

**MICROBIAL CHARACTERIZATION OF AN UNMINED PORPHYRY DEPOSIT AT
BOSQUE DEL PUEBLO (ADJUNTAS, PR)**

By

Katherine Carrero-Figueroa

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

MASTER OF SCIENCE

in

BIOLOGY

**UNIVERSITY OF PUERTO RICO
MAYAGUEZ CAMPUS
2011**

Approved by:

Carlos J. Santos Flores, Ph.D.
Member, Graduate Committee

Date

Carlos Rodríguez Minguela, Ph.D.
Member, Graduate Committee

Date

Arturo Massol Deyá, Ph.D.
President, Graduate Committee

Date

Dr. Victor Siberio Torres
Representative of Graduate Studies

Date

Nanette Difffoot Carlo, Ph.D.
Chairperson of the Department

Date

ABSTRACT

The results from this work demonstrated the relationship between vegetation cover and the diversity of soil bacterial communities from two sites in an abandoned unmined porphyry deposit at *Bosque del Pueblo, Adjuntas, PR*. The sampling sites were named as *Estación Bosque* (EB), located within the secondary forest and cover by vegetation; and *Estación Mina* (EM) which is deforested and lies adjacent to the forest. The aims of this study were i) to characterize the diversity of soil bacterial community at two sites using PCR-amplified 16S rDNA clone libraries and terminal restriction fragment length polymorphism (T-RFLP, or TRF) analyses, and ii) to correlate the sites' bacterial diversity to physical and chemical properties of soil and water streams at these locations. Both 16S rDNA phylogenetic analysis and T-RFLP profiles revealed that the bacterial communities in secondary forest soil samples were distinct from those in deforested soils and that similarities in bacterial community composition were detected within the forested soil samples as well as within the deforested soil samples. The collection of sequences from *Estación Bosque* were dominated by members of *Acidobacteria* (35%), followed by *Proteobacteria* (31%) members. Whereas the clone sequences detected in *Estación Mina* were more closely related to members of phylum *Chloroflexi* (40%). According to Shannon index (H) the bacterial composition was more diverse in the *Estación Bosque* than in *Estación Mina*. In terms of the physico-chemical properties of soil and water streams at the sites, the organic matter, organic carbon, and total nitrogen content in the forested soil samples were higher than the corresponding content in deforested soil samples. Furthermore, water samples taken from the deforested site exhibited higher conductivity, acidity, and metal ions concentrations than samples from the forested site. It is concluded that a shift in vegetation cover from secondary forest to deforested soil induces changes in the physico-chemical

properties of soil and water streams in *Bosque del Pueblo* that in turn influence the soil bacterial diversity.

Resumen

Los resultados de este trabajo demostraron la relación entre la cobertura vegetal y la diversidad de las comunidades bacterianas en el suelo de dos lugares en un depósito porfídico abandonado y no minado en el Bosque del Pueblo (Adjuntas, PR). Los lugares de muestreo fueron nombrados como: Estación Bosque (EB), localizada dentro del bosque secundario y con cobertura vegetal; y Estación Mina (EM) la cual esta deforestada y localizada adyacente al bosque. Las metas de este estudio fueron i) caracterizar la diversidad bacteriana del suelo en dos estaciones usando los análisis de las librerías de clones del gen 16S rDNA amplificado por PCR y los perfiles del polimorfismo de la longitud del fragmento terminal de restricción (T-RFLP), y ii) correlacionar la diversidad bacteriana de los lugares muestreados a las propiedades físicas y químicas de los suelos y arroyos en estos lugares. Ambos, análisis filogenéticos del 16s rDNA y los perfiles de T-RFLP, revelaron que las comunidades bacterianas en las muestras de suelo del bosque secundario fueron distintas de aquellas en suelos deforestados y que similitudes en la composición bacteriana fueron detectadas dentro de las muestras de suelos forestados, así como, dentro de las muestras de suelo deforestado. La colección de secuencias de la Estación Bosque fue dominada por miembros de *Acidobacteria* (35%), seguido por miembros de Proteobacteria (31%). Mientras que las secuencias de clones detectadas en la Estación Mina fueron relacionadas más cercanamente a miembros del phylum *Chloroflexi* (40%). De acuerdo al índice Shannon (H), la composición bacteriana fue más diversa en la Estación Bosque que en la Estación Mina. En términos de las propiedades físico-químicas del suelo y los arroyos en los lugares muestreados, el contenido de materia orgánica, carbono orgánico y nitrógeno total fueron mayores en las muestras de suelo forestadas que el contenido correspondiente en las muestras de suelos deforestados. Más aun, las muestras de agua tomadas de la estación deforestada

exhibieron mayor conductividad, acidez y concentración de iones metálicos que las muestras de la estación forestada. Se concluye que una alteración en la cobertura vegetal, de bosque secundario a suelo deforestado, induce cambios en las propiedades del suelo y los arroyos del Bosque del Pueblo que a su vez influyen en la diversidad bacteriana del suelo.

DEDICATORY

To Casa Pueblo

ACKNOWLEDGMENTS

I would like to gratefully acknowledge the personnel from Biology Department and the support from University of Puerto Rico Mayaguez Campus in the form of teaching and research assistantships. I express my gratitude to Dr. Carlos J. Santos, Dr. Rafael Montalvo and Dr. Carlos Rodríguez Minguela for their guidance. Special acknowledgement is express to Dr. Arturo Massol, for providing me the opportunity to work in his laboratory, for his guidance and patience.

To all members in the Tropical Microbial Laboratory, thanks for giving me the opportunity to share with all of you. Especially thanks to Ernie Pérez Almodóvar for his help with clone libraries' richness and diversity analyses, Alejandro Caro for assistance with TRFLP analyses and Dubiezel Medina for many useful discussions of phylogenetic analyses. All of them always knew *who, what, where, when* and *how*. Thanks for always being so willing to help. I also want to express my gratitude to Elba Díaz for the atomic absorption determinations of heavy metals, Gladys Toro for running T-RFLP polyacrilamide gels, and Delvis Pérez Almodóvar from the Tropical and Agriculture Research Station (TARS) for soil physicochemical measures.

I would like to thank Dr. José E. Cortés-Figueroa for his very helpful comments, suggestions, improvements, and corrections to the manuscript. To Dr. Ivette García Castro, to whom I am deeply indebted your guidance, support, and good nature, motivated me to pursue graduates studies. To my cousin, Ivelisse Colón, for being a role model during my infancy and adolescence years.

Table of Contents

List of Tables	x
List of Figures	xi
Chapter I: Introduction	1
a. Geology of Puerto Rico.....	1
b. Porphyry copper deposits	2
c. Porphyry copper deposits of Puerto Rico.....	2
d. Rock acid drainage.....	5
e. Soil microbial ecology.....	7
f. Characterization of soil microbial communities by molecular approaches.....	10
g. Site under study.....	11
Objectives	14
Chapter II: Literature Review	15
Chapter III: Materials and Methods.....	21
a. Water sample collection.....	21
b. Soil sample collection.....	22
c. DNA extraction from soil samples.....	23
d. PCR amplification, cloning and sequencing.....	23
e. Phylogenetic analyses of clone libraries.....	25
f. Diversity indices and statistical analysis of the 16S clone libraries.....	26
g. Terminal Restriction Fragment Length Polymorphism (T-RFLP) of the DNA soil samples	27
h. Data analyses of T-RFLP.....	28

Chapter IV: Results	30
a. Physicochemical analysis of soil samples	30
b. Physicochemical analysis of water samples	31
c. Nickel, Copper and Lead concentrations in water samples.....	31
d. DNA extraction and PCR amplification from soil samples.....	35
e. Phylogenetic groups of Bacteria detected in soil samples.....	36
f. Diversity statistics of the clone libraries.....	40
g. Phylogenetic Analyses of clone libraries.....	47
h. T-RFLP analyses of the soil bacterial communities from Bosque del Pueblo.....	64
Chapter V: Discussion	72
Conclusions	90
Literature Cited	92
Appendixes.....	102
Appendix I.....	102
Appendix II.....	106
Appendix III.....	109
Appendix IV.....	123
Appendix V.....	136
Appendix VI.....	186

List of Tables

Table 1. Soil samples chemical and physical parameters at two locations in Bosque del Pueblo, Adjuntas, PR.	30
Table 2. Elemental analysis of nickel and copper content in water samples at Bosque del Pueblo, Adjuntas, PR.	34
Table 3. Diversity and richness estimates from 16S rDNA clone libraries	41
Table 4. Occurrence of T-RFs in T-RFLP profiles of 16S rDNA amplified gene in soil samples at Bosque del Pueblo, Adjuntas, PR.	64
Table 5. TRF diversity data from Bosque del Pueblo's soil samples bacterial community profiles	70
Table 6. Similarity value (%) of Bosque del Pueblo's soil samples T-RFLP bacterial community profiles	70

List of Figures

Figure 1. Map of the geology of Puerto Rico	3
Figure 2. Sampling site in the Estación Mina at Bosque del Pueblo Adjuntas, PR	13
Figure 3. Temporal variation of (a) temperature, (b) conductivity, and (c) pH values of water samples obtained at Bosque del Pueblo Adjuntas, PR	33
Figure 4. DNA isolated from soil samples at Bosque del Pueblo	36
Figure 5. Distribution of phylogenetic groups of <i>Bacteria</i> detected in soil samples by 16S rRNA clone libraries. (a) EBS, (b) EBP, (c) EMS and (d) EMP	38
Figure 6. Plots of Good Coverage index vs. Library size (Good Coverage Curve) for the 5' end terminal data-set of the 16S rDNA clone libraries.	42
Figure 7. Plots of Good Coverage index vs. Library size (Good Coverage Curve) for the 3' end terminal data-set of the 16S rDNA clone libraries.	43
Figure 8. Plots of OTUs number vs. Library size sampling for 16S rDNA clone libraries: EBS, EBP, EMS and EMP.	45
Figure 9a. Phylogenetic tree corresponding to <i>Acidobacteria</i> 5' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.	52
Figure 9b. Phylogenetic tree corresponding to <i>Proteobacteria</i> 5' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.	53
Figure 9c. Phylogenetic tree corresponding to <i>Chloroflexi</i> 5' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.	54
Figure 9d. Phylogenetic tree corresponding to various phyla 5' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.	55
Figure 10a. Phylogenetic tree corresponding to <i>Acidobacteria</i> 3' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.	61
Figure 10b. Phylogenetic tree corresponding to <i>Proteobacteria</i> 3' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.	62
Figure 10c. Phylogenetic tree corresponding to various phyla 3' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.	63

Figure 11. Representative TRF profiles of bacterial communities from Bosque del Pueblo’s soil samples a) EBS, b) EBP, c) EMS and d) EMP for *Rsa I* restriction analyses..... **67**

Figure 12. Representative TRF profiles of bacterial communities from Bosque del Pueblo’s soil samples a) EBS, b) EBP, c) EMS and d) EMP for *Rsa I* restriction analyses. **68**

Figure 13. Representative TRF profiles of bacterial communities from Bosque del Pueblo’s soil samples a) EBS, b) EBP, c) EMS and d) EMP for *Rsa I* restriction analyses..... **69**

Figure 14. Dendrogram based on Morisita similarity comparisons for EBS, EBP, EMS, EMP soil bacterial communities. **71**

CHAPTER I

INTRODUCTION

Geology of Puerto Rico

Puerto Rico is the easternmost island of the Greater Antilles in the Caribbean Sea. It lies in the trade-wind belt 18° 15' north of the equator and 66° 30' west of the Greenwich meridian. Puerto Rico is a complex island arc terrain with a geological record of about 195 million years. Together with the northern Virgin Islands, it represents the subaerially exposed parts of the Puerto Rico-Virgin Islands microplate (PRVI) (Byrne et al., 1985), which lies within the seismically active Caribbean-North American Plate boundary zone. The northern border of the Caribbean Plate is marked by the 8,516 meter deep Puerto Rico Trench, the deepest point in the Atlantic Ocean. For tens of millions of years this trench was the site of the subduction of the North American Plate beneath the northward moving Caribbean Plate. The subduction of the North American Plate into the Puerto Rico Trench resulted in a series of volcanic eruptions that piled tuff, ash, and magma onto the ocean floor. These highly folded, faulted, and eroded volcanic and sedimentary rocks now form the core of the Cordillera Central, an east-west trending spine of mountains that cross the island from west to east (Barabas, 1982).

The geology of Puerto Rico is dominated by an east-west belt of Cretaceous to Eocene volcanic and sediments intruded by irregular bodies of ultramafic to granitic rocks. The volcanic-plutonic complex underlies about two-thirds of the island and makes up its mountainous central core. Flat-lying Oligocene limestones, in turn overlain by recent alluvial, beach and swamp deposits, form flanking east-west belts along the north and south coasts.

Porphyry copper deposits

Porphyry ore deposits are the most important source of valuable ore minerals, usually copper or gold. Nearly 60% of the world's copper is extracted from porphyry copper deposits (Cline, 2003).

Porphyry copper deposits are large bodies of rocks that occur in the roof zones of igneous intrusions in island arc and continental margins settings. They are characterized by intense hydrothermal alteration. Many of these hydrothermal deposits are associated with boiling hydrothermal fluids transport metals upward. The principal component of hydrothermal solutions is water containing dissolved salts such as NaCl, KCl, CaSO₂ and CaCl₂. Hydrothermal fluids are capable of dissolving small amounts of elements. The fluids react with the rocks, cool and change its chemistry; as a result chalcopyrite (CuFeS₂) and other ore minerals are precipitated within the rock. The porphyry copper deposits typically contain copper, molybdenum and gold, disseminated through the host rock.

Porphyry copper deposits of Puerto Rico

The existences of copper deposits throughout the central volcanic-plutonic belt have been known since colonial times. In fact, Puerto Rico was a significant source of gold for the Spanish during the 16th century.

All the known porphyry copper deposits in Puerto Rico occur along the southern margin of the Utuado batholiths in the west-central part of the island, Figure 1. These deposits were associated with a series of hydrothermally altered bodies of quartz diorite stocks of Eocene age that have intruded the batholiths or volcanic rocks adjacent to the batholiths.

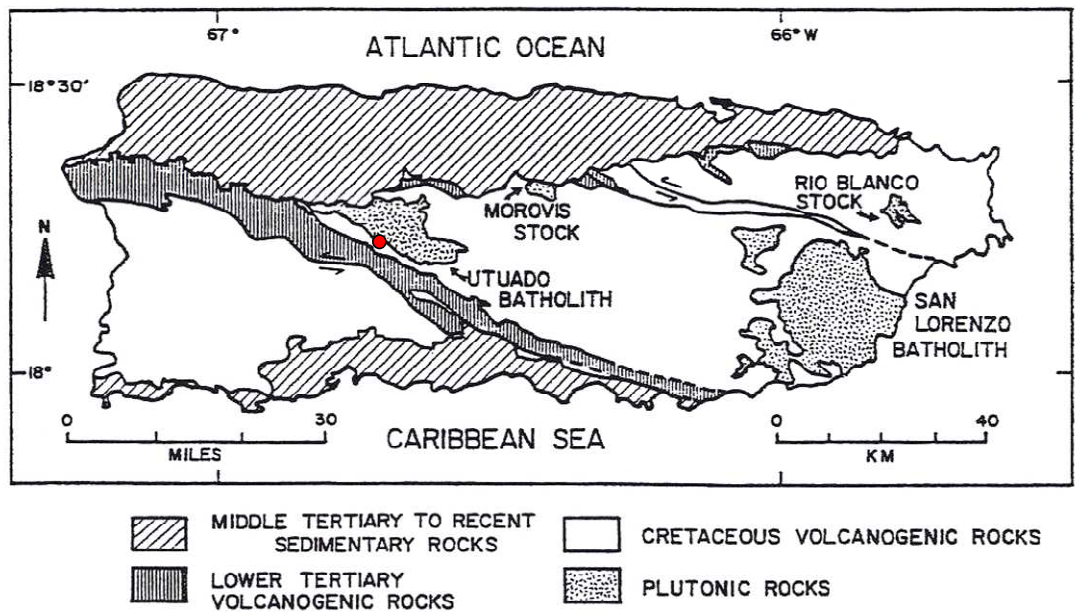


Figure 1. Simplified geologic map of Puerto Rico from Barabas, 1982. The red circle indicates the location of the Rio Viví porphyry copper deposits.

The principal porphyry copper deposits of Puerto Rico were discovered during the late 1950's. In 1958, Kennett Copper Corporation implemented an exploration program in the Lares' area northwest of Utuado. The copper Creek and Laundry creek porphyry copper prospects were identified near the Tanamá River. Later, between 1961 and 1965, Tanamá and Helechos deposits were discovered and explored by drilling. Reserves at Río Tanamá mineralized area were estimated at about 139 millions tons grading 0.64 percent copper.

As a result of explorations that began in 1957, an economically recoverable amount of copper was found approximately 2.75 miles northeast of the town of Adjuntas by W.R. Bergey. Two ore bodies, contained in quartz-diorite intrusions at the edge of the Utuado Batholith, were identified and referred to as the Cala Abajo ore body and the Piedra Hueca ore body. A third ore body was discovered 8 km southeast of Utuado which was given the name Sapo Alegre. The

ages of the deposits are estimated to be between 39.1 to 43.7 Ma (mega-annum= million years) for Calá Abajo, between 35 to 43.1 Ma for Piedra Hueca, and between 41.8 to 42.1 Ma for Sapo Alegre (Barabas, 1982). The Cala Abajo, Piedra Hueca and the Sapo Alegre are grouped within 0.5 km of each other along the Río Viví.

The Cala Abajo ore body is irregularly elongated, approximately 1,200 x 200 meters in surface dimension, and contains approximately 71 million tons of ore assayed at an average of 0.82 percent copper. The upper portion of the ore body is a soft, spunky material containing copper primarily as covellite (CuS) at one percent. Lower layers of ore become progressively harder and contain smaller percentages of copper, mostly as chalcopyrite (CuFeS₂).

The Piedra Hueca ore body is irregularly oval in surface dimension, is approximately 450 x 200 meters, and contains about 33 million tons of ore at approximately 0.82 percent copper. Copper is present mainly as chalcopyrite with some pyrite involved and is contained in very hard plutonic rock.

The Sapo Alegre is an elongated porphyry copper-molybdenite deposit, approximately 500 x 100 meters in plan dimension. Sapo Alegre differs from Cala Abajo and Piedra Hueca deposits in having high molybdenum content and a low copper content. Subsequent geological studies revealed that Sapo Alegre was too small for financially viable mining (Cox et al., 1975).

Amax and Kennett performed economic and environmental feasibility studies on Cala Abajo and Piedra Hueca deposits. In 1967, these companies presented a series of proposals to the government of the Commonwealth of Puerto Rico for a large-scale mining of these reserves. The companies failed to reach an agreement for a mining lease with the Commonwealth of Puerto Rico. In 1987 they abandoned interest in the project.

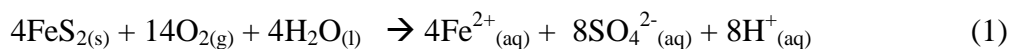
In 1992, the company Southern Gold resources (USA), Inc. received permission from the Puerto Rico Department of Natural Resources to explore the Cala Abajo deposits. Later, in 1995 the permit was amended and enlarged to cover the adjacent Piedra Hueca deposit. Based on results from these explorations Southern Gold Resources proposed to mine the deposits by standard open-pit mining techniques.

In the early 1980s, the community of Puerto Rico under the leadership of Adjuntas' Art and Culture Workshop, today known as Casa Pueblo, opposed the proposed mining activity on the account of the negative impact of the mining to the environment. In June 1995, the Puerto Rico legislative amended Law number 9 of August 18, 1933 prohibiting any kind of mining activities in Puerto Rico. Finally, in 1996 the communitarian organization Casa Pueblo reached the position of Forest administrator and the zone, previously impacted by mining activities, was transformed by law in a Puerto Rican protected forest, *Bosque del Pueblo*.

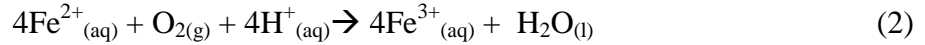
Rock Acid drainage

Pyrite (FeS_2) is the most abundant sulfide mineral in earth's crust (Edwards et al., 1998). When the solid-phase pyrite rock is exposed to water and atmospheric oxygen, metals ions and hydrogen ions are released into the environment. The products generated by this reaction alter the pH and heavy metal composition of nearby soils and streams. When acid is generated at mining sites it is referred as acid-mine drainage; whereas acid-rock drainage occurs naturally within some environments as part of the rock weathering process usually seen within rocks containing an abundance of sulfide minerals.

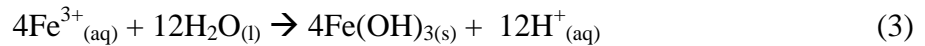
The oxidation of pyrite begins with the following reaction:



As a result of this reaction, the concentration of reduced Iron II ions and hydrogen ions increases significantly in the nearby soils and streams. Iron II ions are oxidized to Iron III ions through the reaction:

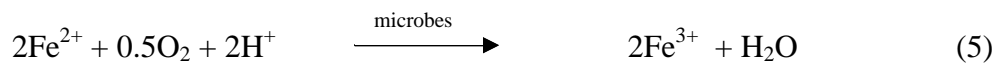


Ferric ions (III) are soluble in acidic runoff with pH values below 3. When the acidic runoff is diluted by other water sources, the pH will increase. This increase in pH causes iron III ions to precipitate out of solution in the form of Iron III hydroxide as described by reaction:



The iron-rich acidic drainage kills aquatic life and renders the contaminated stream unsuitable as a water supply or for recreational use.

Some microbial processes are very important in catalyzing the chemical reactions that produce acid mine drainage (Edwards et al., 2000). The oxidation of metal sulfides proceeds by various known mechanisms. Microorganisms obtain energy by the oxidation of either iron or sulfur; the reactions may be represented by the following schemes:



Elemental sulfur is oxidized to sulfate by sulfur-oxidizing microorganisms such as *Acidithiobacillus thiooxidans* (previously *Thiobacillus thiooxidans*) or *Acidithiobacillus caldus* (previously *Thiobacillus caldus*). The ferrous iron generated in reaction (1) is reoxidized to ferric iron by iron-oxidizing microorganisms such as *Acidithiobacillus ferrooxidans*, *Leptospirillum ferrooxidans*, *Leptospirillum ferriphilum* and *Sulfobacillus thermosulfidooxidans*. The role of iron-sulfur oxidizing microorganisms is to provide an acidic environment (schemes 4) to accelerate the rate of oxidation of Fe^{2+} to Fe^{3+} (schemes 5).

Microorganisms are widely used in bioleaching processes to convert insoluble metal sulfides to water soluble metal sulfates or as a biooxidation pretreatment process to expose the mineral (Rawling, 2005).

