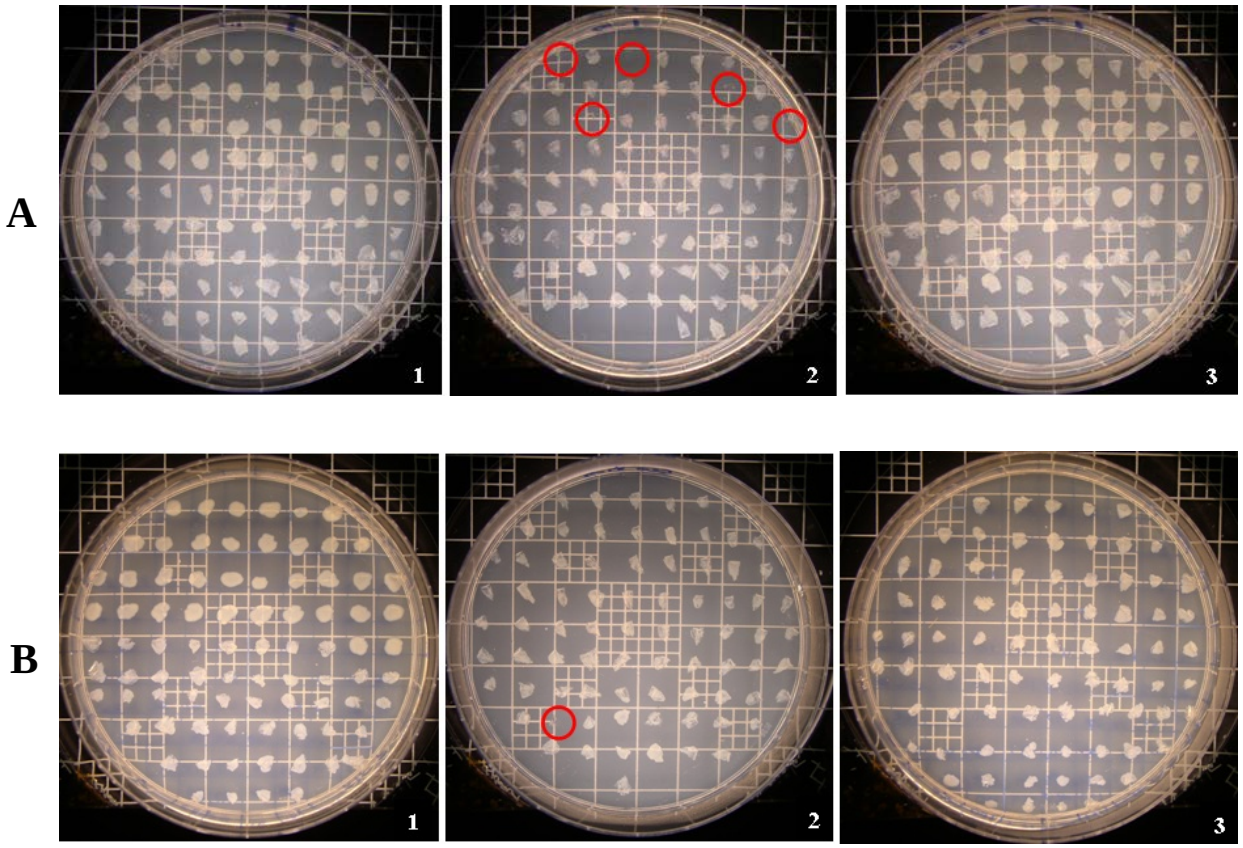

Figure 3.14 Patching of the transposon mutagenesis clones. **A**- Patching of the first 81 clones (1-81). **B**- Patching of the rest 83 clones (82-164). **A1, B1**- Master plates (patch #1): Luria Bertani agar media with chloramphenicol. **A2,B2**- Experimental plates: Luria Bertani agar media with chloramphenicol and gentamicin, **A3,B3**- Confirmation plates: Luria Bertani agar with chloramphenicol. The red circles indicate the clones that lost the resistance phenotype.

3.3.2b Primer walking and ORF's determination of the insert sequence

The mutated clone DNAs were sent to be sequenced with the primers provided in the Genome Priming System kit in order to obtain the first sequence extension in both directions: forward (to the right of the transprimer insertion) and reverse (to the left of the transprimer insertion). These first sequences were used as templates for generating the following primer sets to continue with the sequences extension. Three "primer walking" rounds were completed and the reverse and forward sequence of each clone was integrated into a complete sequence of approximately 4,500 bp. Five ORFs with the same direction of transcription were identified in

the insert sequence fragment (Figure 3.15). The transposon insertions were found to be positioned inside or in close proximity to specific genes associated with aminoglycoside antibiotic resistance.

GPS15 and GPS27 mutations were located inside the first ORF for which the in *silico* analysis suggested it encodes for a leucyl-tRNA synthetase. The analysis *in silico* suggested a 35 % of identity with the leucyl-tRNA synthetase of *Mycoplasma mobile*. GPS4, GPS8, and GPS21 mutations were inserted in the second ORF with a 45 % of identity with a putative tRNA guanine transglycosylase and a complete coding sequence of a ribosyltransferase-like gene (*rmtA*) of *Pseudomonas aeruginosa*. GPS6 mutation occurred inside the third ORF which shares a 30 % of identity with a ribosyl RNA methyltransferase from *Chloroflexus aurantiacus*. Downstream ORF3, two more ORFs were identified where no transposon insertion took place. The fourth ORF apparently encodes for a transposase with an ISX02-like domain that functions as an integrase and shares 72 % of identity with a transposase of *Francisella novicida*. The *in silico* analysis of the fifth ORF suggested the presence of a gene encoding an N-acyltransferase that shares a 28 % of identity with an uncultured prokaryote (Table 3.4).

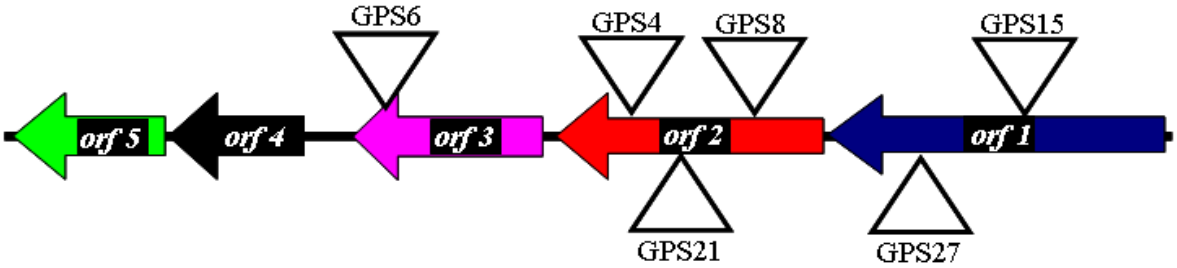

Figure 3.15 Apparent location and genetic environment of the genes present in the gentamicin resistance insert fragment sequence under study. *orf1*- leucyl-tRNA synthetase with 2 transposon insertions; *orf2*- tRNA guanine transglycosylase with 3 transposon insertions; *orf3*-ribosomal RNA methyltransferase with 1 transposon insertion; *orf 4*- tranposase; *orf5*- N-acyltransferase.

Table 3.1 Primers designed in this study for the first primer walking

Primer ID	Clone	Sequence 5' ⟶ 3'
GPS4-1R	GPS4	TGGGGTGCTTACTTCACAAG
GPS4-1F	GPS4	GAGTCTAGCCCTCGAATCACC
GPS6-1R	GPS6	GAGAGGGAAACATGGTCGTG
GPS6-1F	GPS6	GTGAAGTAAGCACCCCAAAC
GPS8-1R	GPS8	GCCACTATACGGATCGTGCTA
GPS8-1F	GPS8	GAGACGGTAGACATGCTCGT
GPS15-1R	GPS15	CGTTCACAAGTACCTCCTTCG
GPS15-1F	GPS15	GATGTTGGCTCCTTTTGCTC
GPS21-1R	GPS21	GGATGCCAAGTGAGATTGAG
GPS21-1F	GPS21	TGGGTTTGGGCAGTCATC
GPS27-1R	GPS27	CAGGCTCCTCAACATCCAAG

Table 3.2 Primers designed in this study for the second primer walking

Primer ID	Clone	Sequence 5' ⟶ 3'
GPS4-2R	GPS4	GGGACGCTTCTGAGATTGTG
GPS4-2F	GPS4	CCATGTCAGGAGCATTAGCC
GPS6-2R	GPS6	TGTGCGTAAGTCCTGCTGAG
GPS6-2F	GPS6	TGGGTTTGGGCAGTCATC
GPS8-2R	GPS8	CATTAGCCAGGCGTTTTGAG
GPS8-2F	GPS8	GGATGCCAAGTGAGATTGAG
GPS15-2R	GPS15	CTCCTCAAGGATATGGTCTGC
GPS15-2F	GPS15	GGCTTTCAGCTTCTGTCTC
GPS21-2R	GPS21	GAGAGGGAAACATGGTCGTG
GPS21-2F	GPS21	GCCACTATACGGATCGTGCTA
GPS27-2R	GPS27	CAGTTGCGTCTTTGGGATG
GPS27-1F	GPS27	CCCCAGAGCTTACTTAAACCA

Table 3.3 Primers designed in this study for the third primer walking

Primer ID	Clone	Sequence 5' ⟶ 3'
GPS4-3R	GPS4	TTGAAGGAGTCTGAGCAGGA
GPS4-3F	GPS4	CAGTTGCGTCTTTGGGATG
GPS6-3R	GPS6	GAGGGGAAAGTTCTTGAAGG
GPS6-3F	GPS6	GCTTTTGACTTTCGATCTAGCC
GPS8-3R	GPS8	CAGTTGCGTCTTTGGGATG
GPS8-3F	GPS8	GAGGGGAAAGTTCTTGAAGG
GPS15-3R	GPS15	ACCCGACAGGAAGCAAAT
GPS15-3F	GPS15	TGGGGTGCTTACTTCACAAG
GPS21-3R	GPS21	CCCGAAACAAAAACCTCTGC
GPS21-3F	GPS21	GGCTCTGTCGTATTTTGAGG
GPS27-3R	GPS27	GGGTTTATCCACAGACACCTC
GPS27-3F	GPS27	GGGACGCTTCTGAGATTGTG

Table 3.4. Open reading frames (ORF) found in the sequenced 4,500 bp insert fragment and their *in silico* analysis using blastp.