Soil microbial ecology

The vast majority of life on Earth is microbial, distributed among three primary domains: Archaea, Bacteria and Eucarya (Pace, 1997; Hugenholtz et al., 1998). The interactions between microorganisms and their non-living surrounding have been going on for about 3.8 billion years (Torsvik et al., 2002). The inorganic nutrition sustained the earliest form of life on Earth. Later the phototrophy and organotrophy were established.

The soil is a very complex habitat in which the microbial population is very diverse. In 1996, Torsvik and collaborators calculated the presence of about 6,000 different bacterial genomes per gram of soil (Torsvik et al., 1996). Another important characteristic of soil is that less than 5% of the overall available space in soil is occupied by living microorganisms (Nannipieri et al., 2003). Important biological molecules such as proteins and nucleic acids are absorbed and protected against degradation by nucleases into the solid phase of soil. Biological molecules can be incorporated by bacterial cells and inserted into its chromosomal DNA in a process known as horizontal gene transfer.

Many chemical reactions are catalyzed in the surfaces of soil mineral components. Soil microorganisms, especially algae, bacteria, fungi and lichens, contribute to the soil formation process. These organisms are able to solubilize silicates and other minerals through production of organics acids and chelating agents (Atlas and Bartha, 1997).

Microorganisms can break down decayed plant and animal tissue into energy, carbon, and nutrients sources. The organic matter generated by microbial decomposition improves the water-holding capacity and the ion exchange capacity of soil. In the nitrogen cycles, soil microorganisms mediate the reduction of gaseous nitrogen to ammonia. They also contribute to denitrification process by converting nitrate through nitrite to nitric oxide (NO), and nitrous oxide (N₂O) to molecular nitrogen.

Plant community composition greatly influenced the community composition of soil microorganisms. Microorganisms in soil have sensory mechanisms to recognized nutrients exudates by plants. These exudates act as chemoattractants for bacteria. Low molecular weight exudates include amino acids, organic acids, sugar, aromatics and various other secondary metabolites, whereas high molecular weight exudates include polysaccharides and proteins (Brencic and Winans, 2005) The biomass and activity of microorganisms in soil around plant roots is termed rhizosphere and was first described by Hiltner in 1904 (Bonkowski, 2004). The exudates liberated from plants create a favorable habitat for microbial growth, which benefit by obtaining carbon and other nutrients from their hosts. Within the rhizosphere the density of microorganisms is higher than in the nearby bulk soil. Because different plants exudates a great diversity of organics components to soil the interaction between plant and microbes is specific. During plant development, microorganisms make adjustment to their own physiology in response to changes in the physiology of their host. Microorganisms also undergo successional changes as the plant grows (Atlas and Bartha, 1997).

Just as plants have an effect on the microbial community in soil, microorganisms also interact with plants in a variety of ways (White et al., 2005). Plants obtain essential minerals elements from soil. These include six macronutrients (N, K, P, S, Mg and Cu) and eight

micronutrients (B, Cl, Cu, Fe, Mn, Mo, Ni and Zn). Microbes can stimulate plant growth by maintaining the recycling and solubilization of mineral nutrients; through the synthesis of vitamins, amino acids; providing resistance against pathogens; and to confer tolerance to toxic compounds (White et al., 2005).

Although nitrogen is an essential plant nutrient, it is metabolically unavailable directly to higher plants and animals. Microbes in soil are capable of taking atmospheric dinitrogen (diazotrophs) and combining it with hydrogen to make ammonia by the enzyme nitrogenase, in a process known as Biological Nitrogen Fixation (BNF). This process supplies the fixed forms of nitrogen required for bacterial and plant growth (Atlas and Bartha, 1997). Most nitrogen-fixing bacteria are free-living in soil, but some are in association with plant's roots forming a tumor like growth called a nodule. Nitrogen fixation by free-living, associated and symbiotic diazotroph is the most important source of nitrogen input in natural ecosystems (Burgmann et al., 2004; Poly et al., 2001).

Many nitrogen-fixing bacteria have been shown to fix nitrogen more efficiently at low oxygen concentrations. These conditions are found in subsoil and sediment habitats. Some genera of free-living nitrogen fixing soil bacteria are *Azotobacter*, *Beijerinckia*, *Azospirillum*, *Desulfovibrio*, *Desulfotomaculum*, *Klebsiella*, *Bacillus*, *Clostridium*, *Pseudomonas*, *Vibrio* and *Thiobacillus* (Atlas and Bartha, 1997). Actinomycetales are also able to fix atmospheric nitrogen as free-living organisms or in association with compatible plants. Although free living nitrogen fixing bacteria are widespread in soils, their rates of nitrogen fixation are higher in soils with vegetation cover than in nude soil.

The participation of microorganisms in the biochemical cycles that operate on Earth motivates the interest to study the distribution of microbial activities across different

environments. Traditionally, an ecosystem is described by measuring the number, frequency, and functional role of species present. Sometimes, this strategy is not practical for microbial ecologist. This is particularly true in soil microbial communities, which are the most complex and diverse (Kuske et al., 1997). The information about how microorganisms influenced and are influenced by the lithosphere environments or about the details of how chemical processes carried out by microorganisms on surfaces and at subsurfaces contribute to geobiological events is scarce.

Characterization of soil microbial communities by molecular approaches

The number of microorganisms within a community can be assayed through cultivation and cultivation-independent approaches. However, many studies have confirmed that cultured isolates are unrepresentative of the soil bacterial community, less than 1% (Liles et al., 2003; Ritchie et al., 2000; Nüsslein and Tiedje, 1998). Culture-independent phylogenetic techniques based on surveys of genes after polymerase chain reaction (PCR) amplification provide a more complete view of the microorganisms that are present in the environment. In the 1980s, Woese and collaborators introduced the use of rRNA genes as molecular chronometers (Woese, 1987). Today, 16S rRNA gene is the most commonly used of the molecular markers to relate prokaryotic organisms. The 16S rRNA gene (i) occurs in all prokaryotic organisms, (ii) shows a high degree of conservation, (iii) is large enough to contain useful phylogenetic information and (iv) is present in low-copy number within the bacterial genome.

Various molecular techniques, such as 16S rRNA cloning (Dunbar et al., 2000), denaturing gradient gel electrophoresis (Peixoto et al., 2002), and TRFLP analysis (Blackwood et al., 2003) have been used to study the bacterial community structure in soil environments.

These approaches allow us to make comparisons of the following three elements of diversity in environmental microbial communities: the types of bacteria present (composition), the number of types (richness), and the frequency distribution or relative abundance of types (structure).

Singh (2006) developed a fingerprinting method, multiplex terminal restriction fragment length polymorphism (M-TRFLP), in which the diversity and structure community of two or more microbial taxa can be simultaneous analyze (Singh et al., 2006). Investigators used soil samples collected from different habitats to examine the reproducibility and robustness of the method. The DNA from soil samples was PCR amplified using taxon-specific primers for bacteria, archaea and fungi. The same soil DNA samples undergo a multiplex PCR with primers for all three taxa. TRFLP profiles were generated for: PCR products for individual taxa, pooled PCR products generated using individual PCR primers and multiplex PCR products. The results showed that single TRFLP and M-TRFLP analysis produced almost identical profiles with a similarity value over 90%.

Site under study

The natural regeneration of vegetation of lands beaten by mining exploration at *Bosque del Pueblo* has been very slow. Even after more four decades, some areas at *Bosque del Pueblo* remained bare ground where vegetation had not regenerated. The vegetation regeneration following other types of land uses have indicated much faster forest recovery than on abandoned mining areas. In 2001 Peterson and Heemskerk (Peterson and Heemskerk, 2001), found that vegetation cover between abandoned mined sites and forest in the Amazon was significantly different, but little differences were observed between abandoned mining areas. The impact of mining activities is probably more severe than other types of anthropogenic disturbance. The

massive repeated soil movement affects soil, water quality and microbial diversity, which greatly slow vegetation regeneration, Figure 2.

Vegetation provides primary resources for growth and therefore is one environmental factor thought to be a major determinant on microbial community structure (Nusslein and Tiedje, 1999). Others environmental factors that influence microbial community structure in soil are soil moisture, temperature, pH, organic carbon concentration, and organic matter content (Ellis et al., 2003; Girvan et al., 2003; Wawrik et al., 2005).

This study represents the first approach using culture-independent techniques to examine the microbial diversity in an abandoned unmined porphyry deposits at *Bosque del Pueblo*, Adjuntas PR. We investigated the influence of vegetation cover on the structure and composition of soil microbial communities.

a.



b.



Figure 2. Sampling site at Bosque del Pueblo Adjuntas, PR. Photos a. and b. were taken in 2002 and 2005, respectively.

OBJECTIVES

The specific aims of this study are: (i) to characterize the soil bacterial community diversity in an abandoned unmined porphyry deposits at *Bosque del Pueblo*, Adjuntas PR by performing PCR-amplified 16S rRNA clone libraries and Terminal Restriction Fragment Length Polymorphism (T-RFLP) analysis and (ii) to correlate the microbial diversity with physical chemical properties of soil and water streams at these locations.

CHAPTER II

LITERATURE REVIEW

A fundamental question in ecology is how many species occur within a given area. Although tropical forests cover only 7% of the Earth's land surface, they contain more than half of the world's species. Tropical forests are well known for their high diversity of flora and fauna as the most genetically diverse terrestrial communities on Earth (Plotkin et al., 2000). However, very few studies have examined specific soil microbial populations in tropical forests.

Soil is a highly heterogeneous matrix. The proportion of cultivable microorganisms is estimated as less of 1% of the bacteria present in environmental samples. Our perspective on microbial diversity has improved enormously over the past few decades due in significant part to the application of culture-independent molecular ecological methods, especially those based on surveys of genes after PCR amplification. Studies using reassociation kinetics in DNA extracted from soil showed the presence of 2,000 to 18,000 different bacterial genomes in a gram of soil (Torsvik et al., 1996; Sandaa et al., 1999)

In 1997, Borneman and Triplett reported the first description of the microbial diversity in Amazonian soils using clone libraries of the small-subunit rRNA genes (Borneman and Triplett, 1997). The DNA used to construct the libraries was extracted from a mature forest soil and an adjacent pasture soil. Fifty pasture and fifty forest clones were sequenced. None of the 100 SSU rRNA clones sequenced were duplicates and none of the sequences had been previously described. They study showed a high diversity of microorganisms in the soils from eastern Amazonia.

Nusslein and Tiedje (1999) investigated the influence of vegetative cover on the structure and composition of a microbial community by comparing two adjacent soils from the Kohala

Forest reserve on the Big Island of Hawaii (Nusslein and Tiedje, 1999). One soil habitat was continuously covered by a native tropical forest and the other soil habitat was covered by a grass pasture. DNA was extracted from both habitats and fractionated by its guanine (G) and cytosine (C) content. The amount of DNA was quantified in fractions with 63% G+C and 35% G+ C contents. The small subunit rRNA genes were amplified by PCR from both DNA fractions and used to construct clone libraries. The ribosomal DNA clone libraries were screened by amplified rDNA restriction analysis (ARDRA) to determine pattern abundance profiles. They also performed a phylogenetic analysis of the dominant clones in the 63% G+C fraction from each soil samples. A significantly higher G+C content profile was found in the pasture soil community than in the forest community. The ARDRA abundance patterns showed differences between habitats. The dominant clones from the forest soil were related to members of the phylum *Fibrobacter* while the dominant phylotypes of the pasture soil were found to be related to members of the *Proteobacteria*. All methods used in this study showed clear differences in microbial communities between the forest and pasture soils but little differences were observed among replicates from the same soil sample. This study suggests that at least half of the total soil microbial mass was replaced as a consequence of changes in vegetation cover.

A similar study was done by Nusslein and Tiedje (1998) at an undisturbed montane rainforest on the island of Hawaii within Volcano National Park (Nusslein and Tiedje, 1998). They analyzed the soil bacterial community diversity of a 200 years old tropical rainforest soil formed from volcanic ash. The soil DNA was fractionated on the basis of guanine-plus-cytosine content. The total soil DNA and two discrete fractions, one with 63% G+C content and one with 35% content, were used to construct clone libraries of the SSU rRNA genes. To screen for SSU rDNA diversity an amplified ribosomal DNA (rDNA) restriction analysis of the amplified inserts

was performed. The restriction analysis from unfractionated soil DNA produced 64 different restriction patterns for the 67 clone sequences examined. The 63% G+C fraction resulted in 46 different patterns for the 76 clone sequences examined. Of the 59 clones analyzed from the 35% G+C fraction, 47 different patterns were obtained. The amplified inserts were also sequenced and analyzed in order to determine phylogenetic affiliations. The majority of the soil DNA was found with G+C contents ranging from 52 to 68% but minor quantity of DNA was in the range from 30 to 50%. The nucleotide sequence analysis of the rDNA clones identified taxa that were expected for DNAs having G+C contents of 63% and 35%. Phylotypes in the 63% fraction were members of the *Pseudomonas*, *Rhizobium-Agrobacterium*, and *Rhodospirillum* assemblages, while phylotypes in the 35% G+C fraction were affiliated with *Clostridium* assemblage.

Several studies (Peterson and Heemskerk, 2001; Edwards et al., 1998) had been done at acid mine drainage production sites to better understand the pyrites dissolution process and to determine which organisms are actively involved in the acid production. Distribution and abundance of *Thiobacillus ferrooxidans* and *Leptospirillum ferrooxidans* were analyzed by fluorescent *in situ* hybridization (FISH) at Iron Mountain, California (Edwards et al., 1998). Although both species had been considered to play key roles in pyrite dissolution, they found that these species were not directly involved in acid generation. *T. ferrooxidans* was found in peripheral slime-based communities with pH values over 1.3 and temperature below 30°C but was not present at the acid generation site with pH values between 0.3 and 0.7 and a temperature between 30°C to 50°C. *L. ferrooxidans* was the predominant form of microbial life associated with the ore body, at a temperature over 40°C and pH values between 0.4 and 0.7. They also used domain-level probes to evaluate the proportion of all cells in the domains *Bacteria*, *Eukarya* and *Archaea*. At least 75% of cells belonged to Bacteria at both environments.

Bond and coworkers (2000) investigated microbial communities at an extreme acid mine drainage (AMD) production site located in Iron Mountain mine, California. Environmental samples were collected at five locations within the sulfide ore body where pyrite actively dissolves and analyzed by FISH. The Cy 3- labeled oligonucleotide probes used were specific for AMD organisms previously detected at the Iron Mountain mine. The majority of the organisms detected in these microbial communities belong to species of *Ferroplasma*, *Leptospirillum*, *Sulfobacillus* and to a lesser extent *Acidimicrobium*. Most of the organisms detected at the mine were uncultured or recently described. They suggest that geochemical and physical properties at the sampled environments determinate the microbial community composition.

Ferroplasma acidarmanus was isolated and characterized in 1997 from Richmond five-way at the Iron Mountain, California (Edwards et al., 2000). At the time of collection temperature, pH, conductivity, and iron concentrations measurements were 40°C, 0 to 1, 120mS cm⁻¹ and 111g liter⁻¹, respectively. In this study, enrichment cultures were prepared with sediments and mine waters collected. DNA was extracted from two of the enrichments cultures and clones libraries were constructed. The representative clones from each library were sequenced. The majority of the clones isolated formed a cluster with *Ferroplasma acidiphilum*. Fluorescent in situ hybridization analysis with a genus-specific probe FER656 confirmed that isolates belong to *Ferroplasma* genus. Edwards and collaborators suggested a new species name to the isolate because its physiology differs from that of *F. acidiphilum*. In order to determine the proportion of *F. acidiphilum* at Iron Mountain sediment and water samples were analyzed by FISH. Around 85% of the total population in a biofilm hybridized with the *Ferroplasma* specific probe FER656, the remainder of the biofilm consisted of eukaryote filaments.

The microbial community from a natural acidophilic biofilm of the Richmond mine at Iron Mountain was studied using random shotgun sequencing approach (Tyson et al., 2004). Briefly, with the whole-genome shotgun analysis, the DNA in a sample is shattered into fragments of specific sizes, sequenced each one, and then assembled these sequences together by matching the ends of the DNA with a powerful overlap-hunting computer program. In principle, this approach allows the reconstruction of entire genomes of the different organisms in a sample. Genome reconstruction at this site was possible because the biofilm was dominated by a low number of species and a low frequency of genomic rearrangements, gene insertions and deletions. *Leptospirillum* group II and *Ferroplasma* type II genomes were near-completely recovered. Three other genomes were partially recovered. However, the microbial community in soil environments was too diverse. Studies at the Iron Mountain revealed that microorganisms not previously reported at AMD sites, such as novel organisms from the *Leptospirillum*, *Thermoplasmales*, *Sulfobacillus*, *Acidimicrobium*, and *Acidiphilium* groups are important contributors to the oxidative portion of the terrestrial sulfur and iron geochemical cycles.

Rosado et al. (2002) used *rpoB* and 16S rRNA genes as biomarkers to analyze microbial community from a Brazilian tropical soil by PCR and DGGE methods. From the study they concluded: (i) the fingerprints generated by *rpoB* PCR-DGGE were different from the 16S rDNA PCR-DGGE fingerprints, (ii) both the *rpoB* and the 16S rDNA communities profiles were affected by soil depth but no differences were observed between dry and wet season and (iii) the single copy of *rpoB* gene rather than the multiple copies of 16S rRNA gene in bacterial genomes resulted in single bands on *rpoB* DGGE gels which provided a more accurate reflection of microbial communities than 16S rDNA-DGGE gels.

Molecular ecological studies that rely on probes or primers specific for the SSU rRNA gene, such as PCR-amplification (Janssen, 2006), fluorescent *in situ* hybridization (Edwards et al., 2000; Cottrell et al., 2000) and slot blot hybridizations have been provided invaluable information about the diversity and abundance of *Acidobacteria*. However, little is known about the genetic and physiology of this uncharacterized group. *Acidobacteria* are expected to have an important ecological role in soil habitats and an extensive metabolic versatility. 16S rDNA clone libraries from nutrient-poor calcareous grassland from Germany revealed that around 15% of all sequences recovered from this habitat belonged to Acidobacteria (Quaiser et al., 2003). In order to obtain additional information about *Acidobacteria*, an environmental genomic approach based on large-insert libraries was conducted from DNA of this calcareous grassland soil. Specific 16S rDNA probes were used to identify genomic fragments of *Acidobacteria*. Six acidobacterial clones were sequenced, representing in total more than 210,000 base pairs (bp).

Since its first use with soil samples, (Osborn, 2000), the T-RFLP analysis of PCR-amplified genes has become one of the most powerful methods in microbial ecology to obtain a genetic fingerprint of the composition of a microbial community (Blackwood et al., 2003). Briefly, the community DNA was PCR-amplified using one or more fluorescently labeled primers, followed by restriction endonuclease digestion and the resulting patterns were analyzed using a DNA sequencer. In principle, each terminal restriction fragment correspond can be defined as a single population within a community (Moyer, 2003).

CHAPTER III

MATERIALS AND METHODS

Sampling Site

Two sampling sites were established at *Bosque del Pueblo* (18°11'42.30"N, 66°41'11.85" W), Adjuntas Puerto Rico. The sampling sites were referred to as *Estación Bosque* (EB) and *Estación Mina* (EM). *Estación Bosque* is located within the secondary forest and cover by tropical vegetation. *Estación Mina* was cleared in the 1960's for mining activities and lies adjacent to the forest.

Water Sample Collection

Sample collection and environmental measurements of water were carried out monthly from January to November 2005.

Water samples of approximately 250 ml were collected in triplicate at each sampling site at ten minutes intervals. Conductivity, pH, and temperature measurements were made *in situ* with a Horiba U-10 Water Quality Checker. Electrodes were calibrated according to manufacturer's instructions.

For heavy metals analyses, water samples were collected in triplicate at each sampling site at ten minutes intervals. Water samples were collected in 500 mL polyethylene bottles that were first doubly rinsed with sample water. After collection, water samples were preserved by acidification using concentrated nitric acid (HNO₃). Water samples were transported on ice to the laboratory and refrigerated at about 4°C until the analyses were performed. Each sample of 500-mL of water was transferred into a 600mL-beaker contains 3.0 mL of concentrated nitric acid. Water samples were evaporated by heating under boiling point with a hot plate, until a volume of 15-20 mL of sample remained in the beaker. Samples were allowed to cool and an

additional 3.0 mL of concentrated HNO₃ were added. Samples were covered with a watch glass and heated on a hot plate until boiling. Samples were passed through Whatman #40 filters and washed with 10% HCl. Additional 10% HCl was added analytically to a final volume of 50.00 mL. Samples were sealed until analysis. Atomic-absorption determinations for copper (Cu), nickel (Ni) and lead (Pb) were obtained with a Perkin-Elmer Model Analyst 100 spectrophotometer, using an air-acetylene flame. Standards were prepared by dilution from 1,000 ppm stock solutions (Fisher Sci., ACS grade). The concentration of acid-extractable elements were determinate as described previously (Massol and Díaz, 2003).

Statistical significance for physical and chemical parameters for water samples from *Estación Bosque* and *Estación Mina* were assessed by a one-way analysis of variance (ANOVA). One-way ANOVA was performed using a software program posted at: [<http://faculty.vassar.edu/lowry//anova1u.html>].

Soil Sample Collection

At *Estación Mina* and *Estación Bosque* sampling sites, one square meter quadrant was randomly located. Ten soil cores measuring 10 cm deep and 2.54 cm wide were collected using soil-coring devices that encase the soil in a stainless steel sleeve. The soil samples were removed from the cores, and the top (0-5 cm) section and bottom (5-10 cm) section were separated from each other. Soil subsamples collected at the same section were mixed and placed in sterile Nalgene™ bags. EBS and EBP soil samples were obtained from the same location (*Estación Bosque*) but at different depths, 0-5 cm and 5-10 cm, respectively. Likewise EMS and EMP were obtained at different depths of the same location (*Estación Mina*), 0-5 and 5-10 cm, respectively. Samples were immediately placed on ice, transported to the laboratory, and stored

at -20°C until further analysis. The physicochemical parameters in soil samples were determined by standard methods at the Tropical Agriculture Research Station (TARS), Mayagüez, Puerto Rico.

DNA Extraction From Soil

Nucleic acids were extracted from 1.0 g aliquots of each of the four soil samples. DNA extraction was performed with the Ultra Clean Soil DNA kit™ (MoBio Laboratories®), according to the manufacturer's protocol.