ORF Number	Predicted Function	Query Coverage (%)	E-Value	Maximun Identity (%)	Best-match organim
1	Leucyl-tRNA Synthetase	99	$1e^{-127}$	35	*Mycoplasma*
2	tRNA ribosyltransferase	95	$5e^{-114}$	45	*Pseudomonas*
3	16S rRNA methylase	97	$6e^{-33}$	30	*Chloroflexus*
4	Transposase	87	$1e^{-70}$	72	*Francisella*
5	N-acetyltransferase	85	$8e^{-17}$	28	Uncultured prokaryote

3.4 DISCUSSION OF RESULTS

The final objective of generating large-insert metagenomic libraries from subtropical hypersaline microbial mats was to use them as tools for the search of activity functions. Since mats are extreme ecosystems proven to have high biological diversity (Spear, 2003), there were great expectations of finding novel sequences with biomedical application potential. The antibiotic resistance screening performed to the BML's and the EML's led to the isolation of two antibiotic resistant clones, one resistant to gentamicin (MIC > 512 μg/mL) and one to kanamycin (MIC > 1024 μg/mL). The restriction enzyme analysis of clones followed by a pulsed-field gel separation showed unique patterns of DNA bands with diverse molecular weights, suggesting each clone is a different aminoglycoside resistance determinant alternative (Figures 3.4 and 3.8).

Aminoglycosides, in general, inhibit bacterial growth by the impairment of protein synthesis. Gentamicin and kanamycin are 4,6-disubstituted deoxystreptamine aminoglycosides that bind to the A-site of the 30S subunit of ribosomes interfering with mRNA translation and cause the incorporation of incorrect aminoacids to polypeptide (Park et al., 2008). Different resistance mechanisms have been reported for both antibiotics: the acquisition of antibiotic-modifying enzymes (present on plasmids and transposons) by environmental bacteria and clinical isolates and the methylation of ribosomal RNA as a self-protection strategy (Zembower et al., 1998). Gentamicin modifying-enzymes genes found in clinical isolates and in the environment include N-acetyltranferase clusters [*aac(3)-I*, *aac(3)-II/VI*, *aac(3)-III/IV*, *aac(6')-II/Ib*], a nucleotidyltransferase gene [*ant(2")-I*] and a phosphotransferase gene [*aph(2")-I*] (Heuer et al., 2002). Also, it has been reported that the gentamicin resistance methylase gene (*grm* and *grmb*) that encodes for a 16S rRNA methylase is originally found in gentamicin-producers from *Micromonospora spp.* Several encoding genes for kanamycin resistance that

have been reported are: *aac(3)-III*, *ant(2")-I*, *aph(3')-I*, *aph(3')-Iia*, *aph(3')-III*, *aph(3')-Iv*, *aph(3')-Vi*, and *aph(3')-VII* (Shaw et al., 1993). The kanamycin-producer *Streptomyces kanamyceticus* have also been found to express a 16S rRNA methylase [kanamycin resistance methylase (*krm*) gene], which protects the cell from the antibiotic action (Cundliffe and Demain, 2010). To compare the resistance determinants isolated from hypersaline microbial mats with the previously described genes, further characterization of clones included the determination of MIC's for the control strain (EPI300 Phage T1-resistant with an empty fosmid), the GRC and the KRC, and the analysis *in silico* of the resistance encoding sequences.

The MIC's for the control strain were 2.0 µg/mL (gentamicin) and 4.0µg/mL (kanamycin). The kanamycin MIC was higher than the value for the same control strain (2.0 µg/mL) reported by Donato et al. (2010). A possible explanation for this discrepancy is that they used a 96-well plate microdilution antibiotic susceptibility assay and the bacterial strains were incubated at 28 ºC. In contrast, a macrodilution assay was employed in our study and bacterial cells were incubated at 37 ºC. These differences, in surface area and temperature, could cause changes in bacterial growth behaviour, therefore, altering the susceptibility results. Similarly, it has been demonstrated that MIC values of a specific antibiotic for *E. coli* could also vary between strains (Monteiro et al., 2003). To our knowledge, there is no comparable report of the MIC value of gentamicin for the metagenomic libraries host strain. Using the MIC values for the control strain as a reference to determine the resistance capability of resistant clones, it was concluded that the GRC and the KRC both express high-level aminoglycoside resistance (MIC >256 µg/mL, Yan et al., 2004). Aminoglycoside high-level resistance has been associated with both the action of antibiotic-modifying enzymes (Abbassi et al., 2009) and the ribosomal

methylation by 16S rRNA methylase (Kharel et al., 2004). An analysis *in silico* of the gentamicin insert sequence was performed to identify the possible genes encoding for resistance.

The *in silico* analysis (protein query-protein database) of the sequenced 4500 bp from the GRC illustrated the presence of five different ORF's (Figure 3.15). After a protein-BLAST, all the ORF's except for one (the predicted transposase) had a maximum amino acid identity less than 45%, suggesting these could be novel sequences (Table 3.2) (Donato et al., 2010). Moreover, a nucleotide-BLAST of the ORF's failed to identify significant matches, which sustains that there are great differences at the nucleotide level between these sequences and the nucleotide database. The predicted transposase had a higher amino acid identity (72%) with *Francisella* sp. in the database; this was expected since transposases have highly conserved domains (i.e. ISXO2 domain). As an example, Vishnivetskaya and Kathariou (2005) found transposases from the gram-positive permafrost bacterium *Exiguobacterium* sp. to have highly conserved homologs in *Exiguobacterium* species from other diverse habitats. Two ORF's (ORF 3 and ORF 5) were predicted as possible gentamicin resistance determinants: a 16S rRNA methylase and a GCN5-related N-acetyltransferase, respectively. Although the expression of each gene could lead to aminoglycoside resistance, they encode for enzymes that exhibit different mechanisms of action. The question that remains is: which of the two ORF's is the actual coding unit for the gentamicin resistance?

One of the functions of GCN5-related N-acteyltransferases is the acetylation of aminoglycosides, conveying to resistance (Dyda et al., 2000). However, its presence does not imply it is involved in the resistance mechanism presented here. Recently, Donato et al. (2010) isolated a kanamycin-modifying bi-functional enzyme which consisted of two domains: an aminoglycoside acetyltransferase (at the N-terminal) and a GCN5-related acetyltransferase (at

the C- terminal). They determined the resistance phenotype was encoded only by the N-terminal domain and that the GCN5-related acetyltransferase does not utilize kanamycin as a substrate; there is a possibility this could be our case. In contrast, 16S rRNA methylase mechanism of action is directly involved in antibiotic resistance. This enzyme originates from aminoglycoside producers (i.e. *Micromonospora* sp. and *Streptomyces* sp.) and acts as a self-protection mechanism by the methylation of 16S rRNA at the positions where the antibiotic binds (Kojic et al., 2007). The data presented here was not enough to confirm which ORF is conferring the antibiotic resistance for which further experimentation should be performed to answer this question.

A possible experiment to identify which gene encodes for the resistance phenotype would be to generate primers in order to amplify the putative 16S rRNA methylase and N-acetyltransferase genes. Each individual amplicon and a selected plasmid could be digested with restriction enzymes and ligated to generate inducible constructs. Afterward, the EPI300 strain could be transformed with the resulting clones and then subjected to the antibiotic susceptibility test. Donato et al. (2010) used this method and successfully identified the encoding ORF of a bi-functional kanamycin-resistance enzyme. In the case that one or none of the clones carrying the amplified resistant genes express the resistance phenotype, another experiment could be performed. To verify if the 16S rRNA methylase gene (ORF3) is itself an encoding unit for antibiotic resistance, the ORF5 of the GRC could be inactivated by transposon insertion, followed by the transformation of EPI300 with this mutant and its subjection to the susceptibility test. Transposon insertions presented in this study apparently had a polar effect on downstream genes expression. Therefore, to determine if the N-acetyltransferase gene (ORF5) encodes for antibiotic resistance, it would be necessary to transform the EPI300 strain with the GRC

inactivated at the 16S rRNA methylase gene (ORF3) and with an inducible construct carrying ORF4 and ORF5 in trans. Moreover, it would be excellent to continue sequencing the insert fragment. This will verify if the N-acteyltransferase is a single enzyme or has a bi-functional mode of action and gain better understanding about the genetic environment of the resistance mechanism. Although it could not be determined in this study exactly which is the ORF responsible for the GRC resistance to gentamicin and kanamycin, the ORF's predicted functions, their approximated location to each other, the cross-resistance results and some data reported on literature suggest the 16S rRNA methylase could be the resistance determinant.

Initially, the 16S rRNA methylase expression was reported as a self-protection mechanism of resistance for the actinobacteria aminoglycoside producers. Then, a ribosomal RNA methylase (named RmtA) was found in *P. aeruginosa* conferring its high-level aminoglycoside resistance. The *rmtA* gene showed to be similar to the ribosomal methylase gene found in aminoglycoside producers and it was suggested to be carried by the transposon Tn5041. This sustains the possibility of intergeneric lateral transfer of resistance genes from aminoglycoside-producing bacteria to *P. aeruginosa*. This movement of resistance genes could have been promoted by the increasingly clinical use of arbekacin, a novel kanamycin-related aminoglycoside, which is rarely inactivated by antibiotic-modifying enzymes (Yokoyama et al., 2003). Our data showed the presence of a transposase encoding ORF, suggesting the 16S rRNA methylase gene could be associated with a transposon from a resistance plasmid (R-plasmid). Several types of plasmid-mediated 16S rRNA methylases have been discovered since 2003: *rmtA* (*P. aeruginosa*), *rmtB* (*Serratia marcescens*), *rmtC* (*Proteus mirabilis*), *rmtD* (*P. aeruginosa*), *rmtE* (*E. coli*), *armA* (*Enterobacteriaceae*) and *npmA* (*E. coli*). Today, these resistance genes are globally distributed in clinically important Gram-negative bacteria which are of human health

concern (Zhou et al., 2010). The ribosomal methylase found in subtropical hypersaline microbial mats is downstream an ORF that encodes for a queuine-tRNA ribosyltransferase as observed in *rmtA* and *rmtD* genetic environments [Yamane et al. (2004) and Doi et al. (2008)]. The queuine-tRNA ribosyltransferase gene has been related to the gentamicin biosynthetic gene cluster, which supports that 16S rRNA methylases come from amionoglycoside-producing microorganisms (Ayra, 2007). The cross-resistance result where the GRC also showed high-level resistance to kanamycin (> 1024 μg/mL) sustains the possibility that the 16S rRNA methylase is the gene responsible for resistance. There are two different families of 16S rRNA methylases based upon their target nucleosides: the kanamycin-gentamicin methyltransferase (Kgm) and the kanamycin-apramycin methyltransferase (Kam) (Vojnović et al., 2010). The Kgm confers high-level resistance to both gentamicin and kanamycin as observed in this study.

Since the isolated antibiotic resistant determinant is novel (has not been previously described) and it confers high-level resistance to gentamicin and kanamycin, it is strongly recommended to confirm which ORF is responsible for the resistant phenotype in order to be able to continue characterizing the mechanisms of resistance. The study of novel mechanisms of resistance will help with the future development of new chemotherapeutics to control the effect of antibiotic resistant pathogens in human health.

CHAPTER 4
CONCLUSIONS, RECOMMENDATIONS AND LITERATURE CITED

4.1 CONCLUSIONS

- The described indirect total DNA extraction method using agarose plugs ("worms") and a chemical/enzymatic cells lysis demonstrated to be an excellent alternative to isolate high molecular weight DNA fragments from benthic and ephemeral tropical hypersaline microbial mats.
- There is an apparent effect of the dry and rainy season on the number of clones in metagenomic libraries that could be due to physical-chemical parameters.
- This study has generated the first metagenomic libraries from tropical hypersaline microbial mats in the Antilles and, to the best of our knowledge, the first generated from an ephemeral microbial mat.
- Resistance genes for different antibiotics were found in both benthic and ephemeral metagenomic libraries.
- A kanamycin resistant clone was isolated from all the generated metagenomic libraries and it showed a high-level resistance of a minimum inhibitory concentration > 1024 µg/mL.
- A gentamicin resistant clone was isolated from the dry season benthic metagenomic library and it showed a high-level resistance of a minimum inhibitory concentration >512 µg/mL. The *in silico* analysis of the sequenced insert fragment suggests the presence of a novel enzyme with a chemical modification activity.

4.2 RECOMMENDATIONS

- It is highly recommended to continue optimizing the indirect DNA extraction method presented here to improve the concentration and quality of DNA extracted and to increase the number of clones obtained in large-insert metagenomic libraries generated from the microbial mats under study at rainy season.
- It is recommended to run DNA extractions with BMM and EMM samples using the PEG indirect method described by Bey et al. (2010) and the chemical/enzymatic method described here followed by a pulsed field electrophoresis of the DNA, the verification of its purity index and a 16S rDNA diversity study in order to determine which of the two is the best method for hypersaline microbial mats high molecular weight DNA isolation. Also, a direct method could be performed to make comparisons between direct and indirect methods.
- The metagenomic libraries generated could be subjected to 16S rDNA diversity studies to recognize some of the groups of microorganisms in our collections of genomes.
- It is strongly recommended to sequence the KRC using another transposon alternative.
- It is also advised to continue with the characterization of the GRC; this includes the determination of the exact coding unit for the resistant phenotype followed by the purification of the enzyme (s).

4.3 LITERATURE CITED

Abbassi, M. S., W. Achour and A. B. Hassen. 2009. High-level Gentamicin-Resistant Enterococcus Faecium Strains Isolated from Bone Marrow Transplant Patients: Accumulation of Antibiotic Resistance Genes, Large Plasmids and Clonal Strain Dissemination. *International Journal of Antimicrobiasl*. 29: 658–664.

Alanis, A.J., 2005. Resistance to Antibiotics: Are We in the Post-Antibiotic Era? *Archives of Medical Research*. 36: 697–705.

Allen, H. K., K. A. Cloud-Hansen, J. M. Wolinski, C. Guan, S. Greene, S. Lu, M. Boeyink, N. A. Broderick, K. F. Raffa and J. Handelsman. 2009. Resident Microbiota of the Gypsy Moth Midgut Harbors Antibiotic Resistance Determinants. *DNA and Cell Biology*. 28(3): 109-117.

Amann, R.I., W. Ludwig and K. Shleifer. 1995. Phylogenetic Identification and In Situ Detection of Individual Microbial Cells without Cultivation. *Microbiological Reviews*. 59(1): 143-169.

Andrews, J. M. 2001. Determination of Minimum Inhibitory Concentrations. *Journal of Antimicrobial Chemotherapy*. 48(Suppl. S1): 5-16.

Angly, F., B. Rodríguez-Brito, D. Bangor, P. McNairnie, M. Breitbart, P. Salomon, B. Felts, J. Nulton, J. Mahaffy, and F. Rohwer. 2005. PHACCS, an Online Tool for Estimating the Structure and Diversity of Uncultured Viral Communities Using Metagenomic Information. *BMC Bioinformatics*. 6: 41

Angly, F. E., B. Felts, M. Breitbart, P. Salamon, R. A. Edwards, C. Carlson, A. M. Chan, M. Haynes, S. Kelley, H. Liu, J. M. Mahaffy, J. E. Mueller, N. J. , R. Olson, R. Parsons, S. Rayhawk and C. A. Suttle, F. Rohwer. 2006. The Marine Viromes of Four Oceanic Regions. *Public Library of Science Biology*. 4(11): 2121-2131.

Baskar, S., R. Baskar, L. Mauclaire, and J. A. McKenzie. 2006. Microbially Induced Calcite Precipitation in Culture Experiments: Possible Origin for Stalactites in Sahastradhara Caves, Dehradun, India. *Current Science*. 90(1): 58-64.

Baquero, F., J. L. Martínez and R.l Cantón. 2008. Antibiotics and Antibiotic Resistance in Water Environments. *Current Opinion in Biotechnology*. 19:260–265.

Béjà, O., L. Aravind, E. V. Koonin, M. T. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich and E. F. DeLong. 2000. Bacterial Rhodopsin: Evidence for a New Type of Phototrophy in the Sea. *Science*. 289(5486): 1902–1906.

Beneviste, R. and J. Davies. 1973. Aminoglycoside Antibiotic-Inactivating Enzymes in Actinomycetes Similar to Those Present in Clinical Isolates of Antibiotic-Resistant Bacteria. *Proceedings of the National Academy of Sciences.* 70(8): 2276-2280.

Bey, B.S., E. B. Fichot, G. Dayama, A. W. Decho and R. S. Norman. 2010. Extraction of High Molecular Weight DNA from Microbial Mats. *BioTechniques* 49: 631-640

Bhaya, D., A. R. Grossman, A. Steunou, N. Khuri, F. M. Cohan, N. Hamamura, M. C. Melendrez, M. M. Bateson, D. Ward4 and J. F. Heidelberg. 2007. Population Level Functional Diversity in a Microbial Community Revealed by Comparative Genomic and Metagenomic Analyses. *The ISME Journal*. 1, 703–713.

Bisht, R., A. Katiyar, R. Singh and P. Mittal. 2009. Antibiotic Resistance: A Global Issue of Concern. *Asian Journal of Pharmaceutical and Clinical Research*. 2(2): 34-39.

Brady, S. F. 2007. Construction of Soil Environmental DNA Cosmid Libraries and Screening for Clones that Produce Biologically Active Small Molecules. *Nature Protocols*. 2(5): 1297-1305.

Breitbart, M., B. Felts, S. Kelley, J. M. Mahaffy, J. Nulton, P. Salamon and Forest Rohwer. 2003a. Diversity and population structure of a near-shore marine-sediment viral community. *Proceedings of the Royal Society London B.* 271(1539): 565-574.

Breitbart M., B. Felts, S. Kelley, J. M. Mahaffy, J. Nulton, P. Salamonand F. Rohwer. 2004. Diversity and Population Structure of a Near-Shore Marine-Sediment Viral Community. *Proceedings of the Royal Society London B.* 1-10.

Breitbart, M., I. Hewson, B. Felts, J. M. Mahaffy, J. Nulton, P. Salamon, and F. Rohwer. 2003b. Metagenomic Analyses of an Uncultured Viral Community from Human Feces. *Journal of Bacteriology*. 185 (20): 6220-6223.

Breitbart, M., P. Salamon, B. Andrese, J. M. Mahaffy, A. M. Segall, D. Mead, F. Azam and F. Rohwer. 2002. Genomic Analysis of Uncultured Marine Viral Communities. *Proceedings of the National Academy of Sciences.* 99(22): 14250–14255.

Breitbart, M. and F. Rohwer. 2005. Here a Virus, There a Virus, Everywhere the Same Virus? *TRENDS in Microbiology* 13(6): 278-284.

Brown, M. G. and D. L. Balkwill. 2009. Antibiotic Resistance in Bacteria Isolated from the Deep Terrestrial Subsurface. *Microbial Ecology*. 57:484–493.

Brulc, J. M., D. A. Antonopoulos, M. E. B. Miller, M.K. Wilson, A. C. Yannarell, E. A. Dinsdaled, R. E. Edwards, E. D. Frank, J. B. Emerson, P. Wacklin, P. M. Coutinho, B. Henrissat, K. E. Nelson and B. A. White. 2009. Gene-centric metagenomics of the fiber-adherentbovine rumen microbiome reveals forage specific glycoside hydrolases. *Proceedings of the National Academy of Sciences*. 106(6): 1948-1953.

Bryant, D. A. and N. Frigaard. 2006. Prokaryotic Photosynthesis and Phototrophy Illuminated. *Trends in Microbiology*. 14(11): 488-496.

Bryant, D. A., A. M. García Costas, J. A. Maresca, A. Gomez Maqueo Chew, C. G. Klatt, M. M. Bateson, L. J. Tallon, J. Hostetler, W. C. Nelson, J. F. Heidelberg, and D. M. Ward. 2007. Candidatus *Chloracidobacterium thermophilum*: An Aerobic Phototrophic Acidobacterium. *Science*. 317(5837): 523 – 526.

Cann, A. J., S. E. Frandrich, and S. Heaphy. 2004. Analysis of the Virus Population Present in Equine Faeces Indicates the Presence of Hundreds of Uncharacterized Virus Genomes. *Virus Genes*. 30(2): 151-156.

Cantón, R. 2009. Antibiotic Resistance Genes from the Environment: A Perspective through Newly Identified Antibiotic Resistance Mechanisms in the Clinical Setting. *Clinical Microbiology and Infection*. 15 Suppl 1: 20-25.