PCR Amplification, Cloning and Sequencing:

PCR amplification of bacterial 16S rDNA was carried out by using the forward primer 8F (5'-AGAGTTTGATCMTGGCTCAG-3') and the reverse 1392R (5'-ACGGGCGGTGTGTACA-3') universal primers. Amplification was performed in a 25 µl total volume reaction mixture containing ~80 ng of soil DNA, 5.0 µl of 10X Go Taq polymerase reaction buffer B, 0.85 µl of 25mM MgCl₂, 2.0 µl of dNTP's mix [2.5 mM (1:1:1:1 proportion)], 0.5 µl of each 16S rDNA primer [50 pmol/ µl], 0.25 µl of 1.25U Go Taq DNA polymerase (Promega®) and ddH₂O. The amplification from total environmental DNA extracts were performed with a Perkin Elmer Gene Amp PCR System 2400™ as follows: a denaturation step of 2 minutes at 95°C, thirty-five cycles of denaturation at 94 °C for 1 min, primer annealing at 52 °C for 1 min, and primer extension at 72 °C for 2 min; followed by a final extension step of 10 minutes at 72 °C. Five replicate reactions were conducted for each of the four soil DNAs and then pooled to neutralize the effects of bias in any single thermal cycling reaction. The amplified products were visualized by 1.0% [wt/vol] agarose gel in 1X Tris-Acetate-EDTA buffer. The gel

was stained with ethidium bromide and then photographed under UV light. A 1.5-kb fragment was excised, purified and concentrated as described in the Wizard[®] SV Gel and PCR Clean-up system (Promega[®]) and stored at -20 °C until cloning step. PCR products were cloned by using the pGEM-T vector system II (Promega[®]) according to the manufacturer's protocol. The ligation were transformed into JM109 High Efficiency Competent Cells and plated on Luria-Bertani agar with ampicillin to a final concentration of 100 µg/mL. The LB plates were supplemented with 100 µl of 100mM IPTG (isopropyl-β-D-thiogalactopyranoside) and 20 µl of 50 mg/mL X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) according to manufacturer's instructions (Promega[®]). White colonies were grown overnight in 2 mL Luria-Bertani broth containing ampicillin [50mg/mL] and incubated at 37 °C in a rotatory shaker at 120 rpm. Plasmids were purified using a Wizard[®] Plus SV Minipreps DNA purification system (Promega[®]). Primers used for amplification of 16S rDNA inserted sequences were SP6 (5'-ACGATTTAGGTGACACTATAG-3') and T7 (5'-TAATACGACTCACTATAGGG-3') corresponding to vector DNA flanking cloned insert (Promega[®]). PCR was conducted as described above except that the final step consisted of incubation for 7 minutes at 72 °C. The PCR products were visualized by agarose gel electrophoresis, and the products obtained were purified by using Wizard[®] SV Gel and PCR Clean-Up System. Each clone was designated EBS, EBP, EMS or EMP, followed by the clone number. The single-strand nucleotide sequencing was performed using the T7 promoter primer at the Nevada Genomics Center, University of Nevada-Reno on an ABI Prism 3730 DNA analyzer[™]. A total of 100 ng/ µl of each PCR product were used to prepare the samples; dilutions were done as described in the sequencing preparation protocol of Nevada Genomics Center [www.ag.unr.edu/genomics]. The Nevada Genomics Center used Phred, a base call quality assessment software, in conjunction with the ABI Prism

3730 DNA analyzer. The Phred software reads DNA sequencer trace data and assigned a quality score to each base in a sequence (Ewing et al., 1998). In this study, bases that were given a Phred score below 20 were eliminated in all analyses.

Each sequence was submitted to the CHECK_CHIMERA program version 1.8 of the Ribosomal database Project (RPD) to detect the presence of possible chimeric artifacts. A preliminary identity was done by compared a total of 239 partial clone sequences with sequences obtained from the BLAST, National Center for Biotechnology Information [<http://www.ncbi.nih.gov/BLAST/>].

In order to conduct phylogenetic and statistical analyses of the 16S clone libraries, the sequence data set from each library were divided in those sequences inserted to the plasmid in a 5' end orientation and those inserted in a 3' end orientation.

Phylogenetic Analyses of clones libraries

Seventeen, thirty-one, twenty-five and seventeen representative clones (OTUs) inserted to the plasmid in a 5' end orientation from EBS, EBP, EMS and EMP libraries, respectively, underwent phylogenetic analysis. Likewise twenty, eighteen, sixteen and fifteen representative clones (OTUs) from EBS, EBP, EMS and EMP libraries, respectively, from the 3' end orientation sequences data set also underwent phylogenetic analysis. The sequences from each library in each end-orientation-data-set were individually aligned with sequences obtained from the Greengenes Database (DeSantis et al., 2006), using ClustalW function of the BioEdit software version 7.0.5 [<http://greengenes.lbl.gov>]. A phylogenetic tree was constructed for each end-orientation-sequences-data-set. Phylogenetic trees were constructed with the Molecular Evolutionary Genetics Analysis package (MEGA version 4.0), (Tamura et al., 2007). The

confidence level of the phylogenetic tree topology was evaluated by performing 2000 bootstrap replications with the Neighbor-Joining method. *Aquifex aeolicus* served as the outgroup.

Diversity Indices and Statistical Analyses of the 16S Clone Libraries:

The 16S rDNA sequences recovered from each libraries in each end-orientation- data-set were uploaded to the ClustalW program [<http://www.ebi.ac.uk/clustalw/>] of the European Bioinformatics Institute (EMBL-EBI) for alignment. A distance matrix was constructed from each alignment using the DNADIST program, available in the phylogeny inference package PHYLIP (Version 3.65; J.Felsenstein, Department of Genetics, University of Washington [<http://evolution.genetics.washington.edu/phylip.html>]). The distance matrix was used as an input file to DOTUR (Distance based OTU and Richness determination software [www.plantpath.wisc.edu/fac/joh/dotur.html]), which placed sequences into OTUs. An OTU is defined as containing sequences with greater than 97% similarity. DOTUR was also used to calculated diversity indices: Shannon-Weiner index (H), Simpson's reciprocal index (1/D), richness (S), evenness (H), Jackard index, Schao, SAce. ASLO (Limnology and Oceanography: Methods [<http://www.aslo.org/lomethods/free/2004/01114.pdf>]) was used to construct rarefaction and coverages curves from the values previously calculated by DOTUR.

The 16S rDNA libraries from each end-orientation-sequences-data-set were compared using the Web-LIBSHUFF (an abbreviation of LIBrary SHUFFling) computer program (version 1.22; D.Singleton, Department of Microbiology, University of Georgia [<http://libshuff.mib.uga.edu/>]) to determine if they are significantly different from one another. The LIBSHUFF program analyzes two libraries at a time. In order to compare the level of differences among the four libraries in this study five individual analysis were necessary: EBS

vs. EMS, EBS vs. EBP, EMS vs. EMP, EBP vs. EMP and EB vs. EM. For each analysis, a distance matrix was generated using DNADIST program. Web-LIBSHUFF analyzed the distance matrix generated by DNADIST program by comparing coverage curves as described by Cramér-von Mises test statistic,

$$\Delta C_{xy} = \sum |C_x(D) - C_{xy}(D)|^2,$$

where D increase in increment of 0.01, (Kemp et al. 2004; Singleton et al, 2001).

Terminal Restriction Fragment Length Polymorphism (t-RFLP) of the DNA Soil Samples

Total genomic DNA was extracted from 1 g aliquots of soil samples, as described previously. The labeled-forward 27F' and the reverse 1392R' primers were used to amplify approximately 1,500 bp fragment of the 16S rDNA. Amplification was performed in a 50 µl (total volume) reaction mixture containing approximately 100 ng of the soil community DNA, 10 µl of 10X Go Taq DNA polymerase reaction buffer B, 1.7 µl of 25mM MgCl₂, 4.0 µl of dNTP's mix [2.5 mM (1:1:1:1 proportion)], 1.4 µl of [10mg/ml] BSA, 1.0 µl of each 16S rDNA primer [50 pmol/ µl], 1.0 µl of 1.25U Go Taq DNA polymerase (Promega®) and ddH₂O. The PCR was performed with a Perkin Elmer Gene Amp PCR System 2400. Reaction conditions were the same as those described above.

The quality of amplified 16S rDNAs was examined in 1% [wt/vol] agarose gel electrophoresis in 1X Tris-Acetate-EDTA and ethidium bromide staining. The PCR products were purified using the Wizard SV Gel and PCR clean-up system (Promega®), according to manufacturer's protocol. Purified PCR products (approximately 200 ng) were digested separately with 5 U of tetrameric restriction endonucleases *HaeIII* [GG'CC] [where the prime shows the site of cleavage], *MspI* [C'CGG], and *RsaI* [GT'CA] (Promega®) in a 20 µl total volume according to manufacturer's recommended reaction buffers. Restriction digests were

incubated at 37 °C for 4 h, after the reaction was completed, the restriction endonucleases were deactivated by heating the reaction mixture to 65°C for 10 min.

Following reaction digestion, 2.5 µl of each digest was mixed with 2.5 µl of IR2 stop solution (LI-COR Biosciences®). The mixture was denatured at 94°C for 5 min and cooled on ice, aliquots (1.0 µl) were loaded onto a 0.25 mm thick and 25-cm- long 5.5% acrylamide gel (acrylamide gel matrix KB^{Plus}-LICOR) using the LI-COR Biosciences NEN®DNA Analyzer Model 4300. Electrophoresis was performed for 3 h with the following parameters: voltage 1,500 (V), current 40 (mA) and power 40 (W).

Data Analysis of TRFLP

T-RFLP profiles were generated using the Gel Pro Analyzer software, version 4.5. In order to avoid TRFs caused by primer dimmers, we only considered peaks at positions between 50 and 700 bp. The relative abundance of a TRF in a T-RFLP profile was calculated by dividing the peak height of the TRF by the total peak height of all TRFs in the profile. All peaks with heights that represented less than 0.5% of the total peak height were excluded in further analysis. This step was performed to minimize the effect of variations in the T-RFLP profiles caused by the quantity of DNA analyzed. The data obtained from the T-RFLP profiles were exported to Excel, Microsoft software and converted into a relative abundance table (percentage of individual peaks in each profile). T-RFLP profiles were aligned by inspection of the relative abundance table and by manually grouping of the peaks in to categories. The peaks were aligned by the size of peaks in base pairs. The overall T-RFLP profiles for each DNA sample soil consisted of the individual T-RFLP profile generated by each restriction endonuclease. Statistical differences and cluster analysis for the four DNA soil samples T-RFLP profiles were

performed using Past (Paleontological statistics package) software program, version 2.07,
[<http://www.nhm.uio.no/norlex/past>] (Hammer et al., 2001).

CHAPTER IV

RESULTS

Physicochemical analysis of soil samples

The physicochemical results of soil samples are shown in Table 1. Two soil samples were obtained from each location corresponding to different depths. EBS (0-5 cm) and EBP (5-10 cm) were collected from the secondary forest. Likewise, EMS (0-5 cm) and EMP (5-10 cm) were obtained from the non-vegetated site. The observed phosphorus, potassium and calcium ions' concentrations found in EMS and EMP soil samples were lower than the corresponding concentrations within the secondary forest soil samples. However, the relative order of the magnesium ion concentration in the sites' samples was found to be EBS > EMS > EMP > EBP. The organic matter, organic carbon, and total nitrogen content found in the forested soil samples (EBS and EBP) were higher than the corresponding content found in deforested soil samples (EMS and EMP). As shown in Table 1, the range in pH was 4.05 to 4.69. In general a higher acidity was observed in deforested soil samples (EMS and EMP) with lower organic content (<0.30%).

Table 1. Soil samples chemical and physical parameters at two locations in Bosque del Pueblo, Adjuntas, PR.

Soil samples	pH ¹	pH ²	Organic C ³	Organic M ³	Total N ⁴	P ⁵	Concentration (ppm) of:			
	(H ₂ O)	(CaCl ₂)					(%)	(%)	(%)	K ⁶
EBS	4.69	3.92	3.44	5.93	0.31	12	306	306	80	3.00
EBP	4.62	3.93	2.04	3.52	0.22	10	18	18	17	0.83
EMS	4.05	3.76	ND	ND	0.030	.46	10	10	50	0.79
EMP	4.12	3.91	0.30	0.52	0.030	1	7	7	35	0.68

Physicochemical analysis of water samples

Temperature, conductivity, and pH values from *Estación Bosque* and *Estación Mina* water samples are shown in Figure 4, (individual measurements are given in Appendix II). The average water temperature showed a steady increase from February to June. The highest value was measured in June. After June the average water temperature showed a steady decrease. It can be noticed that the rate of temperature increase between February and June is larger than the rate temperature decrease from June to November. This trend was more pronounced in *Estación Bosque* than in *Estación Mina*. The average water conductivity measured in *Estación Bosque* was 0.054 (mess/cm), whereas the corresponding average for *Estación Mina* was 0.402 (mess/cm). While acidity for the sampling period was higher for *Estación Mina* (pH 4.0) than for *Estación Bosque* (pH 5.3). One-way ANOVA demonstrated that the differences in the temperature, conductivity, and pH average values between the two sampling locations are statistically-significant, ($P < 0.0001 < 0.05$), ($P < 0.0001 < 0.05$), and ($P < 0.0001 < 0.05$), respectively.

Nickel, Copper and Lead concentrations in water samples

Averages values and their standard deviations of nickel and copper concentrations in water samples from *Estación Bosque* and *Estación Mina* are given in Table 2, (individual values were presented in Appendix II). Nickel concentration in *Estación Mina* was almost twice the observed to *Estación Bosque*. While copper concentration was three times higher as well. Lead was not detected at any site during the sampling period. Performance of one-way ANOVA analysis of the metal ions concentrations in water samples from *Estación Bosque* and *Estación Mina* showed statistically significant differences ($P < 0.0001 < 0.05$) in nickel concentrations, and

($P < 0.0001$ < 0.05) in copper concentrations of the water samples from two locations at *Bosque del Pueblo*.

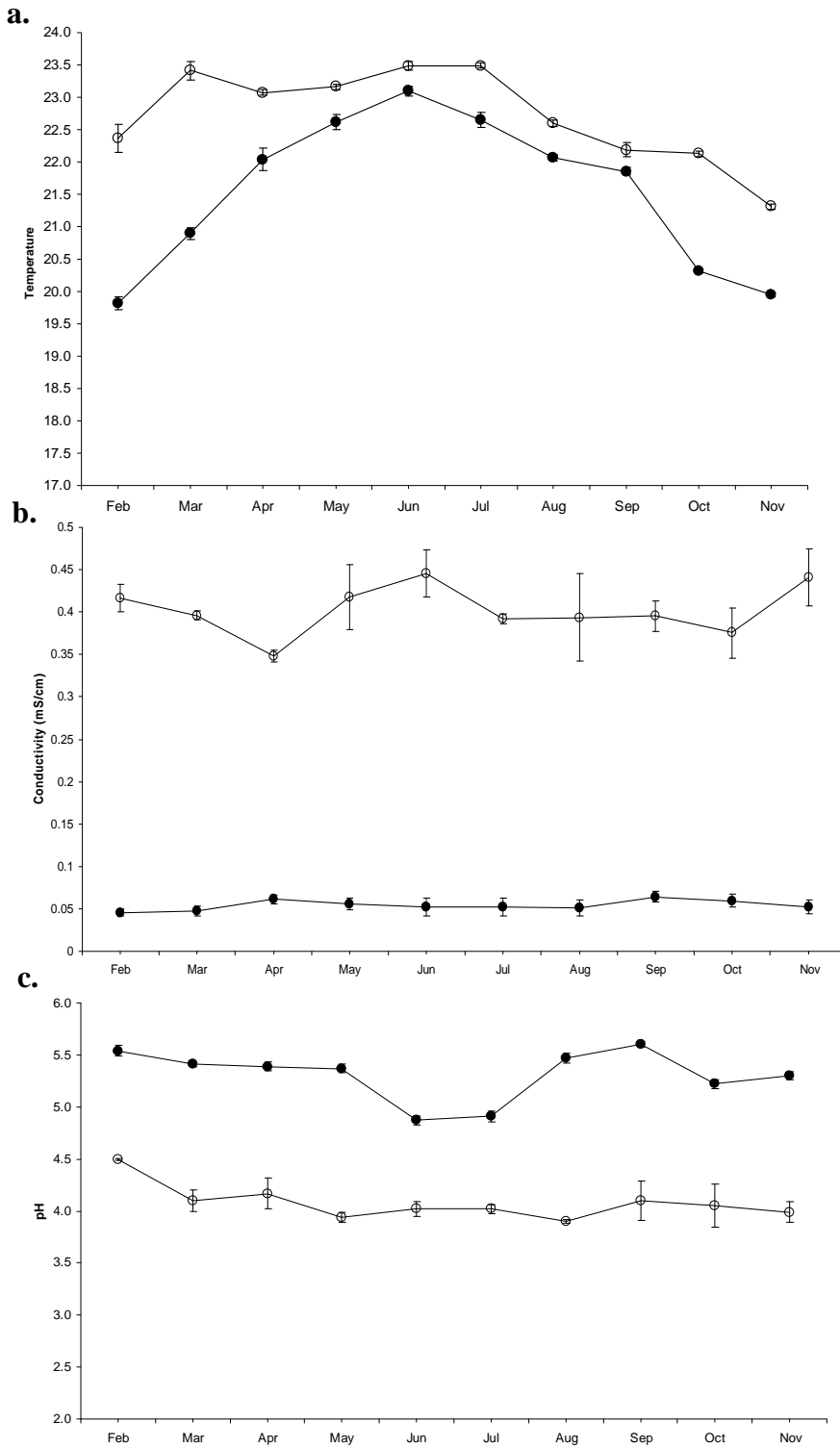


Figure 3. Temporal variation of (a) temperature, (b) conductivity, and (c) pH values of water samples obtained at *Bosque del Pueblo, Estación Bosque* (●) and at *Estación Mina* (○).

Table 2. Elemental analysis of nickel and copper content in water samples at Bosque del Pueblo, Adjuntas, PR

	Chemical element (ppm) ^a	
	Ni	Cu
Estación Bosque		
February	0.0138 (0.0005)	1.7875 (0.0152)
March	0.0199 (0.0005)	1.8619 (0.0027)
April	0.0167 (0.0005)	1.8422 (0.0154)
May	0.0187 (0.0004)	2.1424 (0.0251)
June	0.0167 (0.0004)	2.5975 (0.0224)
July	0.0105 (0.0008)	2.3082 (0.0125)
August	0.0121 (0.0008)	1.8684 (0.0037)
September	0.0113 (0.0003)	1.8166 (0.0040)
October	0.0096 (0.0010)	2.3131 (0.0143)
November	0.0106 (0.0004)	1.6691 (0.0043)
Estación Mina		
February	0.0218 (0.0005)	5.3122 (0.0134)
March	0.0259 (0.0006)	6.1733 (0.0503)
April	0.0225 (0.0007)	6.1800 (0.0371)
May	0.0251 (0.0001)	5.7867 (.0044)
June	0.0259 (0.0005)	6.3955 (0.1195)
July	0.0258 (0.0002)	6.6267 (0.1313)
August	0.0273 (0.0008)	7.9578 (0.0800)
September	0.0237 (0.0006)	6.9777 (0.0734)
October	0.0213 (0.0003)	5.7533 (0.0635)
November	0.0245 (0.0005)	5.3455 (0.0204)

^a Average Standard Deviation.

DNA extraction and PCR amplification from soil samples

DNA extracted from the four soil samples was predominantly of high molecular weight (Fig. 1) and was appropriate for direct PCR amplification. The amounts of DNA extracted from the four soil samples differed greatly. The prefixes EBS, EBP, EMS and EMP indicate that the 16S rRNA gene sequences were obtained from soil samples at *Estación Bosque Superficial* at 0 to 5 cm depth, *Estación Bosque Profundo* at 5 to 10 cm depth, *Estacion Mina Superficial* at 0 to 5 cm depth and *Estación Mina Profundo* at 5 to 10 cm depth, respectively. The highest yields of genomic DNA recovered were from the EBS and EBP soil samples and the estimated amount were approximately twice the yield obtained from EMS and EMP samples.

Out of 59 positive transformants recovered from the EBS library 57 clones tested positive for 16S rDNA genes (97%). The EBP, EMS and EMP libraries yielded 71 clones (96%), 61 clones (94%), and 50 clones (94%), respectively. None of the 16S rDNA sequences were identified as chimeric artifacts based on results from CHECK_CHIMERA version 8.1 program of the Ribosomal Database Project (RDP) and Greengenes program (DeSantis et al., 2006 [<http://greengenes.lbl.gov>]). Ten sequences with less than 450 base pairs in length were excluded from the analysis. The resulting 239 16S rDNA sequences belonged to the kingdom Bacteria, giving a specificity of 100% for the 8F (5'-AGAGTTTGATCMTGGCTCAG-3') and 1392R (5'-ACGGGCGGTGTGTACA-3') primers.

The 239 clone sequences had a length ranging from 450 to 824 base pairs. Eighty-six percent (205) of the clone sequences were about 650 to 749 base pairs in length. Only sequences containing bases with a given Phred Q value ≥ 20 were considered for statistical and phylogenetic analyses.

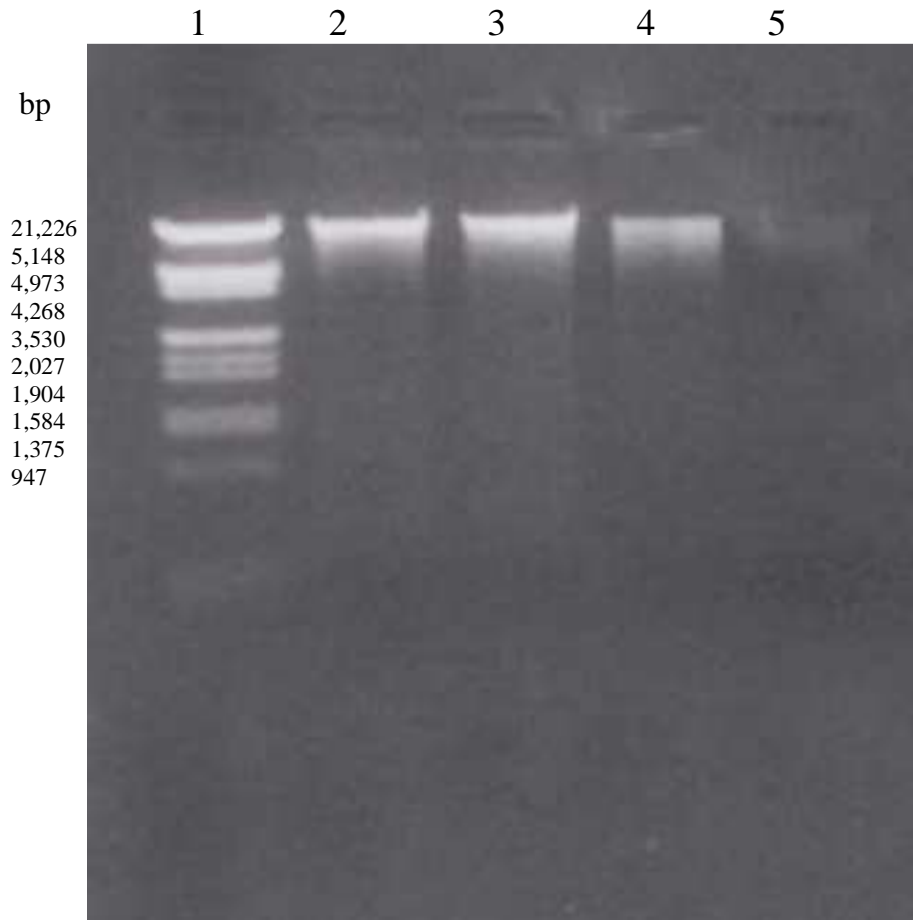


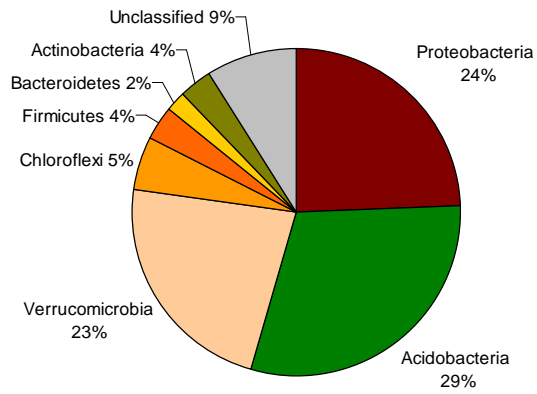
Figure 4. DNA isolated from soil samples. Lane 2, DNA sample from EBS; lane 3, from EBP; lane 4, from EMS; lane 5, from EMP. Lambda digested with *Hind* III and *Eco*RI as used as molecular marker, lane1.

Phylogenetic groups of Bacteria detected in soil samples

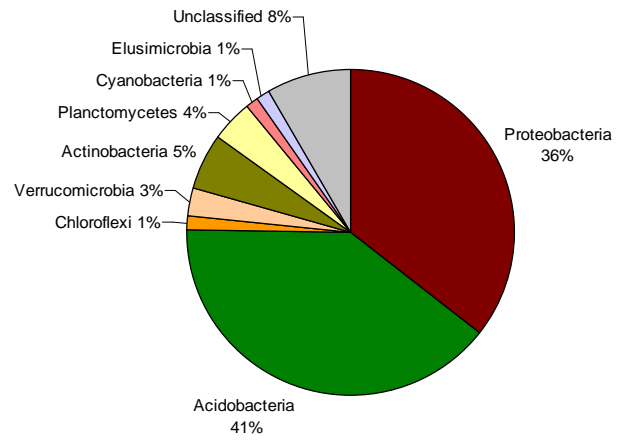
A total of 11 phyla were identified out of four clone libraries. Figure 5 describes the distribution of bacterial groups found in the EBS, EBP, EMS, and EMP clone libraries. A large proportion of clones from EBS clone library were identified as *Acidobacteria* (29%), *Proteobacteria* (24%), and *Verrucomicrobia* (23%); respectively. Smaller proportions of

Chloroflexi (5%), *Firmicutes* (4%), *Actinobacteria* (4%), and *Bacteroidetes* (2%) were identified in EBS clone library. The collection of sequences from EBS clone library did not contain *Planctomycetes*, *Verrucomicrobia*, *Elusimicrobia*, *Cyanobacteria* or *Nitrospira* members. EBP clone library also contained a high proportion of members of *Acidobacteria* (41%) and *Proteobacteria* (36%). Five percent of the sequences from EBP clone library were identified as *Actinobacteria* members. In addition, members of *Elusimicrobia* were found only in the EBP library, which accounted for 1% of the total sequences of that library. In contrast to the EBS and EBP clone libraries, the collection of sequences from EMS and EMP libraries were dominated by *Chloroflexi* members with 32% and 48%, respectively. However, a large proportion of EMS sequences belonged to *Acidobacteria* (29%) and *Proteobacteria* (13%) members. The smallest population of *Proteobacteria* members was found in EMS. Another bacterial group found in a relatively small quantity in EMS library was *Firmicutes* (5%). Members of phylum *Nitrospirae* were found only in EMS clone library. *Cyanobacteria* members were identified only in EMS (5%) and EBP (1%) clone libraries. Furthermore, members of *Acidobacteria* and *Proteobacteria* phyla were common in the EMP clone library; with 15% and 23%, respectively. A small proportion of sequences from EMP clone library were identified as members of *Firmicutes* (6%), *Verrucomicrobia* (2%) and *Planctomycetes* (2%) members. Between 4- 9% of bacterial sequences were not classified.

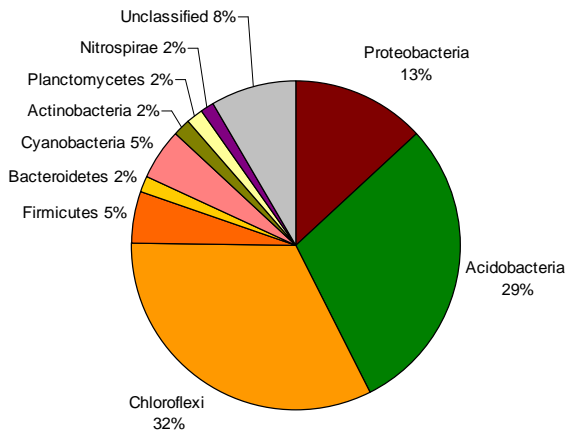
a.



b.



c.



d.

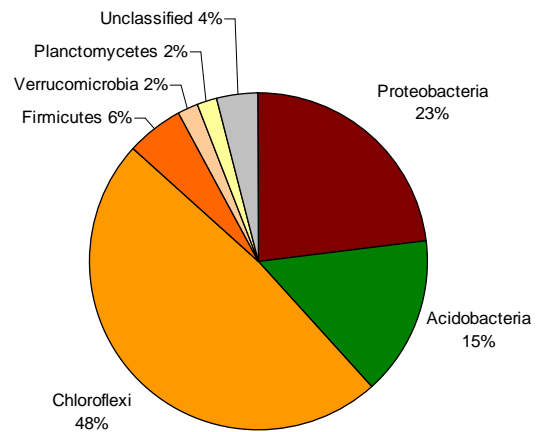


Figure 5. Distribution of phylogenetic groups of *Bacteria* detected in soil samples by 16S rDNA clone libraries. (a) EBS, (b) EBP, (c) EMS and (d) EMP.

In order to conduct richness estimation, the clones' sequences from each clone library were placed into OTUs. The criterion used to define OTUs was a 97% sequence similarity. Nine seven OTUs were identified in the 134 16S rDNA sequences recovered from the 5' end-portion-data-set. Likewise, 71 OTUs were identified in 105 sequences from the 3' end-portion-data-set. The sequence that best matches each OTU was identified by a comparative analysis with sequences deposited in the NCBI database. Each representative clone sequence (OTU) and its closest match according to NCBI databases are given in Appendix III.

Ninety seven OTUs from the 5' end portion-data-set were generated using DOTUR program. None of the generated OTUs was common to the four clone libraries investigated. However an OTU represented by EBP124 was present in EBS, EBP and EMP clone libraries. Four OTUs represented by EBP114, EBP406, EBP404 and EBP417 were presented in EBS and EBP clone libraries. While two OTUs represented by EMP55 and EMS 414 were presented in EMS and EMP soil samples.

Seventy-one OTUs were generated from the 3' end portion-data-set using DOTUR program. As occurred in the 5' end portion-data-set a common OTU in all soil samples was not found. Three OTUs represented by EBP409, EBP110 and EBP167 were found in EBS and EBP soil samples. Two OTUs, EBP 115 and EBP146 were present in EBP ad EMS soil samples. Only one OTU, represented by EBP416 was found in EBP and EMP soil samples. The most dominants OTUs in EBS community were EBS115 and EBS111, each OTU composed of two sequences. Three sequences represented by EBP102 were the most dominant OTU in the EBP community. In the EMS community the most abundant OTU, represented by EMS413, was composed of five sequences. Two OTUs represented by EMP445 and EMP203 were the most

abundant OTUs in the EMP community, each composed of three sequences. In the 3' end portion-data-set forty-seven of the sequences were present in single-member OTUs.

Diversity Statistics of the clone libraries

DOTUR was used to analyze all clone libraries constructed from *Bosque del Pueblo* soil samples. DOTUR analysis requires that the sequences-data-set from each library are aligned by using ClustalW software [<ftp://ftp.ebi.ac.uk/pub/software/unix/clustalw/>], and that a distance matrix is constructed by using the DNADIST program from PHYLIP computer software. The generated distance matrix was used as an input file to DOTUR which rapidly and systematically assigns sequences to OTUs. In this study, OTUs were defined by using a distance level of 3%. The richness estimators and diversity indexes calculated from each library sequence-data-set using DOTUR are presented in Table 3. Data in Table 3 demonstrates that when clone libraries belonging to the 5' end-oriented-data-set were compared in terms of diversity, both Shannon and Simpson's indexes revealed that EBP 5' community had a higher diversity than EBS, EMS, and EMP communities. While in the 3' end-portion-data-set both Shannon and Simpson's diversity indexes revealed that EBS community had a higher diversity than EBP, EMS and EMP communities. According to Shannon index EMP community had the lowest diversity in both 5' and 3' end-oriented-data-set. The 16S rDNA genes richness was calculated for each community using nonparametric estimators ACE, ChaO and Jaccard. According to Jaccard and Chao estimators, richness in clones libraries from the 5' end-oriented-data-set decreases as follows: EBP > EBS > EMS > EMP. While the richness in clone libraries from the 3' end-oriented-data set decreases as follows: EBS > EMP > EMS > EBP as revealed by Jaccard and Chao estimators.

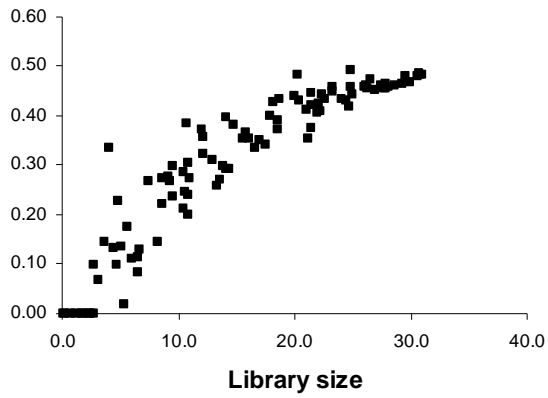
Table 3. Diversity and richness estimates from 16S rDNA clone libraries.

<i>Clone Library</i>	<i>Diversity estimate</i>		<i>Richness estimate</i>				
	<i>Shannon (H)</i>	<i>Simpson's (D)</i>	<i>Observed Richness^a</i>	<i>Jaccard</i>	<i>SAce</i>	<i>Schao</i>	<i>Coverage (%)</i>
EBS 5'	2.88	27.35	21 (31)	58.37	62.8	61	48
EBS 3'	3.02	64.99	22 (26)	88.85	99.8	79	27
EBP 5'	3.36	77.99	31 (40)	58.98	72.94	63	43
EBP 3'	2.71	23.25	18 (31)	25.00	22.91	22	74
EMS 5'	3.16	50.99	26 (34)	53.26	74.71	58	41
EMS 3'	2.60	18.47	16 (27)	26.71	31.37	27	63
EMP 5'	2.78	23.88	19 (29)	48.02	51.30	49	52
EMP 3'	2.49	19.09	15 (21)	28.42	34.36	29	52

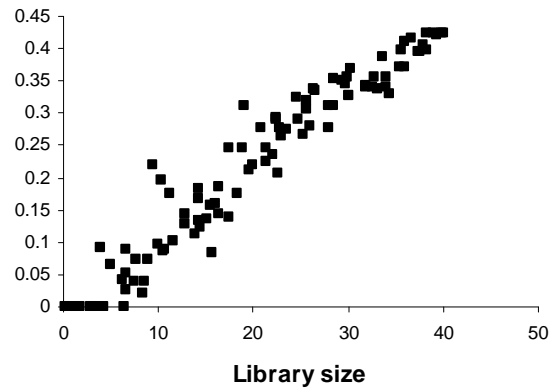
^a The values in parentheses are the number of 16S rDNA clones. Richness is the number of distinct sequences (OTUs). An OTU is defined as a group of sequences that share more than 97% nucleotic similarity.

The Good Coverage Index, calculated using the program ASLO [www.aslo.org/methods/free/2004/0114a.html], was used to evaluate whether a library has captured a sufficient large fraction of diversity in the sampling environment (Kemp and Aller, 2004). Figure 6 shows the good coverage curves for the 5' end terminal data-set of EMP, EBS, EBP, and EMS 16S rDNA clone libraries. The coverage estimated using ASLO for the above clone libraries was 52%, 48%, 43%, and 41%, respectively. In the case of the 3' end-oriented-data-set of the clone libraries EBP, EMS, EMP and EBS the coverage was 74%, 63%, 52% and 27%, respectively, Figure 7. The Good coverage curves showed that EBS, EBP, EMS, and EMP libraries were not sufficiently large to be representative of their respective source environments.

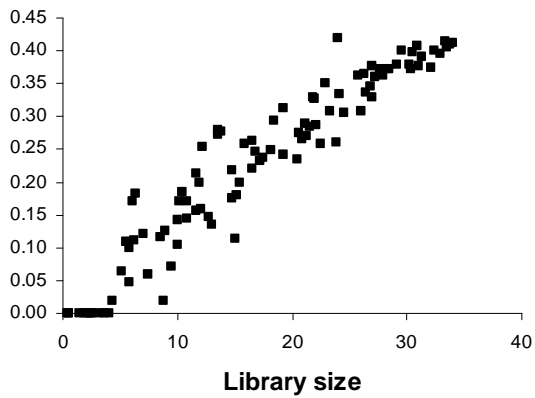
a.



b.



c.



d.

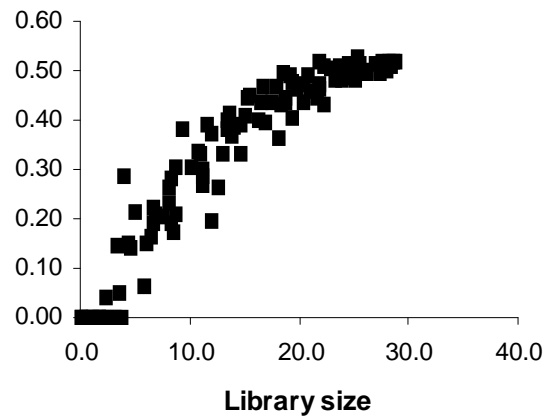
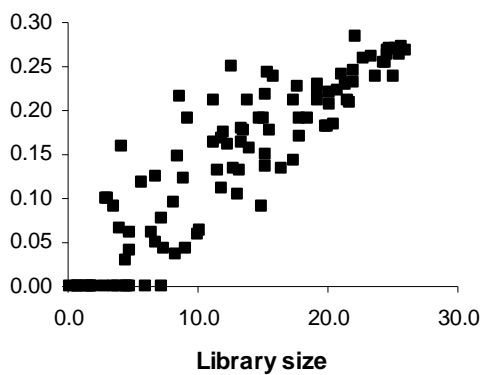
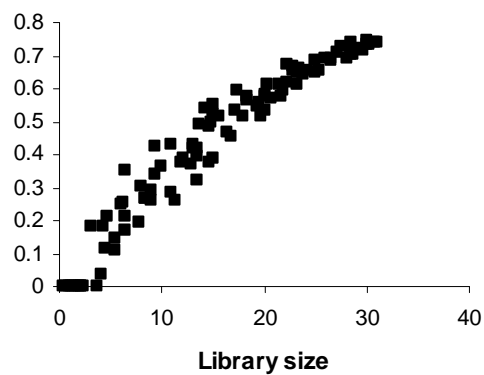


Figure 6. Plots of Good Coverage index vs. Library size (Good Coverage Curve) for the 5' end terminal data-set of the 16S rDNA clone libraries: (a) EBS, (b) EBP, (c) EMS and (d) EMP.

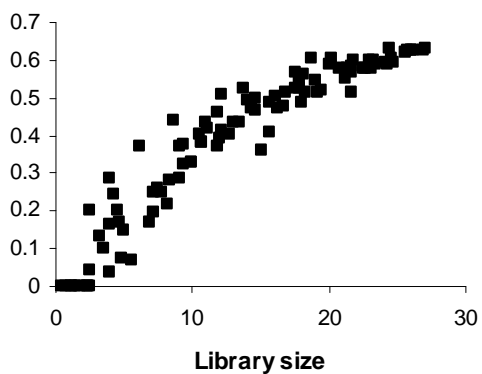
a.



b.



c.



d.

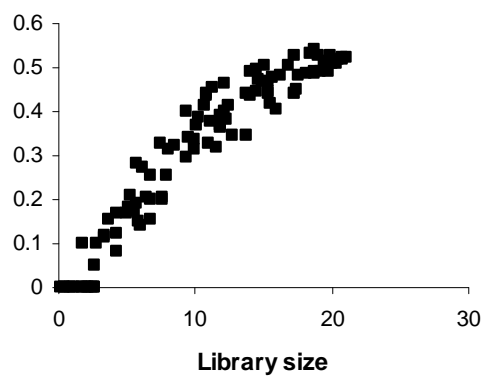
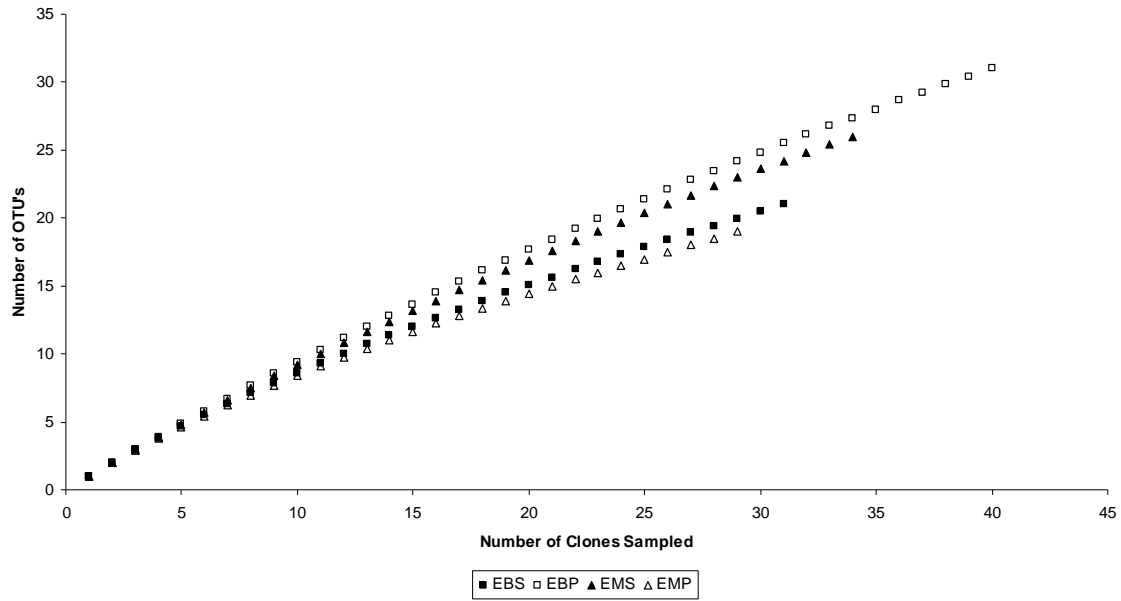


Figure 7. Plots of Good Coverage index vs. Library size (Good Coverage Curve) for the 3' end terminal data-set of the 16S rDNA clone libraries: (a) EBS, (b) EBP, (c) EMS and (d) EMP.

Plots of OTUs number and library sampling size (rarefaction curves) were constructed for each sequence-data-set using DOTUR [<http://www.plantpath.wisc.edu/fac/joh/dotur.html>] to assess the relative richness among libraries with different sampling sizes (Schloss and Handelsman, 2005). The criterion used to estimate the actual bacterial diversity within a sample is the OTUs number independency of library sampling size. This information can be deduced where the plots in Figure 8 levels off (Kemp and Aller, 2004). The plots of OTUs number vs. library sampling size in Figure 8 are concave (downward) indicating a tendency to level off at some library sampling size. The region in the plots where OTUs number is library size-independent were not reached experimentally. However, that region can be estimated mathematically using graph analysis and interpretation.

a



b

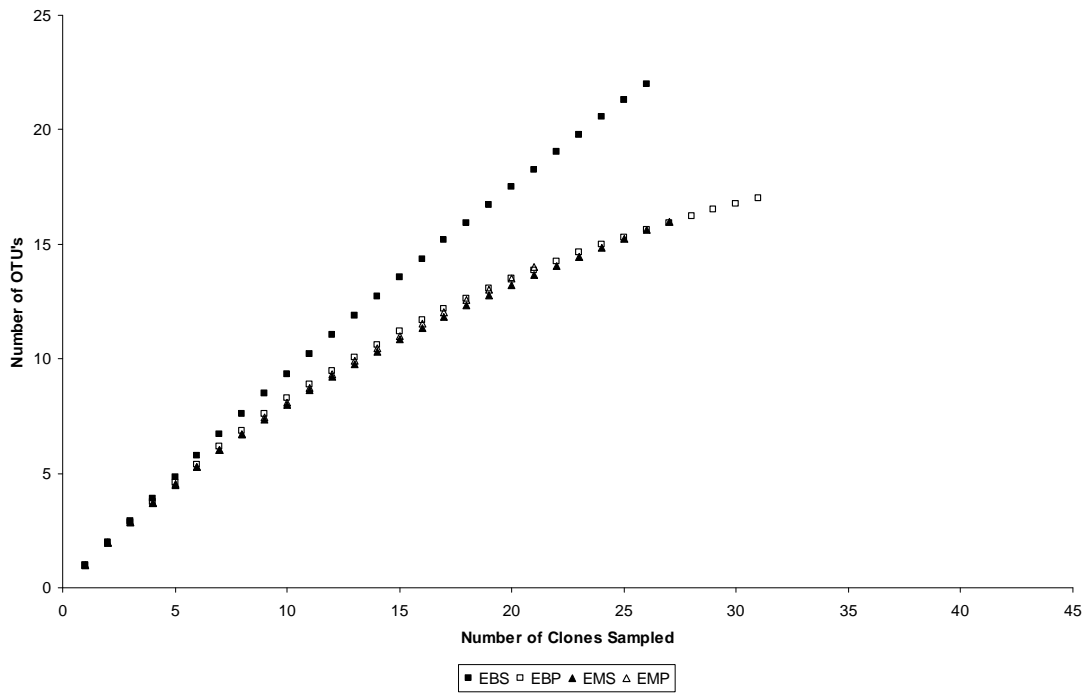


Figure 8. Plots of OTUs number vs. Library size sampling (Rarefaction curves) for the 5' end (a) and the 3' end (b) data sets of the 16S rDNA clone libraries, EBS ■, EBP □, EMS ▲ and EMP Δ.

Whether or not two clone libraries were significantly different in term of their bacterial community composition was elucidated by using LIBSHUFF software. The input file, a distance matrix for each of the two clone libraries, which is needed in the analysis using LIBSHUFF is generated by DNADIST, a program within PHYLIP computer software.