Casillas-Martínez, L., L. M. González, Z. Fuentes-Figueroa, C.M. Castro, D. Nieves, C. Hernández, W. Ramírez, R. E. Sytsma, J. Pérez-Jiménez, and P.T. Visscher. 2005. Community Structure, Geochemical Characteristics and Mineralogy of a Hypersaline Microbial Mat, Cabo Rojo, PR. *Geomicrobiology Journal*. 22: 269-288.

Changhui, G., J. Jianhua, B. R. Borlee, L. L. Williamson, B. Shen, K. F. Raffa, and J. Handelsman. 2007. Signal Mimics Derived from a Metagenomic Analysis of the Gypsy Moth Gut Microbiota. *Applied and Environmental Microbiology*. 73(11): 3669-3676.

Clocksin, K. M., D. O. Jung and M. T. Madigan.2007. Cold-active Chemoorganotrophic Bacteria from Permanently Ice-covered Lake Hoare, McMurdo Dry Valleys, Antarctica. *Applied Environmental Microbiology*. 73(9): 3077-3083.

Cundliffe, E. and A. L. Demain. 2010. Avoidance of Suicide in Antibiotic-Producing Microbes. *Journal of Industrial Microbiology and Biotechnology*. 37:643–672.

Cyr, H. and K. E. Morton. 2006. Distribution of Biofilm Exopolymeric Substances in Littoral Sediments of Canadian Shield Lakes: The Effects of Light and Substrate. *Canadian Journal of Fisheries and Aquatic Sciences*. 63: 1763-1776.

D'Costa, V., M., K. M. McGrann, D. W. Hughes, G. D. Wright. 2006. Sampling the Antibiotic Resistome. *Science*. 311(5759):374-377.

Dang, H., J. Ren, L. Song, S. Sun and L. An. 2008. Dominant Chloramphenicol-resistant Bacteria and Resistance Genes in Coastal Marine Waters of Jiaozhou Bay, China. *World Journal of Microbial Biotechnology*. 24:209–217.

Daniel, R. 2005. The Metagenomics of Soil. *Nature*. 3: 470-478.

Dantas, G., M. O. A. Sommer, R. D. Oluwasegun, G. M. Church. 2008. Bacteria Subsisting on Antibiotics. *Science*. 320(5872): 100-103.

Davidson, S. K., S. W. Allen, G. E. Lim, C. M. Anderson, and M. G. Haygood. 2001. Evidence for the Biosynthesis of Bryostatins by the Bacterial Symbiont "*Candidatus* Endobugula sertula" of the Bryozoan *Bugula neritina*. *Applied and Environmental Microbiology*. 67(10): 4531-4537.

Davies J. E. 2006. Aminoglycosides: Ancient and Modern. *The Journal of Antibiotics*. 59(9): 529–532.

Davies, J. and D. Davies. 2010. Origins and Evolution of Antibiotic Resistance. *Microbiology and Molecular Biology Reviews*. 74(3): 417–433.

Dawson, S. C. and L. K. Fritz-Laylin. 2009. Sequencing Free-living Protists: The Case for Metagenomics. *Environmental Microbiology*. 11(7): 1627–1631.

Day, J. M., L. L. Ballard, M. V. Duke, B. E. Scheffler and L. Zsak. 2010. Metagenomic analysis of the turkey gut RNA virus community. *Virology Journal*. 7: 1-8.

Decho, A.W. 2000. Microbial Biofilms in Intertidal Systems: An Overview. *Continental Shelf Research*. 20: 1257-1273.

Demaneche, S., H. Sanguin, J. Pote´, E. Navarro, D. Bernillon, P. Mavingui, W. Wildi, T. M. Vogel and P. Simonet. 2008. Antibiotic-resistant Soil Bacteria in Transgenic Plant Fields. *Proceedings of the National Academy of Sciences* 105(10): 3957–3962.

Des Marais, D. J. 2003. Biogeochemistry of Hypersaline Microbial Mats Illustrates the Dynamics of Modern Microbial Ecosystems and the Early Evolution of the Biosphere. *Biological Bulletin*. 204: 160–167.

Diaz-Torres, M. L., A. Villedieu, N. Hunt, R. McNab, D. A. Spratt, E. Allan, P. Mullany and M. Wilson. 2006. Determining the Antibiotic Resistance Potential of the Indigenous Oral Microbiota of Humans Using a Metagenomic Approach. *Federation of European Microbiological Societies*. 258(2): 257-262.

Doi, Y., J. M. Adams-Haduch and D. L. Paterson. 2008. Genetic Environment of 16S rRNA Methylase Gene rmtD. *Antimicrobial Agents and Chemotherapy*. 52(6): 2270-2272.

Donato, J. J., L. A. Moe, B. J. Converse, K.D. Smart, F. C. Berklein, P. S. McManus and J. Handelsman. 2010. Metagenomic Analysis of Apple Orchard Soil Reveals Antibiotic Resistance Genes Encoding Predicted Bifunctional Proteins. A*pplied and Environmental Microbiology*. 76(13): 4396–4401.

Duan, Y.,1 L. Zhou, D. G. Hall, W. Li, H. Doddapaneni, H. Lin, L. Liu, C. M. Vahling, D. W. Gabriel, K. P. Williams, A. Dickerman, Y. Sun and Tim Gottwald. 2009. Complete Genome Sequence of Citrus Huanglongbing Bacterium, 'Candidatus Liberibacter asiaticus' Obtained Through Metagenomics. *Molecular Plant-Microbe Interactions*. 22(8):1011–1020.

Dunbar, J., S. Takala, S. M. Barns, J. A. Davis and C.R. Kuske. 1999. Levels of BacterialCommunity Diversity in Four Arid Soils Compared by Cultivation and 16S rRNA GeneCloning. *Applied Environmental Microbiology*. 65(4): 1662-1669.
Edwards R. A. and F. Rohwer. 2005. Viral Metagenomics. *Nature Reviews Microbiology*. 3(6): 504-510.

Durso, L. M., G. P. Harhay, J. L. Bono and T. P. L. Smith. 2010. Virulence-Associated and Antibiotic Resistance Genes of Microbial Populations in Cattle Feces Analyzed Using a Metagenomic Approach. *Journal of Microbiological Methods*. 84(2):278-82.

Dyda, F., D. C. Klein and A. B. Hickman. 2000. GCN5-Related N-acetyltransferases: A Structural Overview. *Annual Review of Biophysics and Biomolecular Structure*. 29:81–103.

Elshahed, M. S., F. Z. Najar, M. Aycock, C. Qu, B. A. Roe, and L. R. Krumholz. 2005. . Metagenomic Analysis of the Microbial Community at Zodletone Spring (Oklahoma): Insights into the Genome of a Member of the Novel Candidate Division OD1. *Applied and Environmental Microbiology*. 71(11): 7598-7602.

Fajardo, A., N. Martínez-Martín, M. Mercadillo, J. C. Galán, B. Ghysels, S. Matthijs, P. Cornelis, L. Wiehlmann, B. Tümmler, F. Baquero and J. L.Martínez. 2008. The Neglected Intrinsic Resistome of Bacterial Pathogens. *Public Library of Science ONE*. 3(2):1-6.

Fenchel, T. 1998. Formation of Laminated Cyanobacterial Mats In the Absence of Benthic Fauna. *Aquatic Microbiology*. Ecol. 14: 235-240.

Feng, Y., C.J. Duan, H. Pang, X.C. Mo , C.F. Wu , Y. Yu , Y. L. Hu, J. Wei, J.L. Tang, J.X. Feng. 2007. Cloning and Identification of Novel Cellulase Genes from Uncultured Microorganisms in Rabbit Cecum and Characterization of the Expressed Cellulases. *Applied and Environmental Microbiology*. 75(2):319-28.

Finkbeiner, S. R., A. F. Allred, P. I. Tarr, E. J. Klein, C. D. Kirkwood, and D. Wang. 2008. Metagenomic Analysis of Human Diarrhea: Viral Detection and Discovery. *Public Library of Science Pathogens*. 4(2): 1-9.

Ferrer, M., O. Golyshina, A. Beloqui1 and P. N. Golyshin. 2007. Mining Enzymes from Extreme Environments. *Current Opinion in Microbiology.* 10:207–214.

Fieseler, L., M. Horn, M. Wagner and U. Hentschel. 2004. Discovery of the Novel Candidate Phylum "*Poribacteria*" in Marine Sponges. *Applied and Environmental Microbiology*. 70(6): 3724-3732.

Fisher, J. F. and S. Mobashery. 2010. Enzymology of Bacterial Resistance. *Comprehensive Natural Products II - Chemistry and Biology*. 8: 443-487.

Fujita, Y., S. Takahashi, M. Ueda, A. Tanaka, H. Okada, Y. Morikawa, T. Kawaguchi, M. Arai, H. Fukuda, and A. Kondo. 2002. Direct and Efficient Production of Ethanol from Cellulosic

Material with a Yeast Strain Displaying Cellulolytic Enzymes. *Applied and Environmental Microbiology*. 68(10): 5136-5141.

Gabor, E. M., E. J. de Vries and D. B. Janssen. 2004. Construction, Characterization, and Use of Small-Insert Gene Banks of DNA Isolated from Soil and Enrichment Cultures for the Recovery of Novel Amidases. *Environmental Microbiology*. 6(9):948-58.

Gabor , E. M., E. J. de Vries, D.B. Janssen. 2003. Eficient Recovery of Environmental DNA for Expression Cloning by Indirect Extraction Methods. *FEMS Microbiology Ecology*. 44: 153-163.

García-Pichel, F., M. Mechling and R. W. Catsenholz. 1994. Diel Migrations of Microorganisms within a Benthic, Hypersaline Mat Community. *Applied Environmental Microbiology*. 60 (5): 1500-1511.

Gaze, W., C. O. Neill, E. Wellington, and P. Hawkey. 2008. Antibiotic Resistance in the Environment, with Particular Reference to MRSA. *Advances in Applied Microbiology*.63: 249-280.

Gillespie, D. E., S. F. Brady, A. D. Bettermann, N. P. Cianciotto, M. R. Liles, M. R. Rondon, J. Clardy, R. M. Goodman and J. Handelsman. 2002. Isolation of Antibiotics Turbomycin A and B from a Metagenomic Library of Soil Microbial DNA. *Applied and Environmental Microbiology*. 68(9): 4301–4306.

Giovannoni, S. J., T. B. Britschgi, C. L. Moyer and K. G. Field. 1990. Genetic Diversity in Sargasso Sea Bacterioplankton. *Nature*. 345:60-63.

Giovannoni, S.J., E.F. DeLong, G.J. Olsen and N.R. Pace. 1988. Phylogenetic Group-Specific Oligodeoxynucleotide Probes for the Identification of Single Microbial Cells. *Journal of Bacteriology*. 170(2): 720-726.

Gloux, K., M. Leclerc, H. Iliozer, R. L'Haridon, C. Manichanh, G. Corthier, R. Nalin, H. M. Blottie`re, and J. Dore´. 2007. Development of High-Throughput Phenotyping of Metagenomic Clones from the Human Gut Microbiome for Modulation of Eukaryotic Cell Growth. *Applied Environmental Microbiology*. 73(11): 3732-3737.

Gordon, D.A. and S.J. Giovannoni. 1996. Detection of Stratified Microbial Populations Related to *Chlorobium* and *Fibrobacter* Species in the Atlantic and Pacific Oceans. *Applied Environmental Microbiology*. 62(4): 1171-1177.

Greninger, A. L., E. C. Chen, T. Sittler, A. Scheinerman, N. Roubinian G. Yu, E. Kim, D. R. Pillai, C. Guyard, T. Mazzulli, P. Isa, C. F. Arias, John Hackett Jr.11, Gerald Schochetman, Steve Miller, Patrick Tang, Charles Y. Chiu. 2010. A Metagenomic Analysis of Pandemic Influenza A (2009 H1N1) Infection in Patients from North America. *PLoS ONE* . 5(10): 1-16.

Grozdanov, L. and U. Hentschel. 2007. An Environmental Genomics Perspective on the Diversity and Function of Marine Sponge-Associated Microbiota. *Current Opinion in Microbiology.* 10:215–220.

Guerrero, M. C. and R. de Wit. 1992. Microbial Mats In the Inland Saline Lakes of Spain. *Limnetica*. 8: 197-204.

Gupta R., Q.K. Beg and P. Lorenz. 2002. Bacterial Alkaline Proteases: Molecular Approaches and Industrial Applications. *Appliedl Microbiology Biotechnology*. 59:15–32.

Hallam, S. J., K. T. Konstantinidis, N. Putnam, C. Schleper, Y. Watanabe, J. Sugahara, C. Preston, J. de la Torre, P. M. Richardson and E. F. DeLong. 2006. Genomic Analysis of the Uncultivated Marine Crenarchaeote *Cenarchaeum symbiosum*. *Proceedings of the National Academy of Sciences.* 103(48): 18296-18301.

Handelsman J. 2004. Metagenomics: Application of Genomics to Uncultured Microorganisms. *Microbiology and Molecular Biology Reviews*. 68(4): 669–685.

Handelsman, J. and K. Smalla. 2003. Conversations with the Silent Majority. *Current Opinion in Microbiology*. 6: 271-273.

Handelsman, J., M. Liles, D. Mann, C. Riesenfeld and R. M. Goodman. 2002. Cloning the Metagenome: Culture-independent Access to the Diversity and Functions of the Uncultivated Microbial World. *Methods in Microbiology*. 33:241-255.

Handelsman, J., M.R. Rondon, S. F. Brady, J. Clardy andR. M. Goodman. 1998. Molecular Biological Access to the Chemistry of Unknown Soil Microbes: A New Frontier for Natural Products. *Chemistry and Biology*. 5(10): R245-R249.

Harvey, R., J. Funk, T. E. Wittum and A. E. Hoet. 2009. A Metagenomic Approach for Determining Prevalence of Tetracycline Resistance Genes in the Fecal Flora of Conventionally Raised Feedlot Steers and Feedlot Steers Raised without Antimicrobials. *American Jurnal of Veterinary research.* 70(2):198-202.

Healy, F. G., R. M. Ray, H. C. Aldrich, A. C. Wilkie, L. O. Ingram, K. T. Shanguman. 1995. Direct Isolation of Functional Genes Encoding Cellulases from the Microbial Consortia in a Thermophilic, Anaerobic Digester Maintained on Lignocellulose. *Applied Microbiology Biotechnology*. 43(4):667-674.

Henne, A., R. A. Schmitz, M. Bomeke, G. Gottschalk, and R. Daniel. 2000. Screening of Environmental DNA Libraries for the Presence of Genes Conferring Lipolytic Activity on Escherichia coli. *Applied Environmental Microbiology*. 66(7): 3113-3116.

Heuer, H., E. Krogerrecklenfort, E.M.H. Wellington, S. Egan, J.D. van Elsas, L. van Overbeek, J.M. Collard, G. Guillaume, A.D. Karagouni. 2002. Gentamicin Resistance Genes in Environmental Bacteria: Prevalence and Transfer. *FEMS Microbiology Ecology* 42: 289-302.

Hildebrand, M., L. E. Waggoner, H. Liu, S. Sudek, S. Allen, C. Anderson, D. H. Sherman and M. Haygood1. 2004. *bryA*: An Unusual Modular Polyketide Synthase Gene from the Uncultivated Bacterial Symbiont of the Marine Bryozoan *Bugula neritina. Chemistry & Biology*. 11: 1543–1552.

Hiraishi, A., T. Umezawa, H. Yamamoto, K. Kato and Y. Maki. 1999. Changes in Quinone Profiles of Hot Spring Microbial Mats with a Thermal Gradient. *Applied Environmental. Microbiology*. 65(1): 198-205.

Hirst, C. N., H. Cyr, I. A. Jordan. 2003. Distribution of Exopolymeric Substances in the Littoral Sediments of an Oligotrophic Lake. *Microbial Ecology*. 46: 22–32.

Hoff, M. 2006. Metagenomics Off ers a Big-Picture View of the Diversity and Distribution of Marine Viruses. *Public Library of Science Biology*. 4(11): 1885-1886.

Hochmuth, T. and J. Piel. 2009. Polyketide Synthases of Bacterial Symbionts in Sponges – Evolution-based Applications in Natural Products Research. *Phytochemistry*. 70 (2009) 1841–1849.

Hochmuth, T., H. Niederkrüger, C. Gernert, A. Siegl, S. Taudien, M. Platzer, P. Crews, U. Hentschel, and J. Piel. 2010. Linking Chemical and Microbial Diversity in Marine Sponges: Possible Role for Poribacteria as Producers of Methyl-Branched Fatty Acids. *ChemBioChem.* 11: 2572-2578.

Hongoh Y.2010. Diversity and Genomes of Microbial Symbionts in the Termite Gut. *Bioscience Biotechnoogy and Biochemistry*. 74(6): 1145-1151.

Hughes D. B., H. Felbeck and J.L. Stein. 1997. A Histidine Protein Kinase Homolog from the Endosymbiont of the Hydrothermal Vent Tubeworm *Riftia pachyptila. Applied and Environmental Microbiology.* 63(9): 3494–3498.

Isenbarger, T.A., M. Finney, C. Ríos-Velázquez, J. Handelsman and G. Ruvkun. 2008. Miniprimer PCR, a New Lens for Viewing the Microbial World. *Applied and Environmental Microbiology*. 74(3): 840–849.

Jannasch, H.W. and G.E. Jones. 1959. Bacterial Populations in Sea Water as Determined by Different Methods of Enumeration. *Limonology and Oceanography*. 4(2) 128-139.

Kapoor, A., J. Victoria, P. Simmondsc, E. Slikas, T. Chieochansinc, A. Naeemd, S. Shaukatd, S. Sharifd, M. Masroor Alamd, M. Angezd, C. Wange, R. W. Shafere, S. Zaidid, and E. Delwarta.2008. A Highly Prevalent and Genetically Diversified Picornaviridae Genus in South Asian Children. *Proceeding of the National Academy of Science*. 105(51): 20482–20487.

Kazimierczak, K. A., K. P. Scott, D. Kelly, and R.I. Aminov. 2009. Tetracycline Resistome of the Organic Pig Gut. *Applied and Environmental Microbiology*. 75(6): 1717–1722.

Kenneddy, J., J. R. Marchesi, and A. D. Dobson. 2007. Metagenomic Approaches to Exploit the Biotechnological Potential of the Microbial Consortia of Marine Sponges. *Applied Microbiology Biotechnology*. 75(1): 11-20.

Kennedy J., J. R. Marchesi and A. DW. Dobson. 2008. Marine Metagenomics: Strategies for the Discovery of Novel Enzymes with Biotechnological Applications from Marine Environments. *Microbial Cell Factories.* 7(27):1-8.

Kharel M. K., B. Subba, D. B. Basnet, J. S. Woo, H. C. Lee, K. Liou,a and J. K. Sohng. 2004. A Gene Cluster for Biosynthesis of Kanamycin from Streptomyces Kanamyceticus: Comparison with Gentamicin Biosynthetic Gene Cluster. *Archives of Biochemistry and Biophysics*. 429: 204–214.

Kim, T. K. and J. A. Fuerst. 2006. Diversity of Polyketide Synthase Genes from Bacteria Associated with the Marine Sponge *Pseudoceratina clavata*: Culture-Dependent and Culture-Independent Approaches. *Environmental Microbiology*. 8(8): 1460-70.

Kim, U., B. W. Birren, T. Slepak, V. Mancino, C. Boysen, H. Kang, M. I. Simon, and H. Shizuya. 1996. Construction and Characterization of a Human Bacterial Artificial Chromosome Library. *Genomics*. 34: 213–218.

Kimura, N. 2006. Metagenomic: Access to Unculturable Microbes in the Environment. *Microbes and Environments*. 21(4): 201-215.

Klatt, C.G., D. A. Bryant and D. M. Ward. 2007. Comparative Genomics Provides Evidence for the 3-hydroxypropionate Autotrophic Pathway in Filamentous Anoxygenic Phototrophic Bacteria and in Hot Spring Microbial Mats. *Environmental Microbiology*. 9(8): 2067-2078.

Klatt, C. G., J. M. Wood, D. B. Rusch, M. M. Bateson, N. Hamamura, J. F. Heidelberg, A. R. Grossman, D. Bhaya, F. M. Cohan, M. Ku¨, D. A. Bryant and D. M. Ward. 2011. Community Ecology of Hot Spring Cyanobacterial Mats: Predominant Populations and Their Functional Potential. *The ISME Journal*. 5: 1262–1278.

Knietsch A., T. Waschkowitz, S. Bowien, A. Henne and R. Daniel. 2003. Metagenomes of Complex Microbial Consortia Derived from Different Soils as Sources for Novel Genes Conferring Formation of Carbonyls from Short-Chain Polyols on Escherichia coli. *Journal of Molecular Microbiology and Biotechnology*. 5:46–56.