One of the LIBSHUFF out puts is a p-value. This p-value is used to assess whether clone sequences from two libraries are different in term of their bacterial composition. The criterion used is that p-values > 0.05 indicate that two libraries may be derived from the same source. All p-values ($= 0.001$) were smaller than 0.05, therefore, results from LIBSHUFF analysis suggest the existence of statistically-significant differences among the clone libraries in this study. Thus, one may conclude that they may come from distinct sources.

Another output generated by LIBSHUFF software is a set of two coverage curves. Coverage curves generated by LIBSHUFF were also useful tools to determine the extent of differences between two given libraries. The homologous coverage curve ($C_x(D)$), was determined by comparing each sequence in the library with the other sequences within the same library. A second analysis was done where each sequence in given library was individually compared to all sequences in another library; the resulting coverage values generated the heterologous coverage curve, ($C_{xy}(D)$). Two libraries are not significantly different if the homologous and heterologous curves appear similar. When two libraries are significantly different, the homologous and heterologous curve will diverge at some points. In all analysis made for libraries under this study, the homologous and heterologous coverage curves revealed that the libraries were significantly different from each other. The homologous and heterologous coverage curves for each analysis are shown in Appendix IV.

Phylogenetic Analyses of clone libraries

The microbial community composition of the four clone libraries was investigated by comparative sequences analysis of partial 16S rDNA genes. Phylogenetic analysis, using the MEGA3.1 computer package, was performed on the 5' portion–data-set corresponding to EBS, EBP, EMS and EMP libraries. Specifically, seventeen, thirty-one, twenty-five and seventeen OTUs from EBS, EBP, EMS and EMP libraries, respectively were analyzed. The same analysis was performed on the 3' end portion-data-set of the clone libraries above. In this case, eighteen, twenty-one, sixteen and sixteen OTUs from EBS, EBP, EMS and EMP clone libraries, respectively were analyzed.

As shown in the 5' end phylogenetic tree in Figure 9, twenty-one out of ninety-two OTUs from EBS, EBP, EMS and EMP clone libraries were affiliated to phylum *Acidobacteria*, Figure 9a. The OTU represented by EBS 94 were closely-related to *Korebacter versatilis* str-Ellin 345. Two OTUs, EBP9 and EMP207 formed a cluster with *Edaphobacter modestus* str-Jb-1 and *Acidobacteria capsulatum*. Most OTUs of the 5' end tree were more closely-related to uncultured environmental clone sequences obtained from different environmental soil samples deposited in NCBI-data base. An OTU represented by EMS414 was affiliated with *Acidobacteria* and shared a high sequence similarity (95%) with a clone sequence from soil iron-manganese nodules (He et al., 2008). Whereas, two OTUs represented by EBS87 and EBP114 were more closely-related to uncultured bacterium clones from limestone soil aggregates (Hansel et al., 2008). OTUs EMS415, EMS146 and EMS112 had a sequence similarity between 96 to 98 % to uncultured bacterium clones from soil rhizospheres (Lesaulnier et al., 2008). Four OTUs were more closely-related to uncultured clones from tropical forest soil samples, two of the four OTUs represented by EBP140, EBP139 were related to clones from Brazilian soil samples (Kim

et al., 2007). The OTU represented by EMP207 was related to a clone sequence from a Taiwan tropical forest soil (Tseng et al., GenBank Accession no. DQ984559.1), and the one represented by EMP55 was related to a clone sequence from China soil samples (Chan et al., 2006). Interestingly, the OTU represented by EBP162 was related to an uncultured clone from a subalpine forest soil in Colorado USA (Jones et al., GenBank Accession no. EU150198.1).

Nineteen of ninety-two OTUs from the 5' portion-data-set were affiliated to the phylum *Proteobacteria*, and were further clustered into α -, β -, δ - and γ - *Proteobacteria*. As shown in Figure 9b. *Alphaproteobacteria* was the most abundant *Proteobacteria* in the EBS, EBP and EMS clone libraries in this study. An OTU represented by EBP107 showed a sequence similarity (99%) to *Bradyrhizobium sp.* OTUs represented by EBP150, EBS106, EBP169 and EMP438 were also similar to common soil bacteria, such as nitrifying bacteria and Rhizobiales. Among the OTUs, the sequences EMP208 and EMS134, classified as γ -*Proteobacteria* were related to clones from Kalahari Shield in South Africa surface water (Gihring et al., 2006) and soil rhizosphere from Wisconsin (Lesaulnier et al., 2008). An OTU represented by EBP 126, and classified as a *Betaproteobacteria*, formed a cluster with a sequence similarity of 100% with *Burkholderia nodosa* str-Br3461, which was isolated from nitrogen-fixing nodules on the roots of woody legumes native to Brazil (Chen et al., 2008). Four OTUs, represented by EBP414, EBP104, EBS74 and EBP163 were affiliated to uncultured members of δ -*proteobacteria* found in various soil environments such as: forest soil in Central Taiwan (Lin et al., GenBank Accession no. GU016086.1), tall grass prairie soil in central Oklahoma (Elshahed et al., 2008), soil aggregates from Oak Ridge, TN (Hansel et al., 2008), and rice paddy field soil (Kuesel et al., GenBank Accession no. AM941477.1). Indeed, OTUs represented by EBP 414 and EBP163 were more related to cultured strains *Chondromyces crocatus* published by

Shimkets and Woese in 1992 and *Stigmatella aurantica* str-Sg a1 published by Spröer in 1999, respectively. Both strains have been reported in sample soils from tropical rain forests worldwide.

OTUs affiliated with the phylum *Chloroflexi* constituted 18 out of the 92 OTUs in the 5' portion-data-set, (20%). Most of the OTUs classified as members of the phylum *Chloroflexi* were retrieved from soil samples at the deforested field (EMS and EMP) in *Bosque del Pueblo*. As revealed in Figure 9c, many of the OTUs affiliated to *Chloroflexi* were clustered between them rather than with clones sequences previously detected in other environmental samples, as is the case of the OTUs represented by EMS449 and EBS79, EMS85 and EMP442, EMS111 and EMS115, and, EMP415 and EMP7. Based on the 5' end phylogenetic tree OTUs represented by EMS12 and EMP413 clustered with clone sequences retrieved from volcanic deposits from Hawaii, USA (Gomez-Alvarez and Nusslein, 2007). While OTUs represented by EMP434 and EMP421 shared a sequence similarity between 95 to 96% with clone sequences found in soil aggregates from Oak Ridge, TN (Hansel et al, 2008). *Ktedobacter racemifer* strain SOSP1-21 was the cultured strain most closely-related to members of *Chloroflexi* that were found in clone libraries at *Bosque del Pueblo*.

Seven OTUs represented by clone sequences EBS85, EBS120, EBS107, EBP109, EBS102, EBS108 and EBP417, affiliated with the phylum *Verrucomicrobia*, and shared a sequence similarity between 91 to 97% with soil clones from various environments, Figure 9d. Based on results from the phylogenetic analysis performed in this study *Chthoniobacter flavus* str- Ellin428 was the only cultured strain that was affiliated with the verrucomicrobial clone sequences from *Bosque del Pueblo*.

Five OTUs were affiliated to members of phylum *Actinobacteria* as shown in Figure 9d. OTUs represented by EBP143, EBP145, EBS116 and EBS175 were more closely-related between them than with uncultured clones from databases; the most closely related isolated strains were *Mycobacterium bouchedurhonense* str 4355387 and *Actinomadura* sp. str Shinshu-MS-03. An OTU represented by EBP147 shared a sequence similarity of 97% with *Mycobacterium bouchedurhonense* str 4355387, both sequences formed a cluster supported by a bootstrap value of 95%.

The clone sequences from *Bosque del Pueblo* that belong to the phylum *Firmicutes* formed two clusters as revealed in Figure 9d. All OTUs classified within this Phylum were related to bacterial clones or bacterial strains detected in soils and sediments environments. OTUs represented by EMS13, EMS107, EMP435 and EMP427 were phylogenetically related to uncultured clones from various soil environments: forest soil from central Taiwan (Lin et al., GenBank Accession no. GU016168.1), Eucalyptus plantation soil from China (Duan et al. GenBank Accession no. GU172181.1), inland dune fields from Georgia, USA (Tarlera et al., 2008), and rhizosphere soil of trembling aspen (*P. tremuloides*) in Rhinelander, WI, USA (Lesaulnier et al., 2008). While OTUs represented by EBS100 and EMS 108 were more related to *Paenibacillus terrigena*.

The remaining phyla present within the clone libraries at *Bosque del Pueblo* included *Cyanobacteria*, *Bacteroidetes*, *Elusimicrobia*, *Planctomycetes* and *Nitrospira*. Each consisted of less than 3.2% of the OTUs of the 5' end-data-set, Figure 9d. In the case of the OTUs affiliated to phylum *Cyanobacteria*, one OTU represented by EMS110 formed a cluster with *Cylindrospermum* sp. str-A1345 with a sequence similarity between them of 92%; another OTU represented by EMS141 was more closely-related to an uncultured clone detected in a soil

sample from Socompa Volcano at Los Andes (Costello et al., 2009). Only one OTU, represented by EMS405, were classified as a member of phylum *Bacteroidetes*. EMS405 shared a 93% sequence similarity with an uncultured clone retrieved from soil samples at Coweeta Forest, North Carolina, USA (Upchurch et al., 2008). An OTU retrieved from the EBP clone library at *Bosque del Pueblo* and represented by EBP120, shared a 94% sequence similarity with an uncultured bacterium clone detected in inland dune fields at Georgia, USA (Tarlera et al., 2008). Based on results from the phylogenetic analysis the closest strain related to EBP120 was *Elusimicrobium minutum* str-Pei191; first cultivated by Herlemann (Herlemann et al., 2009). Among the OTUs from *Bosque del Pueblo* affiliated to phylum *Planctomycetes* were representative clone sequences EBP119 and EMS145. Both OTUs were more closely-related to uncultured clones detected in soil samples from a karst region in China (Chen et al., GenBank Accession no. EU881251.1) and from soil rhizosphere of trembling aspen (*P. tremuloides*) in Rhinelander, WI, USA (Lesaulnier et al., 2008), respectively. The representative clone EMS80 formed a cluster supported by a bootstrap value of 99 with an uncultured bacterium from a Western Amazon forest soil sample (Kim et al., 2007). Based on results from the phylogenetic analysis, the closest isolated strain to EMS80 with a 95% sequence similarity was *Nitrospira sp.* srtGC86, a nitrite-oxidizer (Burrell et al., 1998).

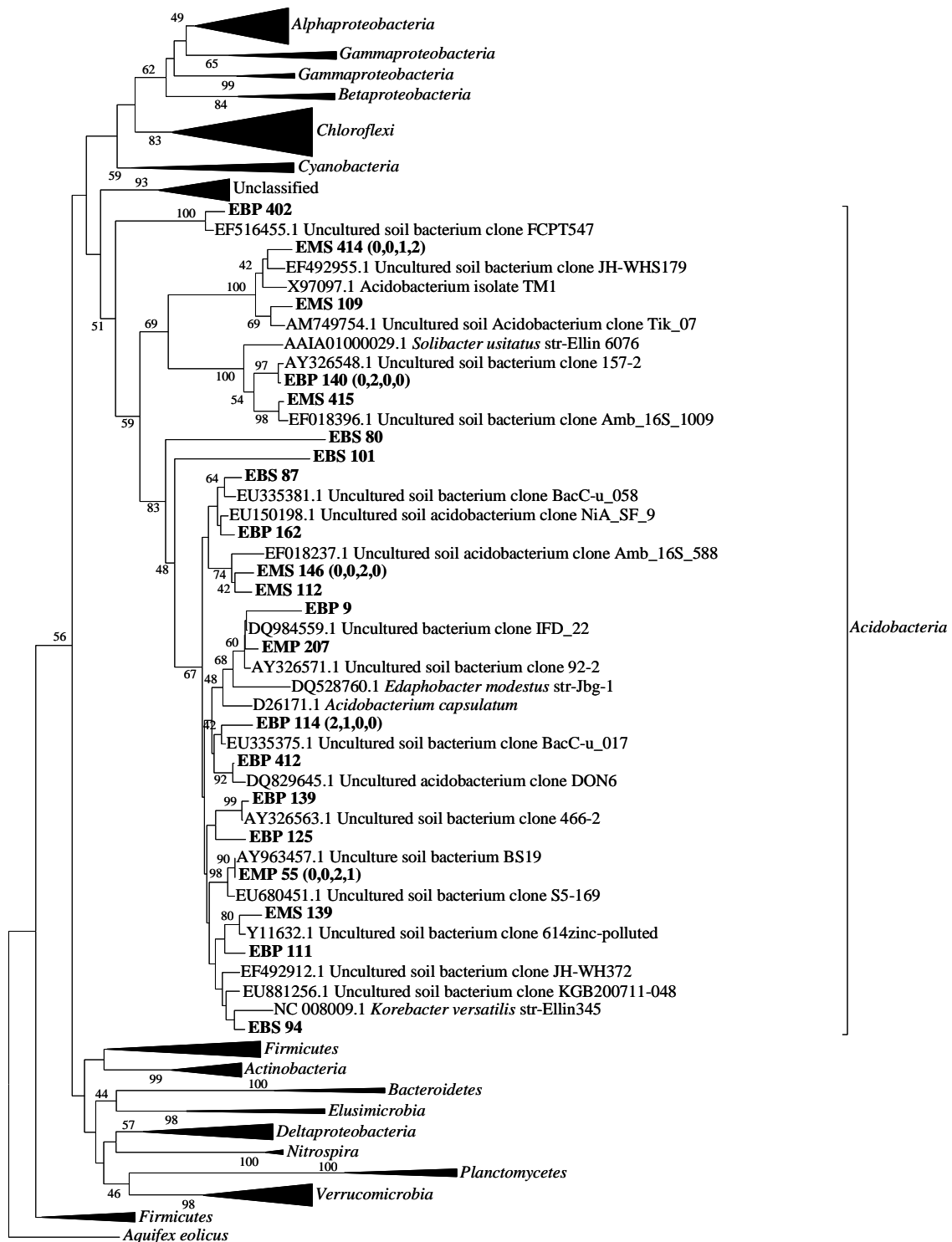


Figure 9a. Phylogenetic relationship of partial 16S rDNA sequences from clone libraries. A total of 286 aligned bases corresponding to 5' end terminal of partial 16S rDNA sequences were used in this analysis. The letters EMS followed by a clone number indicates soil clones. The numbers in parentheses indicate the number of times the OTU was found in each of the libraries e.g., EMS414 (0,0,1,2) was not found in the EBS and EBP communities and was present 1 and 2 times in the EMS and EMP communities, respectively. The scale bar represents 0.05 substitutions per base position. Bootstrap values above 40% are shown. *Aquifex aeolicus* served as the outgroup. See Materials and methods for description of tree construction.

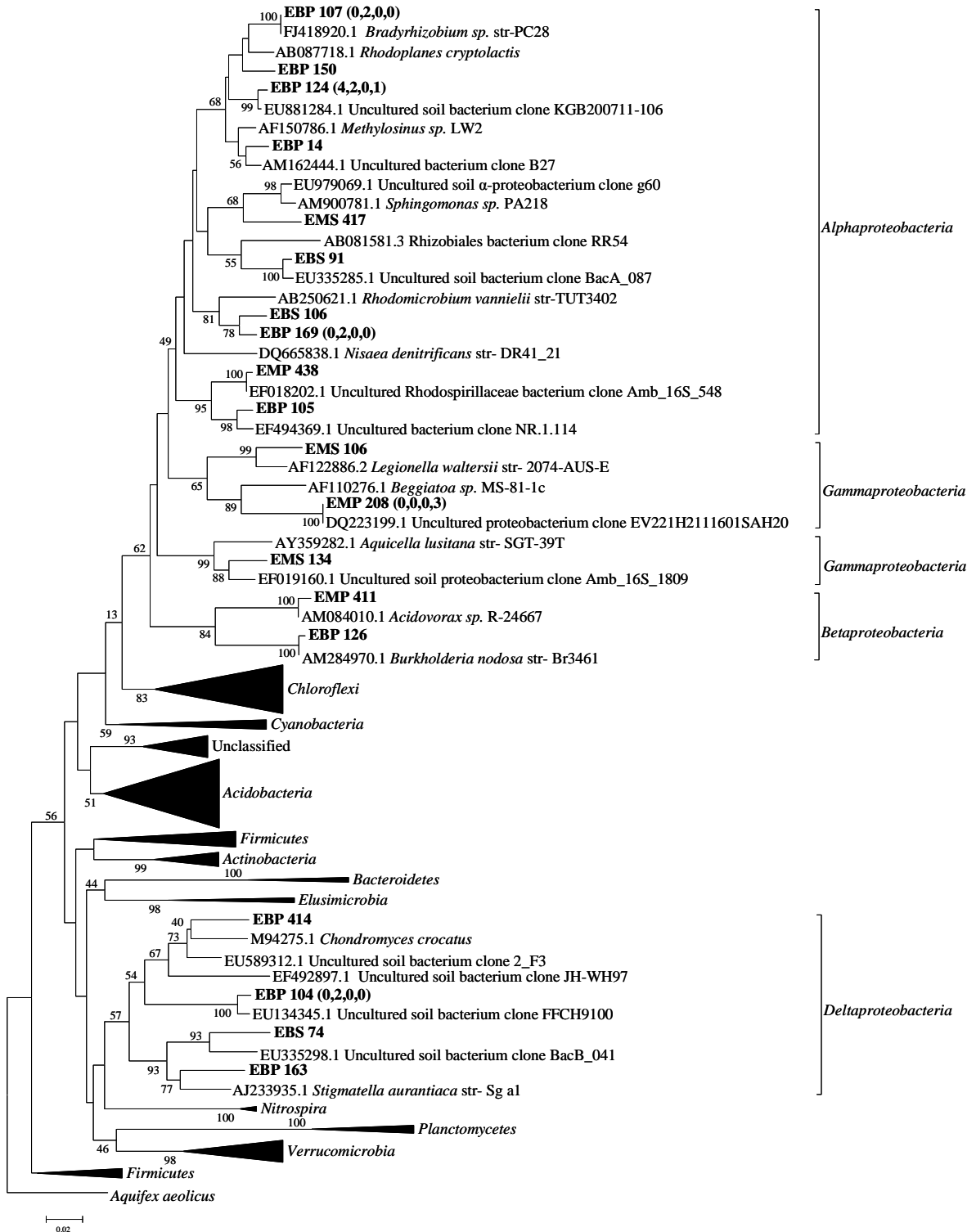


Figure 9b. continued. Phylogenetic tree corresponding to *Proteobacteria* 5' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.

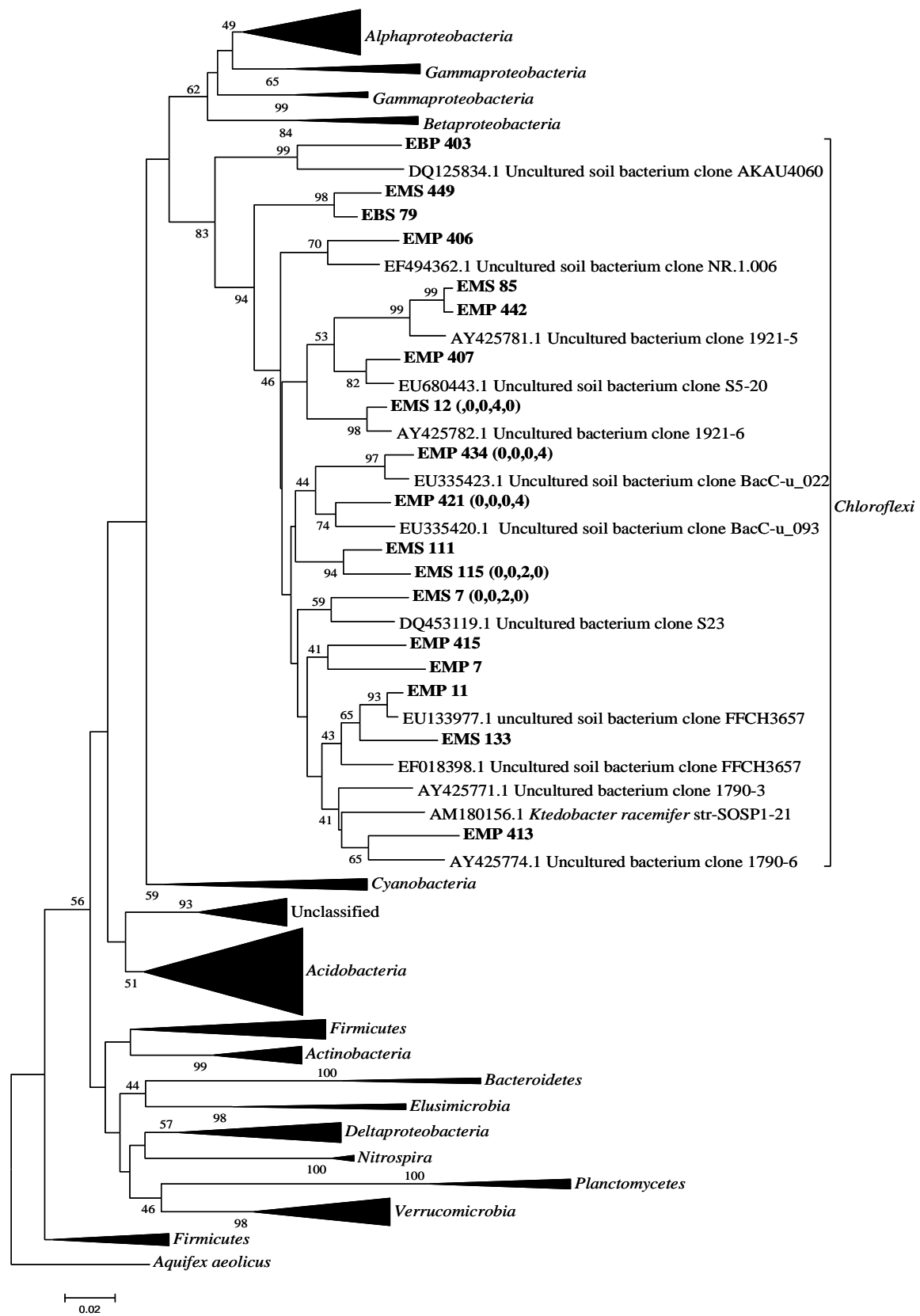


Figure 9c. continued. Phylogenetic tree corresponding to *Chloroflexi* 5' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.

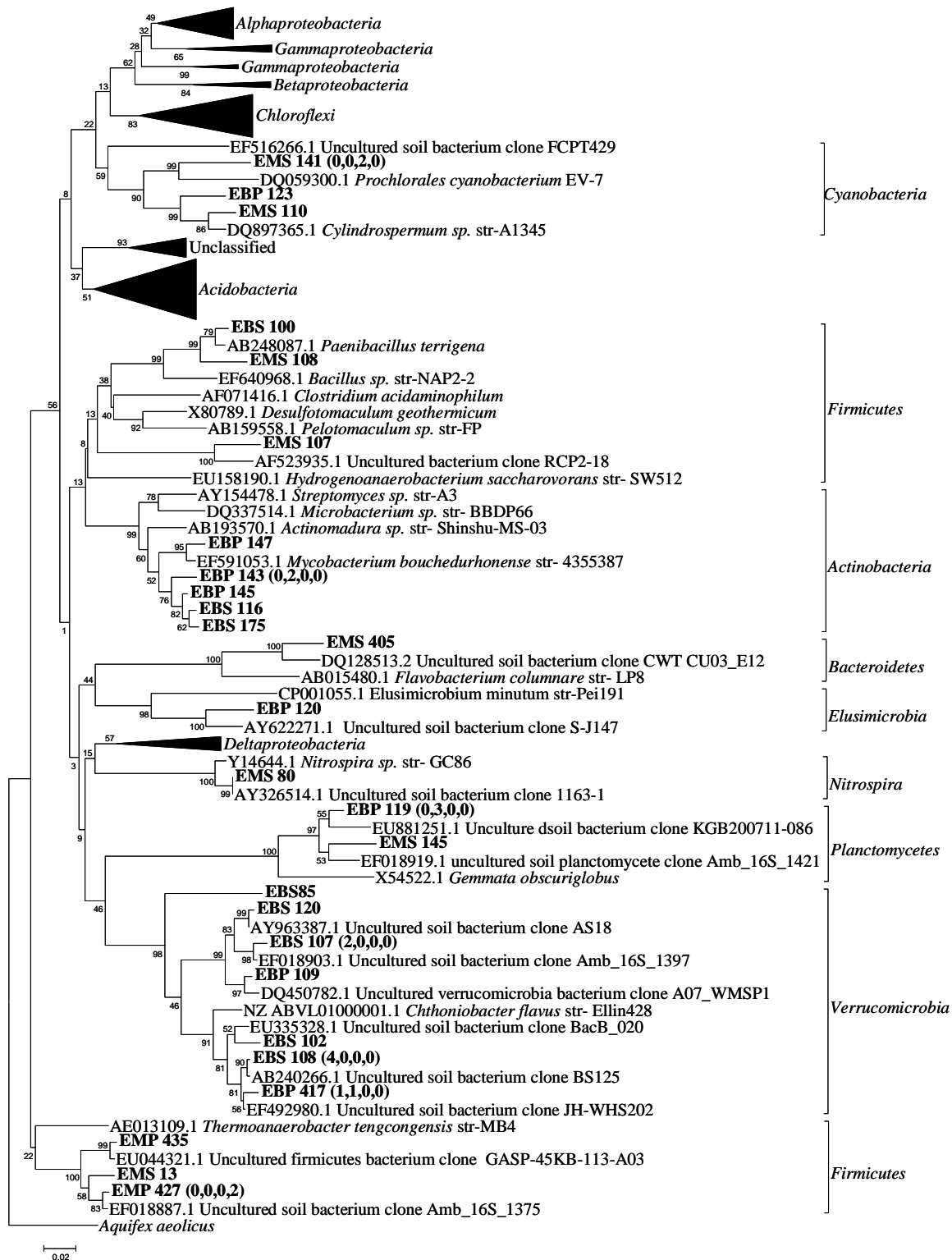


Figure 9d. continued. Phylogenetic tree corresponding to various phyla 5' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.

As revealed in the 3' end phylogenetic tree in Figure 10a, twenty-six out of seventy-one OTUs from the 3' end of the 16S RNA sequences from EBS, EBP, EMS, EMP clone libraries were assigned to the phylum *Acidobacteria*. OTUs representing members of the *Acidobacteria* were the most abundant in the 3'end sequences data-set, comprising 38 % of total OTUs. An OTU composed of three sequences and represented by the clone sequence EBP110 was closely-related to *Korebacter versatilis* strain Ellin-345. Whereas, an OTU represented by EBS126 was more closely-related to *Solibacter usitatus* strain Ellin-6076. As occurred in the 5' end phylogenetic tree, most of the sequences in 3'end tree was closely related to uncultured environmental clone sequences retrieved from different environmental soil samples. OTUs represented by EBP108, EBP115, EBS71, EMS5 and EBP411 had a sequence similarity between 96 to 99% to uncultured bacterium clones from forest uncultured clones from tropical forest soil samples; two of them represented by EBP 416 and EBP151, had a 99% sequence similarity with Taiwan soil samples clone sequences (Tseng et al., GenBank Accession no. DQ984566.1 and DQ984565.1). One OTU represented by EBS105 shared a 99% sequence similarity with a clone sequence from a Taiwan Fushan Forest soil sample (Tsai et al., 2009). An OTU represented by EBS117 shared a sequence 92% similarity with a clone sequence detected in soil a sample from Western Amazon (Kim et al., 2007). However four OTUs represented by EBS86, EBP136, EMP440 and EBP146 were related to uncultured clones from forest soil in Europe; the first three shared a sequence similarity between 97% to 98% with clone sequences from Cologne, Germany soil samples (Roesh et al., GenBank Accession no. AY913380.1, AY913233.1, and AY913350.1) and the latter OTU had a sequence 99% similarity with a clone sequence retrieved from Donana National Park in Spain (Zimmermann et al., GenBank Accession no. DQ829638.1). The OTUs EBP409, EBP110, EMP404 EBP164 and EBS97 affiliated to

Acidobacteria were closely-related to clone sequences detected in a California grassland soil (Cruz et al., 2009). OTUs represented by EMP417, EMS409, EBS77 and EBS81 clustered with clone sequences from various soils environments such as soil rhizosphere (Lesaulnier et al., 2008), peat soil (Akiyama et al., GenBank Accession no. AB238786.1), soil in karst region (Chen et al., GenBank Accession no. EU881239.1) and iron oxide nodules in sedimentary rock (Yoshida et al., GenBank Accession no. AB179518.1), respectively.

A total of twenty-one OTUs out of seventy-one OTUs from the 3' end sequences data-set were assigned to the phylum *Proteobacteria* and further clustered into α -, δ - and γ -*Proteobacteria*, as shown in Figure 10b. In this data-set, OTUs affiliated with the *Alphaproteobacteria* were the most abundant, comprising 48% of total proteobacterial OTUs, and those representing members of *Deltaproteobacteria* were the second most abundant, 28%. An OTU composed of three sequences and represented by EBP167 showed a sequence 98% similarity to *Bradyrhizobium sp.* strain Th-b2. This strain was isolated by Parker from the root-nodules of legumes *Apios americana* and *Desmodium glutinosum* in eastern North America (Parker, 1999). As revealed by the 3' end phylogenetic tree most α -proteobacterial OTUs were more closely related to uncultured clones from different soil environments. OTUs represented by EBP407, EMS450, EMS403 and EBP157 shared a sequence similarity between 92 to 99% with uncultured bacterial clones retrieved from forest soil samples in Taiwan (Tsai et al., 2009), Germany (Roesch et al., GenBank Accession no. AY913376.1) and China (Chan et al., 2006; Huang et al., DQ 453803.1), respectively. One OTU represented by EBP148 formed a clustered supported by a bootstrap value of 93 with an uncultured bacterial clone detected in an acid-impacted lake from Adirondack, NY, USA (Percent et al., 2008), both sequences shared a sequence similarity of 97%. Whereas other α -proteobacterial OTU, EBP165, clustered with an

uncultured clone from pasture soil samples from Australia (Schoenborn et al., GenBank Accession no. AY395444.1), the cluster between the two sequences had a bootstrap value of 99%. OTUs affiliated to *Deltaproteobacteria*, EMP62, EMS1, EBS115, and EBP116 were phylogenetically related to uncultured clones from different soil environments such as, agricultural soil from Mexico (Ceja-Navarro, GenBank Accession no. EU202851.1), grassland soil from California (Cruz et al., 2009), forest soil from western Amazon (Kim et al., 2007), and from soil rhizosphere (Nakamura et al., GenBank Accession no. AB240359.1), respectively. It is interesting to note that OTUs EMP62 and EMS1 from the deforested field at *Bosque del Pueblo* were more phylogenetically related to uncultured clones that came from agricultural and grassland soil samples, and not to uncultured clones from forest soil samples as EBS115 did. Among the *Gammaproteobacteria* two OTUs, represented by EMP203 and EBP154 shared a high sequence similarity between 97 to 98% with uncultured bacterial clones from a subsurface water of the Kalahari Shield in South Africa (Gihring et al., 2006) and a soil sample from an uranium mining waste pile in Saxony, Germany (Satchanska et al., GenBank Accession no. AJ536870.1). Another OTU, EBP102 formed a cluster supported by a bootstrap value of 99% with an uncultured clone from a tropical forest soil from Taiwan (Tseng et al., GenBank Accession no. DQ984542.1), both sequences shared a 99% sequence similarity.

Thirteen OTUs (18%) from the 3'portion data-set of the 16S rRNA sequences were classified as members of *Chloroflexi*. As showed in Figure 10c, two OTUs represented by EBS119 and EMP448 shared a sequence similarity between 91 to 94% with clone sequences detected in forest soil samples from North Carolina-Georgia border, USA (Upchurch et al., 2008). While three representative sequences, EMS451, EMP445 and EMP447 clustered with clone sequences retrieved from soil aggregates from Oak Ridge, TN (Hansel et al., 2008). In the

case of EMP445 and EMS 451, both formed clusters with the clone sequences retrieved from the soil aggregates supported by a bootstrap value of 99%. Two OTUs represented by EMS116 and EMS65 shared a sequence similarity between 95 to 96 % to uncultured clones obtained from volcanic deposits from Hawaii, USA (Gomez-Alvarez et al., 2007). Both OTUs were in clusters supported by a high bootstrap value. Whereas, three OTUs represented by EBS121, EMP412 and EMP409 were more closely related to uncultured clones from various soil environments such as tall grass prairie soil in central Oklahoma (Elshahed et al., 2008), inland dune fields from Georgia, USA (Tarlera et al., 2008), and soils in karst region from Guangxi, China (Chen et al., 2005), respectively. As occurred in the 5' end-portion-data-set, OTUs from clone libraries at *Bosque del Pueblo* that were affiliated to *Chloroflexi* were most closely related to *Ktedobacter racemifer* strain SOSP1-21 than to any other cultured strain within this phylum.

Members of the *Verrucomicrobia*, *Firmicutes*, *Actinobacteria*, *Planctomycetes*, and *Bacteroidetes* phyla contribute to a lesser extent to the bacterial community within clone libraries at *Bosque del Pueblo*. OTUs affiliated to *Verrucomicrobia* constituted 3 % of the OTUs of the 3'end-data-set. Among the OTUs affiliated to *Verrucomicrobia*, a representative sequences, EBS123 clustered supported by a bootstrap value of 99% with an uncultured clone from a Taiwan Fushan Forest soil sample (Tsai et al., 2009), both sequences shared a 98% sequence similarity. A second verrucomicrobial OTU, EBS111, and an uncultured clone from a grassland soil from California (Cruz et al., 2009) shared a 99% sequence similarity and as shown in figure 10c were clustered 86% of the times. Each of the following phyla *Firmicutes*, *Actinobacteria*, *Planctomycetes*, and *Bacteroidetes* were constituted by only one OTU and accounted for less than 1.4 % of the total OTUs in the 3' end-data-set of the bacterial community from *Bosque del Pueblo*. An OTU affiliated with the *Planctomycetes* and represented by the clone sequence

EMP57 shared a sequence 97% similarity with an uncultured clone from a south-west China forest soil sample (Chan et al., 2006). Based on the boot strap analysis both sequences were clustered 84 percent of the times. The closest related strain to the representative sequence EBS82 was *Propionispira arboris* strain 12B4. Both sequences, EBS82 and *Propionispira arboris* strain 12B4, shared a 91% sequence similarity. According to boot strap analysis both sequences were clustered 96 percent of the times.

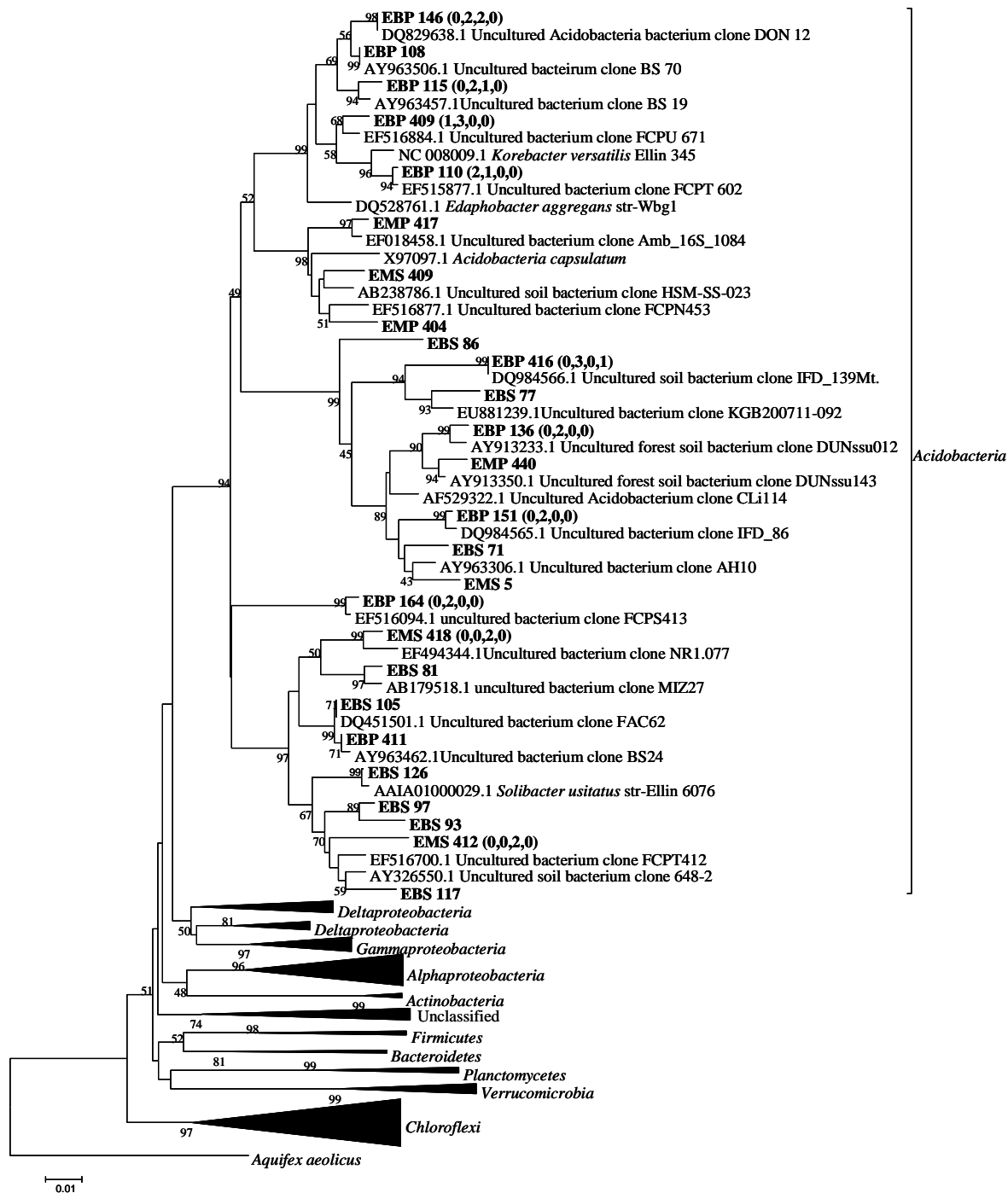


Figure 10a. Phylogenetic relationship of partial 16S rDNA sequences from clone libraries. A total of 422 aligned bases corresponding to 3' end terminal of partial 16S rDNA sequences were used in this analysis. The letters EMS followed by a clone number indicates soil clones. The numbers in parentheses indicate the number of times the OTU was found in each of the libraries e.g., EMS414 (0,0,1,2) was not found in the EBS and EBP communities and was present 1 and 2 times in the EMS and EMP communities, respectively. The scale bar represents 0.02 substitutions per base position. Bootstrap values above 40% are shown. *Aquifex aeolicus* served as the outgroup. See Materials and methods for description of tree construction.

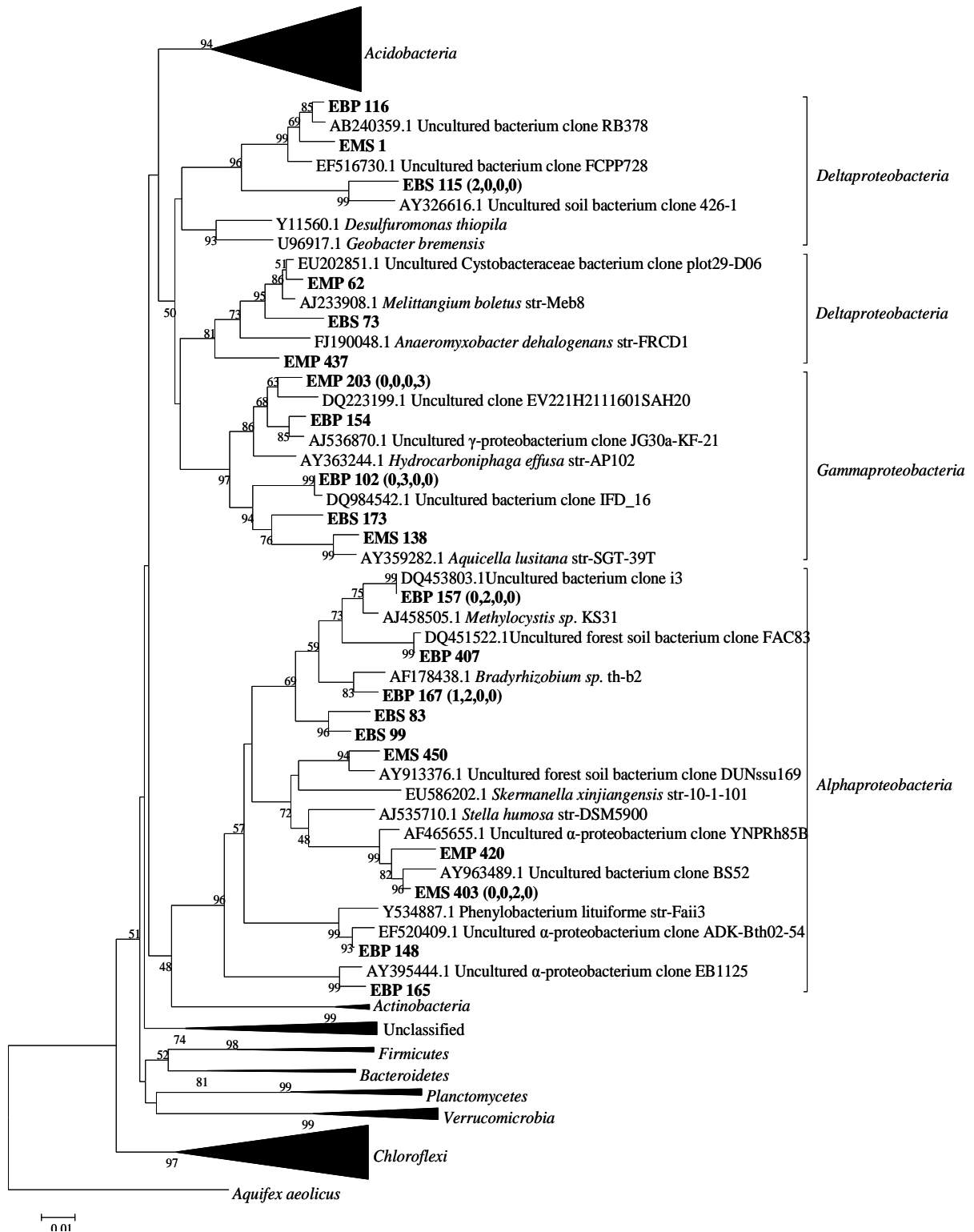


Figure 10b. continued. Phylogenetic tree corresponding to *Proteobacteria* 3' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.

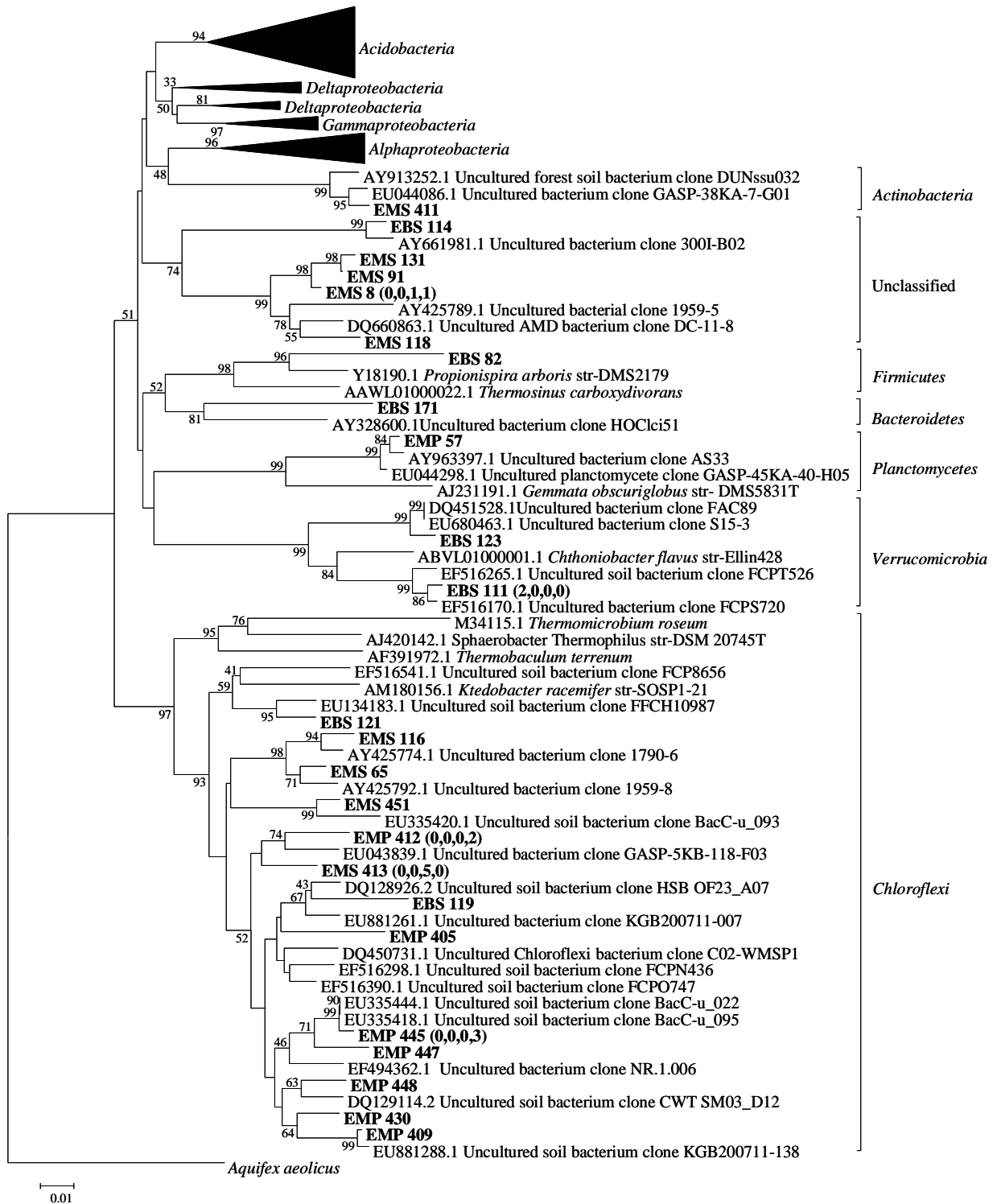


Figure 10c. Phylogenetic tree corresponding to various phyla 3' end terminal of partial 16S rDNA sequences for clone libraries: EBS, EBP, EMS, EMP.