Kojic, M., S. Vojnovic, N. Vukov and B. Vasiljevic.2007. Overexpression of SGM 5' UTR mRNA Reduces Gentamicin Resistance in Both *Escherichia coli* and *Micromonospora melanosporea* cells. *Archives of Biological Sciences*. 59 (4): 273-280.

Kottman, R. 2005. Data Acquisition, Design and Implementation of an Integrative Database for Metagenomics. Max Plank Society. 1-47.

Kristensen D. M.. A. R. Mushegian, V. V. Dolja and E. V. Koonin. 2010. New Dimensions of the Virus World Discovered through Metagenomics. *Trends in Microbiology*. 18(1):11-19.

Kube, M., A. Beck, A. Meyerdierks, R. Amann, R. Reinhardt, and R. Rabus. 2005. A Catabolic Gene Cluster for Anaerobic Benzoate Degradation in Methanotrophic Microbial Black Sea mats. *Systematic and Applied Microbiology*. 28 (4):287-94.

Lang K. S., J. M. Anderson, S. Schwarz, L. Williamson, J. Handelsman and R. S. Singer. 2010. Novel Florfenicol and Chloramphenicol Resistance Gene Discovered in Alaskan Soil by Using Functional Metagenomics. *Applied and Environmental Microbiology*. 76(15): 5321–5326.

Ley, R. E., J. K. Harris, J. Wilcox, J. R. Spear, S. R. Miller, B. M. Bebout, J. A. Maresca, D. A. Bryant, M. L. Sogin and N. R. Pace. 2006. Unexpected Diversity and Complexity of the Guerrero Negro Hypersaline Microbial Mat. *Environmental Microbiology*. 72(5):3685-3695.

Li L., J. G. Victoria, C. Wang, M. Jones, G. M. Fellers, T. H. Kunz and E. Delwar. 2010. Bat Guano Virome: Predominance of Dietary Viruses from Insects and Plants Plus Novel Mammalian Viruses. *Journal of Virology*. 84(14):6955-6965.

Liesack W. and E. Stackebrandt. 1992. Occurrence of Novel Groups of the Domain Bacteria as Revealedby Analysis of Genetic Material Isolated from anAustralian Terrestrial Environment. *Journal of Bacteriology*. 174(15): 5072-5078.

Liles M. R., L. L. Williamson, J. Rodbumrer, V. Torsvik, R. M. Goodman, and J. Handelsman. 2008. Recovery, Purification, and Cloning of High-Molecular-Weight DNA from Soil Microorganisms. *Applied and Environmental Microbiology*. 74(10): 3302-3305.

Liles, M. R., L. L. Williamson, J. Handelsman, and R. M. Goodman.2004. Isolation of High-Molecular-weight Genomic DNA from soil Bacteria for Genomic Library Construction. *Molecular Microbial Ecology Manual*. 2: 839-850.

Liles M.R., B. F. Manske, S. B. Bintrim, J. Handelsman and R. M. Goodman. 2003. A Census of rRNA Genes and Linked Genomic Sequenceswithin a Soil Metagenomic Library. *Applied and Environmental Microbiology*. 69(5): 2684–2691.

Liu K., J. Wang, D. Bu, S. Zhao, C. McSweeney, P. Yu and D. Li. 2009a. Isolation and Biochemical Characterization of Two Lipases from a Metagenomic Library of China Holstein Cow Rumen. *Biochemical and Biophysical Research Communications*. 385: 605–611.

Lu J. and J. S. Domingo. 2008. Turkey fecal microbial community structure and functional gene diversity revealed by 16S rRNA gene and metagenomic sequences. *Journal of Microbiology*. 46(5):469-77.

Lu J., J. S. Domingo and O. C. Shanks. 2007. Identification of Chicken-Specific Fecal Microbial Sequences Using a Metagenomic Approach. *Water Research.* 41: 3561 – 3574.

Lynch S. R., R. L. Gonzalez, Jr. and J. D. Puglisi. 2003. Comparison of X-Ray Crystal Structure of the 30S Subunit-Antibiotic Complex with NMR Structure of Decoding Site Oligonucleotide-Paromomycin Complex. *Structure*. 11: 43–53.

Majernik, A., G. Gottschalk and R. Daniel. 2001. Screening of Environmental DNA Libraries for the Presence of Genes Conferring Na_(Li_)/H_ Antiporter Activity on Escherichia coli: Characterization of the Recovered Genes and the Corresponding Gene Products. *Journal of Bacteriology*. 183(22): 6645-6653.

Massana R, B. Karniol, T .Pommier, I. Bodaker, O. Béjà. 2008. Metagenomic Retrieval of a Ribosomal DNA Repeat Array from an Uncultured Marine Alveolate. *Environmental Microbiology.* 10(5):1335-43.

Massana R., E. F. DeLong and C. Pedrós-Alió. 2000. A Few Cosmopolitan Phylotypes Dominate Planktonic ArchaealAssemblages in Widely Different Oceanic Provinces. *Applied and Environmental Microbiology*. 66(5): 1777-1787.

Matsui T., G. Tokuda and N. Shinzato. 2009. Termites as Functional Gene Resources. *Recent Patents on Biotechnology.* 3: 10-18

Meisinger, D. B., J. Zimmermann, W. Ludwig, K. H. Schleifer, G. Wanner , M. Schmid, P. C. Bennett, A. S. Engel, N. M. Lee,. 2007. In Ssitu Detection of Novel Acidobacteria in Microbial Mats from a Chemolithoautotrophically Based Cave Ecosystem (Lower Kane Cave, WY, USA). *Environmental Microbiology*. 9(6): 1523-34.

Messi, P., E. Guerrieri, M. Bondi. 2005. Antibiotic Resistance and Antibacterial Activity in Heterotrophic Bacteria of Mineral Water Origin. *Science of the Total Environment*. 346: 213–219.

Mocali S. and A. Benedetti. 2010. Exploring Research Frontiers in Microbiology: The Challenge of Metagenomics in Soil Microbiology. *Research in Microbiology*. 161:497-505.

Monier, A., J. M. Claverie and H. Ogata. 2008. Taxonomic Distribution of Large DNA Viruses in the Sea. *Genome Biology*. 9(7): R1061-10615.

Monika Bright and Silvia Bulgheresi. 2010. A Complex Journey: Transmission of Microbial Symbionts. *Nature Reviews*. 8:218-230.

Monteiro A.C.M., R.C.C Ferreira, G. Padilla, L.C.S. Ferreira and S.O.P. Costa. Environmental and Genetic Factors Affecting Mutability to Aminoglycoside Antibiotics Among Escherichia coli K12 Strains. *Genetics and Molecular Biology*. 26(2): 221-227.

Morales-Jiménez, J., G. Zúñiga, L. Villa-Tanaca and C. Hernández-Rodríguez.2009. Bacterial Community and Nitrogen Fixation in the Red Turpentine Beetle, Dendroctonus valens LeConte (Coleoptera: Curculionidae: Scolytinae). *Microbial Ecology*. 58(4):879-91.

Moran N.A., P.H. Degnan . 2006. Functional Genomics of Buchnera and the Ecology of Aphid Hosts. *Molecular Ecology*. 15(5):1251-1261.

Mori, T., S. Mizuta, H. Suenaga and K. Miyazaki1. 2008. Metagenomic Screening for Bleomycin Resistance Genes. *Applied and Environmental Microbiology*. 74(21): 6803–6805.

Moyer, C. L., F. D. Dobbs, and D. M. Karl. 1994. Estimation of Diversity and Community Structure through Restriction Fragment Length Polymorphism Distribution Analysis of Bacterial 16S rRNA Genes from a Microbial Mat at an Active, Hydrothermal Vent System, Loihi Seamount, Hawaiit. *Applied and Environmental Microbiology*. 871-879.

Murphy E. F., P. D. Cotter, S. Healy, T. M. Marques, O. O'Sullivan, F. Fouhy, S. F. Clarke, P. W. O'Toole, E. M. Quigley, C. Stanton, P. R. Ross, R. M. O'Doherty and F. Shanahan. 2010. Composition and Energy Harvesting Capacity of the Gut Microbiota: Relationship to Diet, Obesity and Time in Mouse Models. *Gut.* 59(12):1635-42.

Nanba K., G. M. King, and K. Dunfield. 2004. Analysis of Facultative Lithotroph Distribution and Diversity on Volcanic Deposits by Use of the Large Subunit of Ribulose. *Applied and Environmental Microbiology.* 70(4): 2245-2253.

Nelson, W. C., L. Wollerman, D. Bhaya and J. F. Heidelberg. 2011. Analysis of Insertion Sequences in Thermophilic Cyanobacteria: Exploring the Mechanisms of Establishing, Maintaining, and Withstanding High Insertion Sequence Abundance. *Applied Environmental Microbiology*. 77(15): 5458-5466.

Ng, T. F. F., C. Manire, K. Borrowman, T. Langer, L. Ehrhart and M. Breitbart. 2009. Discovery of a Novel Single-Stranded DNA Virus from a Sea Turtle Fibropapilloma by Using Viral Metagenomics. *Journal of Virology*. 83(6): 2500–2509.

Noonan, J. P., M. Hofreiter, D. Smith, J. R. Priest, N. Rohland, G. Rabeder, J. Krause, J. C. Detter, S. Pa¨a¨bo, and E. M. Rubin. 2005. Genomic Sequencing of Pleistocene Cave Bears. *Science*. 309: 597-600.

Noonan, J.P., G. Coop, S. Kudaravalli, D. Smith, J. Krause, J. Alessi, F. Chen, D. Platt, S. Pääbo, J. K. Pritchard, and E. M. Rubin. 2006. Sequencing and Analysis of Neanderthal Genomic DNA. *Science*. 314: 1113-1118.

Nunoura, T., H. Hirayama, H. Takami, H. Oida, S. Nishi, S. Shimamura, Y. Suzuki, F. Inagaki, K. Takai, K. H. Nealson and K. Horikoshi. 2005. Genetic and Functional Properties of

Uncultivated Thermophilic Crenarchaeotes from a Subsurface Gold Mine as Revealed by Analysis of Genome Fragments. *Environmental Microbiology*. 7(12): 1967-1984.

Nunoura, T., Y. Takaki, J. Kakuta, S. Nishi, J. Sugahara, H. Kazama, G. Chee, M. Hattori, A. Kanai, H. Atomi, K. Takai and H. Takami. 2011. Insights into the Evolution of Archaea and Eukaryotic Protein Modifier Systems Revealed by the Genome of a Novel Archaeal Group. *Nucleic Acids Research*. 39(8): 3204–3223.

Ogle J. M. and V. Ramakrishnan. 2005. Structural Insights Into Translational Fidelity. *Annual Review of Biochemistry*. 74:129–77.

Ohmiya, K., K. Nagashima, T. Kajino, E. Goto, A. Tsukada, and S. Shimizu. 1988. Cloning of the Cellulase Gene from Ruminococcusalbus and its Expression in Escherichia coli. *Applied and Environmental Microbiology*. 54(6): 1511-1515.

Pace, NR, DA Stahl, DJ Lane and GJ Olsen. 1985. Analyzing Natural Microbial Populations by rRNA Sequences. *American Society of Microbiology News*. 51(1): 4-12

Park, J. W., J. S. Joong Hong, N. Parajuli, W. S. Jung, S. R. Park, S. Lim, J. K. Sohng and Y. J. Yoon. 2008. Genetic Dissection of the Biosynthetic Route to Gentamicin A2 by Heterologous Expression of its Minimal Gene Set. *Proceedings of the National Academy of Sciences*. 105(24): 8399–8404.

Piel J., D. Butzke, N. Fusetani, D. , M. Platzer , G. Wen, S. Matsunaga . 2005. Exploring the Chemistry of Uncultivated Bacterial Symbionts: Antitumor Polyketides of the Pederin Family. *Journal of Natural Products.* 68(3):472-47

Piel J., D. Hui, N. Fusetani, S. Matsunaga. 2003. Targeting Modular Polyketide Synthases with Iteratively Acting Acyltransferases from Metagenomes of Uncultured Bacterial Consortia. *Environmental Microbiology.* 6**:** 921-927.

Piel J., D. Hui, G. Wen, D.l Butzke, M. Platzer, N. Fusetani and Shigeki Matsunaga. 2004. Antitumor Polyketide Biosynthesis by an Uncultivated Bacterial Symbiont of the Marine Sponge *Theonella swinhoei*. *Proceedings of the National Academy od Science.* 101(46): 16222–16227.

Piel, J., I. Höfer and D. Hui. 2004. Evidence for a Symbiosis Island Involved in Horizontal Acquisition of Pederin Biosynthetic Capabilities by the Bacterial Symbiont of *Paederus fuscipes* Beetles. *Journal of Bacteriology*. 186(5): 1280-1286.

Poinar, H. N., C. Schwarz, J. Qi, B. Shapiro, R. D. E. MacPhee, B. Buigues, A. Tikhonov, D. H. Huson, L. P. Tomsho, A. Auch, M. Rampp,W. Miller, S.C . Schuster. 2006. Metagenomics to Paleogenomics: Large-Scale Sequencing of Mammoth DNA. *Science*. 311(5759): 392-394.

Raoult D. 2008. Obesity Pandemics and the Modification of Digestive Bacterial Flora. *Eur J Clin Microbiol Infect Dis.* 27:631–634.

Raper, K. B., D. F. Alexander, R. D. Coghill. 1944. Penicillin: II. Natural Variation and Penicillin Production in Penicillium notatum and Allied Species. *Journal of Bacteriology*. 48(6):639-59.

Riesenfeld, C.S., P.D. Schloss and J. Handelsman. 2004. Metagenomics: Genomic Analysis of Microbial Communities. *Annual Review of Genetics*. 38:525–552.

Risatti, J. B., W. C. Capman and D. A. Stahl. 1994. Community Structure of a Microbial Mat: The Phylogenetic Dimension. *Proceedings of the National Academy of Science*. 91: 10173-10177.

Robidart J.C., S.R. Bench , R.A. Feldman , A. Novoradovsky , S.B. Podell , T. Gaasterland , E.E .Allen , H. Felbeck . 2008. Metabolic Versatility of the Riftia pachyptila Endosymbiont Revealed through Metagenomics. *Environmental Microbiology*. 2008. 10(3):727-37.

Rondon M.R., P.R. August, A.D. Bettermann, S.F. Brady, T.H. Grossman, M.R. Liles, K.A. Loiacono, B.A Lynch, I.A. MacNeil, C. Minor, C. L. Tiong, M. Gilman, M. S. Osburne, J. Clardy, J. Handelsman and R. M. Goodman. 2000. Cloning the Soil Metagenome: a Strategy for Accessing the Genetic and Functional Diversity of Uncultured Microorganisms. *Applied and Environmental Microbiology*. 66(6): 2541–2547.

Rondon, M. R., S. J. Raffel, R. M. Goodman, and J. Handelsman. 1999. Toward Functional Genomics in Bacteria: Analysis of Gene Expression in Escherichia coli from a Bacterial Artificial Chromosome Library of Bacillus cereus. *Proceedings of the National Academy of Science*. 96(11): 6451-6455.

Roose-Amsaleg, C.L., E. Garnier-Sillam and M. Harry. 2001. Extraction and Purification of Microbial DNA from Soil and Sediment Samples. *Applied Soil Ecology*. 18: 47–60.

Roossinck M.J., P. Saha, G.B. Wiley, J. Quan, J. D. White, H. Lai, F. Chavarría, G. Chen and B. A. Roe. 2010. Ecogenomics: Using Massively Parallel Pyrosequencing to Understand Virus Ecology. *Molecular Ecology*. 19 (1): 81–88.

Rosario K., C. Nilsson, Y. Wei Lim, Y. Ruan and M. Breitbart. 2009. Metagenomic Analysis of Viruses in Reclaimed Water. *Environmental Microbiology.* 11(11), 2806–2820.

Roszak, D.B. and R.R. Colwell. 1987. Survival Strategies of Bacteria in the Natural Environment. *Microbiological Reviews*. 51(3):365-379.

Saxena, R. K., P. K. Ghosh, R. Gupta, W. Sheba Davidson, S. Bradoo and R. Gulati. 1999. Microbial lipases: Potential Biocatalysts for the Future Industry. *Current Science*. 77(1): 101-115.

Schirmer, A., R. Gadkari, C. D. Reeves, F. Ibrahim, E. F. DeLong and C. R. Hutchinson. 2005. Metagenomic Analysis Reveals Diverse Polyketide Synthase Gene Clusters in Microorganisms

Associated with the Marine Sponge *Discodermia dissolute*. *Applied and Environmental Microbiology*. 71(8): 4840-4849.

Schleper, C., E. F. DeLong, C. M. Preston, R. A. Feldman, K. Wu and Ronald V. Swanson. 1998. Genomic Analysis Reveals Chromosomal Variation in Natural Populations of the Uncultured Psychrophilic Archaeon *Cenarchaeum symbiosum. Journal of Bacteriology*. *180(19): 5003-5009.*

Schloss, P. D. and J. Handelsman. 2003.Biotechnological Prospects from Metagenomics. *Current Opinion in Biotechnology*. 14:303–310.

Schmidt, T. M., E. F. DeLong, and N. R. Pace. 1991. Analysis of a Marine Picoplankton Community by 16S rRNA Gene Cloning and Sequencing. *Journal of Bacteriology*. 173(14): 4371-4378.

Schoenfeld T., M. Patterson, P. M. Richardson, K. E. Wommack, M. Young and D. Mead.2008. Assembly of Viral Metagenomes from Yellowstone Hot Springs. *Applied and Environmental Microbiology*. 74(13): 4164–4174.

Sebat, J. L., F. S. Colwell, and R. F. Crawford. 2003. Metagenomic Profiling: Microarray Analysis of an Environmental Genomic Library. *Applied and Environmental Microbiology*. 69(8):4927–4934

Sharma, A., V. Shukla, R. Verma and J. Shrivastava. 2009. Molecular Characterization and Antibiogram Analysis of Bacterial Community from River Narmada, India. *Journal of Applied Biosciences*. 17: 930 – 940.

Shaw, K. J., P. N. Rather, T. R. S. Hare and G. H. Miller. 1993. Molecular Genetics of Aminoglycoside Resistance Genes and Familial Relationships of the Aminoglycoside-Modifying Enzymes. *Microbiological Reviews*. 57(1): 138-163.

Shigenobu S., H. Watanabe, M. Hattori,Y. Sakaki and H. Ishikawa. 2000. Genome Sequence of the Endocellular Bacterial Symbiont of Aphids Buchnera sp. APS. *Nature*: 407: 81-86.

Shizuya, H., B. Birren, U. Kim, V. Mancino, T. Slepak, Y Tashiiri, and M. Simmon. 1992. Cloning and Stable Maintenance of 300-kilobase-pair Fragments of Human DNA in Escherichia coli Using an F-factor-based Vector. *Proceedings of the National Academy of Science*. USA. 89: 8794-8797.

Singh B., S. K. Gautam, V. Verma, M. Kumar and B. Singh. 2008. Metagenomics in Animal Gastrointestinal Ecosystem: Potential Biotechnological Prospects. *Anaerobe*. 14: 138– 144.

Sommer M. O. A., G. Dantas and G. M. Church. 2009. Functional Characterization of the Antibiotic Resistance Reservoir in the Human Microflora. *Science.* 325: 1128-1131.

Sørensen, K. B., D. E. Canfield, A. P. Teske, and A. Oren. Community Composition of a Hypersaline Endoevaporitic Microbial Mat. *Applied and Environmental Microbiology*. 71(11): 7352–7365.

Spear, J. R., R. E. Ley, A. B. Berger and N. R. Pace. 2003. Complexity in Natural Microbial Ecosystems: The Guerrero Negro Experience. *The Biological Bulletin*. 204: 168–173.

Stackebrandt E., W. Liesack and B. M Goebel. 1993. Bacterial Diversity in a Soil Sample from a SubtropicalAustralian Environment as Determined by 16S rDNAAnalysis. *The FASEB Journal.* 7(1): 232-236.

Stein, J.L., T. L. Marsh, K. Y. Wu, H. Shizuya, and E. F. DeLong. 1996. Characterization of Uncultivated Prokaryotes: Isolation and Analysis of a 40-Kilobase-Pair Genome Fragment from a Planktonic Marine Archaeon. *Journal of Bacteriology*. 178(3): 591-599.

Streit, W. R. and R. A. Schmitz. 2004. Metagenomics – The Key to the Uncultured Microbes. *Current Opinion in Microbiology*. 7: 492–498.