Terminal Restriction Fragment Length Polymorphism (T-RFLP) analyses of the soil bacterial communities at Bosque del Pueblo

In order to study the soil bacterial community structure at *Bosque del Pueblo*, a 16S rDNA gene terminal restriction fragment length polymorphism (T-RFLP) analysis was performed for EBS, EBP, EMS and EMP soil samples. The total community 16S rDNA gene for each soil sample was digested separately with restriction endonucleases *HaeIII*, *MspI* and *RsaI*. After standardization of TRF data, fragment length and peak height were considered as criteria for TRFLP profile comparison. Appendix VI, shows the electropherogram profiles generated with *HaeIII*, *MspI* and *RsaI* restriction enzymes for the soil samples from *Bosque del Pueblo*. Soil samples from *Bosque del Pueblo* harbor a wide range of bacterial communities as revealed by the T-RFs profiles. In all restriction enzymes digestions the T-RFLP profiles show a higher number of peaks for the secondary forest soil samples (EBS and EBP), Table 4. It was also observed in each restriction endonuclease analysis that T-RFLP profiles from secondary forest soil samples showed more similarity among them than with deforested soil samples. The same observation also can be made for deforested soil samples, Figures 11-13.

Table 4. Occurrence of T-RFs in T-RFLP profiles of 16S rDNA amplified gene in soil samples at Bosque del Pueblo, Adjuntas, PR

Soil Sample	Number of T-RFs for restriction enzymes		
	<i>HaeIII</i>	<i>MspI</i>	<i>RsaI</i>
EBS	29	42	27
EBP	32	49	26
EMS	13	24	19
EMP	27	16	21

In *HaeIII* digests, Figure 11, only 2 bacterial populations represented by T-RFs 223 and 196 bp were common for all soil samples. Among the 29 T-RFs generated for EBS soil sample, only three T-RFs (473, 277 and 265 bp) were unique. In the case of EBP soil sample 5 out of 32 T-RFs generated were unique, and were represented by 407, 377, 290, 233 and 179 bp. Bacterial populations represented by T-RFs 526, 523, 312, 262, 256, 240, 237, 216 and 201 bp were common for EBS and EBP soil samples. The soil samples from the deforested soil shared two bacterial communities represented by T-RFs 261 and 236 bp. Among the 13 T-RFs presented for EMS soil sample two T-RFs (71 and 53 bp) were unique. While bacterial populations represented by T-RFs 332, 309, 293, 284, 208, 161 and 117 bp were only found at EMP soil sample.

The T-RFLP profiles generated by *MspI* restriction enzyme revealed 2 bacterial populations, represented by 267 and 257 bp, were common for all soil samples, Figure 12. EBS and EBP soil samples generated 42 and 49 T-RFs, respectively. These secondary forest soil samples shared 32 bacterial populations. Five T-RFs, represented by 506, 371, 280, 221 and 155 bp, were unique to EBS soil sample. While EBP soil sample was characterized by 13 unique T-RFs, represented by 444, 435, 287, 240, 211, 173, 168, 126, 108, 90, 71 and 54 bp. The *MspI* restriction profile revealed that the deforested soil samples generated fewer amounts of T-RFs than the secondary forest soil samples. Thirteen bacterial populations, represented by 522, 326, 297, 285, 251, 242, 234, 225, 215, 208, 196, 192, 187 and 75 bp, were common for EMS and EMP soil samples. In the case of EMS soil sample 4 T-RFs, represented by 322, 311, 306 and 218 bp, were only found in this sample. While none bacterial population were unique for EMP soil sample.

According to the *RsaI* restriction profile one bacterial population, represented by T-RFs 77bp, was common for all four soil samples, Figure 13. Sixteen bacterial populations, represented by T-RFs 658, 641, 490, 476, 466, 459, 454, 451, 442, 423, 118, 111, 105, 92, 89, 84 and 55 bp, were found in both secondary forest soil samples (EBS and EBP). In *RsaI* restriction analyses, 27 T-RFs were found in EBS soil sample, while EBP soil sample generated 26 T-RFs. Among the 27 T-RFs, four bacterial populations were unique for EBS soil sample and were represented by T-RFs 645, 484, 437, 58 bp. While 5 out of 26 T-RFs generated by *RsaI* were found only in EBP soil sample and were represented by T-RFs 568, 472, 462, 433 and 115 bp. Fourteen T-RFs were common for both deforested soil samples (EMS and EMP); and were represented by T-RFs 522, 326, 267, 256, 246, 241, 237, 232, 223, 218, 214, 207, 192 and 187 bp. Only one bacterial population represented by T-RFs 297 was unique for EMS soil sample; and two bacterial populations (T-RFs 612 and 322 bp) were unique for EMP soil sample.

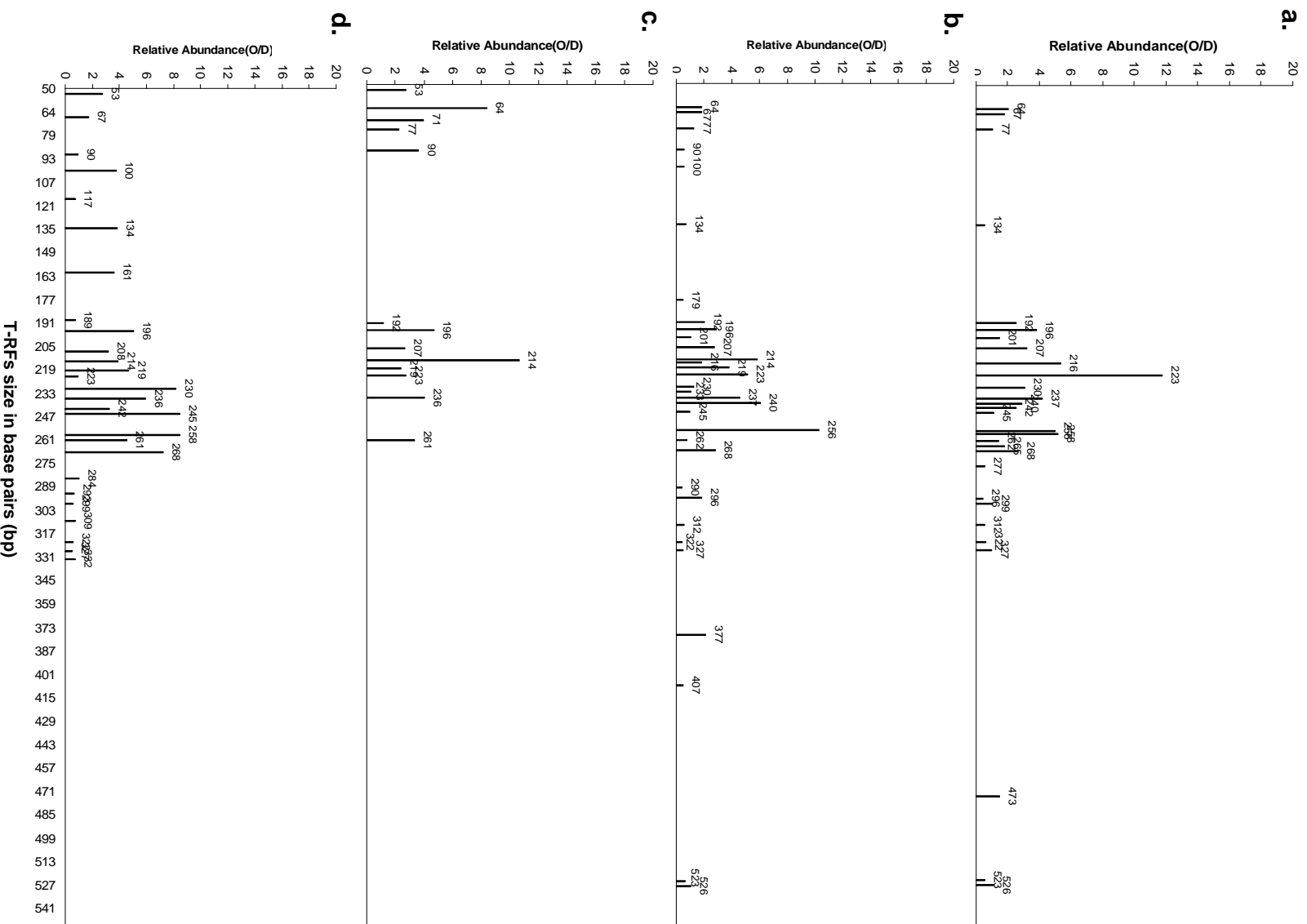


Figure 11. Representative TRF profiles of bacterial communities from Bosque del Pueblo's soil samples a) EBS, b) BBP, c) EMS and d) EMP for *Hae III* restriction analyses.

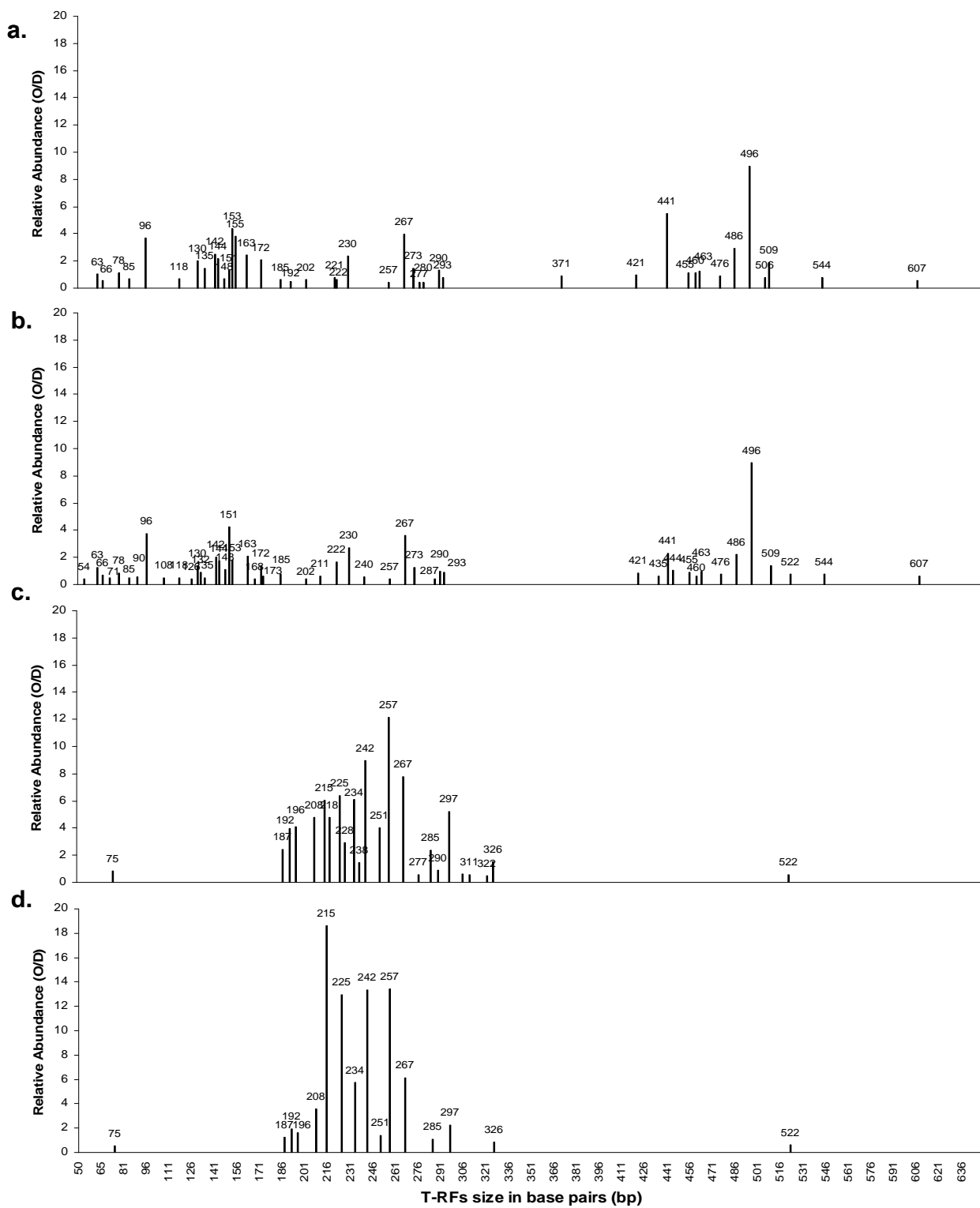


Figure 12. Representative TRF profiles of bacterial communities from Bosque del Pueblo's soil samples a) EBS, b) EBP, c) EMS and d) EMP for *Msp I* restriction analyses.

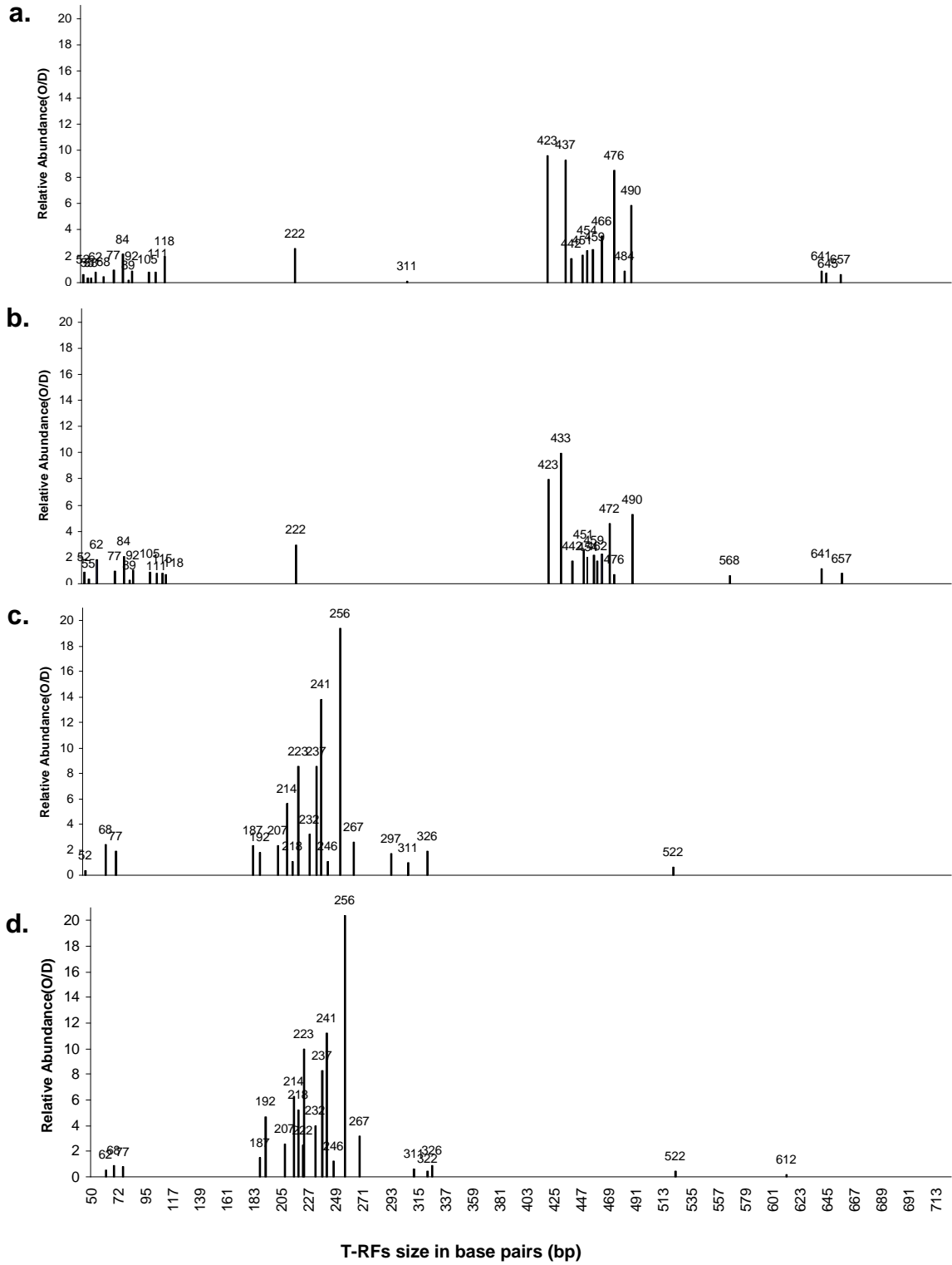


Figure 13. Representative TRF profiles of bacterial communities from Bosque del Pueblo's soil samples a) EBS, b) EBP, c) EMS and d) EMP for *Rsa I* restriction analyses.

The Phylotype Richness, Shannon-Weiner and Simpson's indexes were calculated as a measure of diversity, using Past software, Table 5. Among all the soil samples, secondary forest soils (EBS and EBP) showed the highest phylotype richness values. The values of Shannon-Weiner diversity index were higher for secondary forest soil samples, while no differences were found in the Simpson's index among all soil samples under study.

Table 5. TRF diversity data from Bosque del Pueblo's soil samples bacterial community profiles

Community soil samples	TRF diversity statistics		
	Phylotype Richness (S)	Shannon-Weiner diversity index (H)	Simpson's diversity index (1/D)
EBS	98	4.17	1.02
EBP	107	4.25	1.02
EMS	56	3.65	1.02
EMP	64	3.65	1.02

Furthermore, similarity comparison of the T-RFLP profiles using Past software revealed that soil bacterial communities from secondary forest soil samples (EBS and EBP) were more similar among them (83.2%), in terms of composition, than with deforested soil samples (EMS and EMP), Table 6. The bacterial community composition among deforested soil samples (EMS and EMP) had a similarity value of 79.2 %.

Table 6. Similarity value (%) of Bosque del Pueblo's soil samples T-RFLP bacterial community profiles

T-RFLP	% Similarity			
	EBS	EBP	EMS	EMP
EBS	-	83.2	9.8	13.6
EBP	83.2	-	14.4	12.0
EMS	9.8	14.4	-	79.2
EMP	13.6	12.0	79.2	-

Dendrogram in Figure 14, illustrates the cluster analysis of bacterial community T-RFLP profiles inhabiting the different soil environments (EBS, EBP, EMS, EMP). The dendrogram clearly shows that microbial communities in the secondary forest soil samples clustered and were distinct from the bacterial communities of the deforested soil samples, which clustered separately.

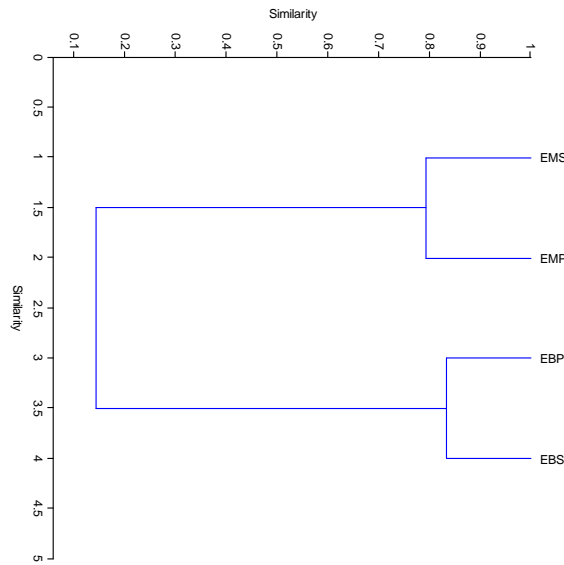


Figure 14. Dendrogram based on Morisita similarity comparisons for EBS, EBP, EMS, EMP soil bacterial communities.

CHAPTER V

DISCUSSION

The aims of this study were i) to use molecular approaches to characterize the diversity and community structure of soil bacterial community at two sampled sites in an abandoned unmined porphyry deposits at *Bosque del Pueblo*, and ii) to correlate sampled sites' microbial diversity with physical and chemical properties of these locations' soil and water streams. In order to assess the bacterial community composition at *Bosque del Pueblo*, 16S rRNA clones libraries were constructed from soil top (0-5 cm, S) and bottom (5-10 cm, P) sections at the two sampling sites, one which was covered with forest vegetation (EB) and the other uncovered by vegetation (EM). Analysis of 16S rRNA gene allowed detection of the most dominant members of the bacterial communities at each sampling sites.

Based on the phylogenetic analysis of the 16S rRNA genes sequences, 29%, 41%, 29% and 15% of the sequences from EBS, EBP, EMS and EMP, respectively, were classified as members of the *Acidobacteria*.

The phylum *Acidobacteria* is recognized on Bergey's Manual of Systematic Bacteriology and includes four validated genera with cultured representatives: *Acidobacterium* (Kishimoto et al., 1991), *Geothrix* (Coates et al., 1999), *Holophaga* (Liesack et al., 1994), and *Terriglobus* (Schmidt et al., 2007). *Solibacter* was recently proposed as a new genus and is pending for validation [<http://jdi.doe.gov>]. The first member of the *Acidobacteria*, *Acidobacterium capsulatum*, was first described in the early 1990s by Kishimoto and collaborators (Kishimoto et al., 1991). This gram negative-acidophilic, heterotrophic bacteria was isolated from acidic mine drainage and acidic mud. While *Acidobacterium capsulatum* is a moderately acidophilic aerobic

heterotroph, *Holophaga foetida* and *Geothrix fermentans* are strict anaerobes that metabolize aromatic compounds and acetate, respectively (Pace et al., 1998).

Members of *Acidobacteria* have been consistently detected through 16S rRNA gene surveys in many environments worldwide, particularly in soils and sediments. In many molecular surveys of soil environments using 16S rRNA clone libraries, *Acidobacteria* represent more than 30% or even 50% of the sequences of the soil bacteria community. Approximately 50% of the clones obtained in four 16S rDNA clone libraries from Arizona soil samples were *Acidobacteria* (Dunbar et al., 2002) as did 64 % of the clones in two soil samples from the arid southwestern United States (Kuske et al., 1997). In a study reported by (Hackl et al., 2004), sequences belonging to *Acidobacteria* were found in high concentration (35%, 28%, 12%) in each of the three 16S rRNA clone libraries from Austrian pine forest soil samples. (Janssen, 2006) analyzed the pool data of thirty-two libraries of 16S rRNA and 16S rRNA genes from different soils samples worldwide. In that study Janssen indicated that members of *Acidobacteria* make up an average of 20% (range, 5 to 46%) of the total soil bacterial community.