Summers Engel, A., M. L. Porter, L. A. Stern, S. Quinlan, and P. C. Bennett. 2004. Bacterial Diversity and Ecosystem Function of Filamentous Microbial Mats from Aphotic (Cave) Sulfidic Springs Dominated by Chemolithoautotrophic ‘‘Epsilonproteobacteria’’. *FEMS Microbiology Ecology*. (Article in Press).

Svraka S., K. Rosario, E. Duizer, H. van der Avoort, M. Breitbart and M. Koopmans. 2010. Metagenomic Sequencing for Virus Identification in a Public-Health Setting. *Journal of General Virology*.91: 2846–2856.

Taton, A., S. Grubisic, E. Brambilla, R. De Wit, and A. Wilmotte. 2003. Cyanobacterial Diversity in Natural and Artificial Microbial Mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a Morphological and Molecular Approach. *Applied and Environmental Microbiology*. 69(9): 5157-5169.

Taylor, M. W., R. Radax, D. Steger and M. Wagner. 2007. Sponge-Associated Microorganisms: Evolution, Ecology, and Biotechnological Potential. *Microbiology and Molecular Biology Reviews*. 71(2): 295-347.

Thomas ,V. and G. Greub. 2010. Amoeba/Amoebal Symbiont Genetic Transfers: Lessons from Giant Virus Neighbours. *Intervirology*. 53:254–267.

Treude, T., V. Orphan, K. Knittel, A. Gieseke, C. H. House and A. Boetius. 2005. Consumption of Methane and CO2 by Methanotrophic Microbial Mats from Gas Seeps of the Anoxic Black Sea. *Applied and Environmental Microbiology*. 73(7): 2271-2283.

Turnbaugh P. J., R. E. Ley, M. Hamady, C. M. Fraser-Liggett, R. Knight, and J. I. Gordon. 2007. The Human Microbiome Project. *Nature*. 449: 804-810.

Turnbaugh P. J., R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis and J. I. Gordon. 2006. An Obesity-Associated Gut Microbiome with Increased Capacity for Energy Harvest. *Nature*. 444(7122):1027-31.

Turnbaugh P. J., V. K. Ridaura, J. J. Faith, F. E. Rey, R. Knight, and J. I. Gordon. 2009. The Effect of Diet on the Human Gut Microbiome: A Metagenomic Analysis in Humanized Gnotobiotic Mice. *Science Translational Medicine.* 1(6):1-19.

Turque, A. S., D. Batista, C. B. Silveira, A. M. Cardoso, R. P. Vieira, F. C. Moraes, M. M. Clementino, R. M. Albano, R. Paranhos, O. B. Martins and G.Muricy. 2010. Environmental Shaping of Sponge Associated Archaeal Communities. *Public Library of Science ONE*. 5(12): 1-10.

Varin, T., C. Lovejoy, A. D. Jungblut, W. F. Vincent, J. Corbeil. 2011. Metagenomic Analysis of Stress Genes in Microbial Mat Communities from Antarctica and the HighArctic. *Applied Environmental Microbiology*. 78(2): 549-559.

Vega R. L., K. L. Barotta, D. Halla, H. Liua, B. Rodriguez-Muellera, C. Desnuesa, R.A. Edwards, M. Haynesa, F. E. Anglya, L. Wegleya, and F. L. Rohwera,e. 2008. Metagenomic Analysis Indicates that Stressors Induce Production of Herpes-Like Viruses in the Coral Porites Compressa. *Proceedings of the National Academy of Science*. 105(47): 18413–18418.

Victoria J. G., C. Wang, M. S. Jones, C. Jaing, K. McLoughlin, S. Gardner and Eric L. Delwart. *2010.* Viral Nucleic Acids in Live-Attenuated Vaccines: Detection of Minority Variants and an Adventitious Virus. *Journal of Virology.* 84(12): 6033-6040.

Villanueva L., "Ecophysiological and Molecular Characterization of Estuarine Microbial Mats" (Ph.D. dissertation, University of Barcelona, 2005), 357.

Vishnivetskaya T. A. and S. Kathariou. 2005. Putative Transposases Conserved in Exiguobacterium Isolates from Ancient Siberian Permafrost and from Contemporary Surface Habitats. *Environmental Microbiology.* 71(11): 6954–6962.

Vojnović S., T. Ilić-Tomić, M. Savić, S. Bajkić, I. Morić and B. Vasiljević. 2010. Comparative Analysis of 16S rRNA Methyltransferases Conferring Resistance to Aminoglycoside Antibiotics in Producing Strains. *Applied Microbiology*. 501-508.

Voget S., H.L. Steele, W.R. Streit. 2006. Characterization of a Metagenome-Derived Halotolerant Cellulase. *Journal of Biotechnology.* 126:26-36.

Voget, S. C. Leggewie, A. Uesbeck, C. Raasch, K. E. Jaeger, and W. R. Streit. 2003. Prospecting for Novel Biocatalysts in a Soil Metagenome. *Applied and Environmental Microbiology*. 69(10): 6235-6242.

Walter J., M. Mangold, and G. W. Tannock. 2005. Construction, Analysis, and

β-Glucanase Screening of a Bacterial Artificial Chromosome Library from the Large-Bowel Microbiota of Mice. *Applied and Environmental Microbiology.* 71(5): 2347–2354.

Ward, D.M., F. M. Cohan, D. Bhaya, J. F. Heidelberg, M. Kühl and A. Grossman. 2007. Genomics, Environmental Genomics and the Issue of Microbial Species. *Nature*. 1-13.

Warnecke F., P. Luginbühl, N. Ivanova, M. Ghassemian, T.H. Richardson, J.T. Stege, M. Cayouette, A.C. McHardy, G. Djordjevic, N. Aboushadi, R. Sorek, S. G. Tringe, M. Podar, H. Garcia Martin, V. Kunin, D. Dalevi, J. Madejska, E.Kirton, D. Platt, E. Szeto, A. Salamov, K. Barry, N. Mikhailova, N. C. Kyrpides, E. G. Matson, E. A. Ottesen, X. Zhang, M. Hernández, C. Murillo, L.G. Acosta, I. Rigoutsos, G.Tamayo, B.D. Green, C.Chang, E. M. Rubin, E.J. Mathur, D. E. Robertson, P. Hugenholtz and J. R. Leadbette. 2007. Metagenomic and Functional Analysis of Hindgut Microbiota of a Wood-Feeding Higher Termite. *Nature* 450: 560-565.

Waschkowitz T., S. Rockstroh and R. Daniel. 2009. Isolation and Characterization of Metalloproteases with a Novel Domain Structure by Construction and Screening of Metagenomic Libraries. *Applied and Environmental Microbiology*. 75(8): 2506–2516.

Woyke T., H. Teeling3, N. N. Ivanova, M. Huntemann, M. Richter, F. O. Gloeckner, D. Boffelli, .I. J. Anderson, K. W. Barry, H. J. Shapiro, E. Szeto, N. C. Kyrpides, M. Mussmann, R. Amann, C. Bergin., C. Ruehland, E. M. Rubin and N. Dubilier. 2006. Symbiosis Insights Through Metagenomic Analysis of a Microbial Consortium. *Nature*. 443: 950-955

Wright G. D. 2005. Bacterial Resistance to Antibiotics: Enzymatic Degradation and modification. *Advanced Drug Delivery Reviews*. 57: 1451– 1470.

Wright, G. D. 2010. Antibiotic Resistance in the Environment: A Link to the Clinic? *Current Opinion in Microbiology*. 13: 589–594.

Xi, C., Y. Zhang, C. F. Marrs, W. Ye, C. Simon, B. Foxman and J. Nriagu1. 2009. Prevalence of Antibiotic Resistance in Drinking Water Treatment and Distribution Systems. *Applied and Environmental Microbiology*. 75(17): 5714–5718.

Yamane K., Y. Doi, K. Yokoyama, T. Yagi, H. Kurokawa, N. Shibata, K. Shibayama, H. Kato and Y. Arakawa. 2004. Genetic Environments of the *rmtA* Gene in *Pseudomonas aeruginosa* Clinical Isolates. *Antimicrobial Agents and Chemotherapy.* 48(6): 2069-2074.

Yan J., J. Wu, W. Ko, S. Tsai, C. Chuang, H. Wu, Y. Lu and J. Li. 2004. Plasmid-Mediated 16S rRNA Methylases Conferring High-Level Aminoglycoside Resistance in Escherichia coli and Klebsiella Pneumoniae Isolates from Two Taiwanese Hospitals. *Journal of Antimicrobial Chemotherapy*. 54: 1007–1012

Yokoyama K., Y. Doi, K. Yamane, H. Kurokawa, N. Shibata, K. Shibayama, T. Yagi, H. Kato and Y. Arakawa. 2003. Acquisition of 16S rRNA Methylase Gene in Pseudomonas Aeruginosa. *The Lancet*. 362: 1888–93.

Yun J., S. Kang, S. Park, H. Yoon, M. Kim, S. Heu and S. Ryu. 2004. Characterization of a Novel Amylolytic Enzyme Encoded by a Genefrom a Soil-Derived Metagenomic Library. *Applied and Environmental Microbiology*. 70(12): 7229–7235.

Zembower T. R., G. A. Noskin, M. J. Postelnick, C. Nguyen and L. R. Peterson. 1998. The Utility of Aminoglycosides in an Era of Emerging Drug Resistance. *International Journal of Antimicrobial Agents*. 10: 95–105.

Zhang, T., M. Breitbart, W. H. Lee, J. Q. Run, C. L. Wei, S. W. Lin Soh, M. L. Hibberd, E. T. Liu, F. Rohwer, Y. Ruan. 2006. RNA Viral Community in Human Feces: Prevalence of Plant Pathogenic Viruses. *Public Library of Science Biology*. 108-118.

Zhang, X., T. Zhang and H. H. P. Fang. 2009. Antibiotic Resistance Genes in Water Environment. *Applied Microbial Biotechnology*. 82: 397–414.

Zhang Y. and V. N. Gladyshev. 2007. High Content of Proteins Containing 21st and 22nd Amino Acids, Selenocysteine and Pyrrolysine, in a Symbiotic Deltaproteobacterium of Gutless Worm *Olavius algarvensis*. *Nucleic Acids Research*. 35(15): 4952–4963.

Zhou Y., H. Yu, Q. Guo, X. Xu, X. Ye, S. Wu, Y. Guo and M. Wang. 2010. Distribution of 16S rRNA Methylases Among Different Species of Gram-Negative Bacilli with High-Level Resistance to Aminoglycosides. *European Journal of Clinical Microbiology and Infectious Diseases*. 29:1349–1353.