A high abundance of members of the *Acidobacteria* has been observed in many other environments such as uranium-contaminated sediments in Tennessee and Colorado (Kuske et al., 2007), microbial mats in Yellow Stone Park (Bryant et al., 2007), acid mine drainage sediments in Yanahara pyrite mine in Japan (Kishimoto et al., 1991), river stream waters in central Germany (Marxsen et al., 2008), lake snow aggregates and activated sludge from waste water treatment plants in Germany (Ludwig et al., 1997), and benzene-contaminated aquifers (Anderson et al., 1998) , among others.

As in other soil environments, *Acidobacteria* was also consistently detected as a major fraction of the microbiota in all four soil clone libraries of *Bosque del Pueblo*. It is interesting to note that results from the phylogenetic analysis as well as in the sequence similarity analysis performed in the present study indicate that many of the sequences from the libraries that can be classified as members of *Acidobacteria*, were further clustered with members of subdivisions 1 and 3 of this phylum. The mild acidity of the soil in both forested and deforested field with pH values in the vicinity of 4.5 and 4.2, respectively seems to favor the growth of *Acidobacteria* members, especially members of subdivision 1. In a study done by (Janssen, 2006), he observed that members of the subdivision 1 of *Acidobacteria* were most abundant in clone libraries from soils with pH values below 6. Members of subdivision 1 of *Acidobacteria* include genera: *Acidobacterium*, *Terriglobus*, and *Edaphobacter*, (Kuske et al., 2009, Janssen et al., 2006). Although, we found sequences closely-related to members of subdivision 3 at *Bosque del Pueblo*, it seems that the preference for acidic pH is a distinctive characteristic of members of subdivision 1 of *Acidobacteria*, as concluded in another study, (Kuske et al., 2009). In a recent study, three *Acidobacteria* strains: *Koribacter versatilis* strain Ellin-345, and *Solibacter usitatus* strain Ellin-6076 were compared in terms of their genomics and culture traits by (Kuske et al., 2009). Kuske's results showed that all three strains have the capability to use a variety of carbon substrates from simple sugars, amino acids, and alcohols to more complex substrates such as xylan, hemicelluloses, pectin, starch and chitin. So, it can be concluded in our study that members of *Acidobacteria* found in forested soil at *Bosque del Pueblo* play an important role in organic matter decomposition. In the genomic analysis of the same study, all three genomes contained a low number of rRNA gene copies, which is a characteristic of organisms that show a slow growth. As physiological and genomic data have revealed, these aerobic heterotrophic

strains are also able to overcome stress or starvation periods and to tolerate fluctuations in hydration (moisture content in soil). This fact could explain the high abundance of members of *Acidobacteria* in the sample soil at the deforested field from *Bosque del Pueblo*. Apparently the low carbon and low nitrogen content in the deforested soil at *Bosque del Pueblo* created a low nutrient environment, where only microorganisms with well-equipped metabolic mechanisms are suited to survive. Based on the results reported by (Kuske et al., 2009) we concluded that the mildly acidity, and a high organic matter and organic carbon content favor the growth of members of *Acidobacteria* in soil forested samples. However, the versatility and sophisticated metabolic mechanisms found in members of *Acidobacteria* enable them to flourish in deforested soil samples characterize by a low nutrient, low organic matter and organic carbon content.

The abundance and wide distribution of *Acidobacteria* as well as its phylogenetic diversity suggest an important ecological role and broad metabolic versatility comparable to *Proteobacteria* (Quaiser et al., 2003). However, little is known about their physiology and potential functions in the environment (Kuske et al., 2009). Although, more experimental data is needed, it seems that *Acidobacteria* play a significant role as decomposers of complex carbon forms in soils.

In EBS and EBP libraries, the sequences of clones belonging to *Proteobacteria* were the second most abundant sequences, 24 % and 36 %, respectively. However, sequences representing this bacterial group were the second and third most abundant in the EMP and EMS, 23 % and 13%, respectively. These results are in agreement with the results in the studies of (Dunbar et al., 2002) and (Zhou et al., 1997) where proteobacterial sequences comprised 15% and 53%, respectively, of the sequences libraries obtained from Arizona soils and Siberian tundra soil. Janssen studied the microbial diversity in a pool data of thirty-two libraries of

16SrRNA genes from different soil samples worldwide. Janssen reported that members of *Proteobacteria* were the most dominant bacterial group in libraries prepared from soil bacterial communities, making up an average of 39% (range, 10 to 77%) of the sequences retrieved (Janssen, 2006). *Proteobacteria* is a very large and diverse group of gram-negative bacteria divided into five subgroups designated as alpha, beta, gamma, delta and epsilon. The first four subgroups are commonly found in soils.

Based on phylogenetic analysis, all 16S rRNA gene sequences related to *Bradyrhizobium* at *Bosque del Pueblo* were detected in the forested field, which could correlate with root biomass present at this site. None 16S rRNA sequences retrieved from deforested site at *Bosque del Pueblo* were phylogenetically related to *Bradyrhizobium*. This slow-growing bacterium lives symbiotically in plant root nodules, especially in root nodules of legumes, and is able to take atmospheric nitrogen and fix it into ammonia (NH₃) or ammonium (NH₄)⁺, forms usable by plants. In exchange the plants provide carbohydrates to *Bradyrhizobium*. So, it seems that 16S rRNA sequences related to this group of bacteria play a main role in the nitrogen fixation process in the rhizosphere at *Bosque del Pueblo*.

Sequences detected in the forested field at *Bosque del Pueblo* were also phylogenetic related to *Rhodomicrobium* genera. According to Bergey's Manual of Systematic Bacteriology, *Rhodomicrobium vannielii* is the only species of this genus. This photoheterotrophic species has been found in freshwater and seawater environments and grow best in slightly acidic conditions with pH values between 5.2 and 6.5. *Rhodomicrobium sp.* lives with simple organic substrates as carbon and electron sources, and prefers dinitrogen and ammonia as nitrogen sources. So apparently the slightly acidic pH and the high carbon and nitrogen content of the soil at the forested field at *Bosque del Pueblo* favor the growth of *Rhodomicrobium sp.*

The 16S rRNA gene sequence from *Stella humosa* strain DSM5900 was the cultivated strain most closely related to the two OTUs, represented by EMP420 and EMS403, from the deforested field at *Bosque del Pueblo*. However, none of the OTU from the forested field at *Bosque del Pueblo* was phylogenetically related to *Stella* strain. *Stella Humosa* is one of the two species of *Stella* genus recognized by Bergey's Manual of Bacteriology. *Stella humosa* was first isolated from soil by Vasilyeva (Vasilyeva, 1985); it was also found in freshwater habitats. The genus *Stella* is characterized by star-shaped cells. In the oligotrophic environment that prevails in the deforested soil at *Bosque del Pueblo* this morphology increases the surface/mass ratio of the bacteria, a competitive advantage to absorb nutrients from the soil. It has been reported (Vasilyeva and Zavarzin, 1995) that this aerobic oligotrophic group utilizes the products of particulate organic matter decomposition formed by other microorganisms; they even are able to take traces of organic substrates from the air. The physiologic and metabolic features of this strain provide them with the tools necessary to successfully compete against fast-growing bacteria that require higher concentrations of organic compounds. Conversely, the inability of this group to catabolize complex carbohydrate or protein polymers such as cellulose, starch and gelatin, inhibits a successful establishment of them in an environment rich in organic compound. Even more, the low growth rate of this strain means they cannot surpass the fast-growing bacteria that inhabit in the forested soil at *Bosque del Pueblo*, an environment rich in nutrients and organic compounds.

An OTU represented by EMS450 formed a cluster with *Skermanella xinjiangensis* strain 10-1-101 of the class *Alphaproteobacteria*. This novel species of the genus *Skermanella* was isolated from a desert soil of Xinjiang, China (Hongli et al., 2009) and its validation is pending.

It is a chemoorganotrophic bacterium unable to hydrolyze complex organic compounds such as starch and cellulose or to fix nitrogen.

16 S rRNA sequences, represented by OTU EBP157 and OTU EBP14, from the forested soil at *Bosque del Pueblo* were also phylogenetically related to alphaproteobacterial *Methylocystis* sp. and *Methylosinus* sp., respectively. These bacteria are methanotrophs, gram-negative obligate aerobes that use only methane and methanol as sole carbon and energy sources. There are two kinds of methane oxidizing bacteria: type I methanotrophs and type II methanotrophs. The first group is composed of members of the *Gammaproteobacteria* subdivision and the latter group belongs to the *Alphaproteobacteria* subdivision (Knief et al., 2005). Methanotrophic bacteria contributes significantly to the global flux of methane by oxidizing atmospheric methane, and also by oxidizing the methane produced by methanogenic bacteria during the decomposition of organic matter in soils. *Methylocystis* and *Methylosinus* have been detected in sample soils from Thailand (Knief et al., 2005) as well as from rice field soils (Le Mer and Roger, 2001), lake sediments (Costello and Lidstrom, 1999), acidic peat soil (Dedysh et al., 2006), and upland forests soils (Degelmann et al, 2010; Amaral et al., 1998). Forest soils are more efficient methane sink than other kind of soil such as cultivated soils. It has been suggested that methane produced in forest soil sub layers, particularly under wet conditions, by methanogenic bacteria may support the methanotrophic activity in the upper aerobic layers of forest soils (Amaral et al., 1998; Le Mer and Roger, 2001). This fact may explain why sequences retrieved from the deforested soil samples at *Bosque del Pueblo* were not phylogenetically related to methanotropic bacteria.

A cluster supported by a bootstrap value of 100% was formed between an OTU represented by EMP411 and *Acidovorax* sp. strain R-24667 isolated from an activated sludge

(Heylen et al., 2006). According to Bergey's Manual of Bacteriology, *Acidovorax* species can be isolated from soil, water, activated sludge, and infected plants. Members of *Acidovorax* are gram-negative, and most species are capable of heterotrophic denitrification of nitrate to nitrous oxide (N₂O) and nitrogen gas (N₂). During the process some N₂O and NO gases are released to the atmosphere. Microbial activity in soils is known to play a key role in the production and consumption of N₂, N₂O, and NO, although the precise mechanism by which it occurs is complicated (Rösch et al., 2002). The OTU represented by EMP411 was retrieved from the 5 to 10 cm soil layer from the deforested soil sample at *Bosque del Pueblo*, so apparently the sequence comes from a bacterium responsible for denitrification at this site. Previous studies showed that denitrifying bacteria occurs abundantly in the upper soil layer (Rösch et al., 2002). In a similar way, another OTU represented by EBP126 and *Burkholderia nodosa* strain Br3461 formed a single cluster supported by a bootstrap value of 100% and with a sequence similarity of 97%. *Burkholderia nodosa* strain Br3461 was first isolated from nitrogen-fixing nodules on the roots of *Mimosa scabrella* in Brazil (Chen et al., 2005). As described by Bergey's Manual of Bacteriology members of *Burkholderia* genus are chemoorganotrophs and are found in a wide range of environments, including soil, water, plants, animals and humans; some are of medical importance. Previous studies have reported that some *Burkholderia* species, such as *Burkholderia nodosa* strain Br3461, are able to fix nitrogen in symbiotic association with herbaceous legumes from different continents (Garau et al., 2009). In the present study the sequence EBP126 was retrieved from the soil surrounding the plant roots in the forested field at *Bosque del Pueblo* where it seems to play an important role in increasing plant nutrient availability via nitrogen fixation. As previously mentioned in this study, the soils samples at *Bosque del Pueblo* reported acidic values. In a study done in South Africa by Garau and

collaborators (Garau et al., 2009) *Burkholderia* strains were isolated from acidic soil (pH 5.5), Garau concluded that these strains play an important role in nitrogen fixation on acidic soils.

Among the gamma-proteobacterial sequences recovered at *Bosque del Pueblo*, an OTU represented by EMP208 formed a cluster supported by a bootstrap value of 100% and a 98% sequence similarity with an uncultured clone retrieved from acid mine drainage (AMD) of Evander mine in South Africa (Gihring et al. 2006). Another OTU that clustered with the same uncultured clone detected at the AMD of Evander mine in South Africa was an OTU represented by EMP203, in this case the bootstrap value for this cluster was 63% and the similarity between the sequences was 97%. The closest cultivated bacterium to OTU EMP411 in the bank database is *Beggiatoa* sp. strain MS-81-1. *Beggiatoa* is a filamentous sulfur oxidizer genus. Some bacterial sulfur oxidizer use sulfide as an energy source producing sulfuric acid as an oxidation product. As expected, the low- nutrients, organic matter, carbon and nitrogen content in the deforested field at *Bosque del Pueblo* created an environmental niche that supports the growth of microorganism able to utilize the sources present in this soil. It seems that the clone sequence EMP411 belongs to a sulfur oxidizer microorganism that uses inorganic sources of energy at *Bosque del Pueblo*.

Most of the bacterial 16s rRNA sequences obtained in the 0-5 cm (top) and 5-10 cm (bottom) from the EB sample soils belong to members of *Acidobacteria* and *Proteobacteria*. However, most of the sequences recovered in the 0-5 cm and 5-10 cm sections from the deforested sample soils were representatives of the *Chloroflexi*. Based on the phylogenetic analysis of the 16S rRNA genes sequences performed in the present study, 32%, and 48% of the sequences from EMS and EMP, respectively, belonged to members of *Chloroflexi*. In phylogenetic trees constructed, in this study, *Ktedonobacter racemifer* str-SOSP1-21 clustered

with clone sequences from EMS and EMP libraries, cluster was supported by bootstrap value of 99%. SOSP1-21 bacterial strain was the first characterized member of this filamentous, spore-forming and gram-positive novel Class (Cavaletti, 2006). This bacterial strain was isolated from a black locust woods in Gerenzano, Italy and was named *Ktedobacter racemifer*, now known as *Ktedonobacter racemifer* (Euzéby, 2007). *Ktedonobacter racemifer* strain SOS1-21 grew at pH values between 4.2 and 7.2; and was able to grow in aerobic and microaerophilic conditions. Other bacterial strains phylogenetically related to SOSP1-21 were recovered from soil samples at different locations: a pine wood in Spain, an ant house in Honduras and a cereal field in Egypt (Cavaletti et al., 2006).

In our study we observed that members of *Chloroflexi* were abundant in deforested soil samples (EMS and EMP). They were nearly absent in forested soil samples (EBS and EBP). It is important to note that as a consequence of mining explorations in deforested soil samples a subsurface layer of soil was exposed. This newly exposed soil could be considered as a young or immature soil. Taking that in consideration, our results coincide with Tarlera's findings (Tarlera et al., 2008). In her study, Tarlera constructed 16S rRNA gene chronolibraries to assess the composition of bacterial communities as a result of their natural development in inland dunes fields at Georgia, USA. Interestingly, she found that members of *Chloroflexi* were more abundant in the youngest soil of her soil collection. This young soil described by Tarlera was also characterized by a low organic matter content (0.95%) and acidic pH value (4.4). Similar organic matter content (0.52%) and soil pH values (ca. pH 4.0) were measured in our study in deforested soil samples at *Bosque del Pueblo*. Tarlera concluded that *Chloroflexi* is a fundamental member in the soil initial colonization process in inland dunes fields and that the new bacterial community differs from the bacterial community in older soils.

Members of phylum *Verrucomicrobia* accounted 23 %, 3% and 2% of the sequences in EBS, EBP and EMP libraries, respectively. No sequences of *Verrucomicrobia* were detected in EMS library. This phylum is an abundant component of soil microbial communities worldwide. *Verrucomicrobia* has been estimated to comprise 1% to 10 % of the total bacteria in soils and on average constitute 5% of all the bacterial 16S rRNA in soils (Janssen, 2006).

As mentioned before the conditions present at the forested soil samples (EBS and EBP) at *Bosque del Pueblo* were characterized by higher soil organic matter, carbon and nitrogen contents than neighboring deforested soil samples (EMS and EMP). The old-growth forest at EBS and EBP sampling sites constantly deposits leaf litter and woody debris to soil resulting in a rich nutrients environment that support the growth of members of the phylum *Verrucomicrobia*.

On the basis of the 16S rRNA sequences similarity analysis, the nearest cultivated relative to OTUS represented by EBS111, EBS102, EBS108, and EBP417 is *Chthoniobacter flavus* strain-Ellin 428. Janssen's group isolated this strain from a rye grass and clover pasture soil in Victoria, Australia (Janssen et al., 2004). *Chthoniobacter flavus* strain Ellin-428, a member of subdivision 2 of the phylum *Verrucomicrobia*, is a free-living gram-negative soil bacterium able to grow aerobically with many of the saccharide components of plant biomass such as glucose, fructose, galactose among others. This strain grows at pH values between 4.0 to 7.0. No genes related with nitrogen fixation were detected in this strain. So it seems that this bacteria, as in others soil bacterial community previously studied, play an important role in the transformation of organic carbon compounds in the surface of forested soil samples at *Bosque del Pueblo*. Other member of subdivision 2 of *Verrucomicrobia* is *Xiphinematobacter spp*, live as an obligate endosymbiont in ectoparasitic nematodes found in soil and freshwater

environments (Janssen et al., 2004). However, none of the sequences recovered from soil samples at *Bosque del Pueblo* were phylogenetic related to this group.

It was observed in our study that verrucomicrobial 16S sequences abundance decreased with soil depth. Verrucomicrobial 16S rRNA sequences obtained from the forested soil samples (EBS and EBP) were more abundant in 0-5 cm deep soil core, 23 % of the total sequences from EBS clone library, than in 5-10 cm deep soil core, only 3% of the total sequences from EBP clone library. This observation can be explained by the decrease of total organic carbon and total nitrogen with depth of soil. Another variable that decrease with soil depth is soil moisture, which also has been associated with *Verrucomicrobia* abundance in soil. In our study we observed that 94% (15 out of 16) of the total verrucomicrobial sequences recovered from *Bosque del Pueblo* were from forested soil samples (EBS and EBP) while deforested soil samples (EMS and EMP) accounted for only 6% (1 out of 16) of the total verrucomicrobial sequences. So it seems that a high soil moisture content supports the growth of *Verrucomicrobia* members. The data obtained by Schmit and Buckley in 2001 suggests that the abundance of *Verrucomicrobia* is related to soil moisture (Schmit and Buckley, 2001). However, Schmit and Buckley determined that more experimentation is required to understand how *Verrucomicrobia* respond to changes in the soil moisture content. It seems that other bacterial groups with different metabolic strategist can flourish at deeper soils where the resources are less available.

Members of *Verrucomicrobia* have been isolated in terrestrial as well as in marine sediments (Polymenakou et al., 2005), aquatic environments (Lindstrom et al., 2004), acid rock drainage (Okabayashi et al., 2005) and associated with eukaryotic hosts (Janssen et al., 2004). However, the knowledge about the ecological contribution of members of this phylum is scarce, as it occurs with many other bacterial groups.

In EBS, EBP and EMS clone libraries, the clone sequences belonging to phylum *Actinobacteria* comprised 4%, 5% and 2%, respectively. Therefore, the soil bacterial communities at *Bosque del Pueblo* showed lower dominance by *Actinobacteria* as compared with the average of 13% for other soil communities (Janssen, 2006). However, in a study done by Tarlera and collaborators 1.9% of the clone sequences from three clone libraries corresponded to members of *Actinobacteria* (Tarlera et al, 2008). In our study an OTU represented by EMS411 shared a sequence similarity of 98% with a clone sequence retrieved from dunes of southeast Georgia, USA (Tarlera et al. 2008). The soil sample at the inland dunes from which the uncultured clone was obtained had a pH of 4.6, similar to pH values from soil samples at *Bosque del Pueblo* (Table 1). As expected, the soil properties are determinant factors that influence the bacterial communities in soils.

In the present study, members of phylum *Firmicutes* constituted 4%, 5%, and 6% of EBS, EMS, and EMP clone libraries, respectively. None of the 16S rRNA sequences from EBP clone library were affiliated to *Firmicutes*. Our results are in agreement with a study done by (Janssen, 2006). In his study Janssen determined that members of *Firmicutes* contribute a mean of 2% (range, 0 to 8%) to thirty-two 16S rRNA clone libraries that were analyzed. As Janssen explained in his study, members of *Firmicutes* have long been considered common members of the soil bacterial community. He argued that *Firmicutes* may be underrepresented in soil clone libraries because DNA's extraction protocols are unable to properly lyse cells or spores from members of this bacterial group. In our study an OTU represented by EBS100 formed a cluster supported by a bootstrap value of 79% with *Paenibacillus terrigena*. EBS 100 and *Paenibacillus terrigena* type strain-A35 sequences shared a 96% sequence similarity. *Paenibacillus terrigena* type strain-A35 was isolated from a soil sample in Chiba, Japan (Cheng-Hui X et al., 2007).

This gram-positive type strain do not fixes nitrogen nor possesses the *nif* gene. Another OTU represented by EMP435 formed a cluster supported by a bootstrap value of 99% with a clone sequence (45KB-113) from a sample soil of Georgia's inland dune fields (Tarlera et al., 2008). Both sequences shared a 97% sequence similarity. In her study Tarlera reported that clone sequence 45k-113 was detected in a soil sample characterize by low organic matter content (0.76%) and mildly acidic conditions (pH ~4.6). These soil properties are very similar in respect to the organic matter content (0.52%) and a soil pH (ca. 4) measured in our study at the bottom section of deforested soil sample (EMP) at *Bosque del Pueblo*.

Soil bacteria are an essential component in tropical forests because they play a major role in nutrient transformation (Smith and Paul 1990). Land-use changes have been shown to have a significance effects on both the physical and the chemical properties of the soil (Girvan et al., 2003; McCaig et al., 1999). These properties in turn affect the soil bacterial composition and the process they mediate (Lauber et al., 2008). Results from the present study indicate that the shift in cover vegetation from forest to deforested soils caused by mining explorations has led to changes in soil properties, including an increase in acidity, and a decrease in nutrients, organic carbon, organic matter and total nitrogen content. The higher nutrients concentration found in the forest soil than in the deforested soil is likely due to its higher soil organic matter content, which in turn increases the soil cation-exchange capacity (Nüsslein and Tiedje, 1999). Similar changes in soil properties have been reported in other studies (Carney et al., 2004).

In the present study the OTUs composition was unique to each clone library studied. Indeed, none of the OTUs was common to all four clone libraries. However, more OTUs were common to EBS and EBP clone libraries from the forested soil. Likewise, more OTUs were common to EMS and EMP clone libraries from the deforested soil. Results from the LIBSHUFF

confirmed that all four clone libraries were statistically different. This difference in bacterial community composition is parallel to the distinct physicochemical properties determined in the soil samples from which each clone library were obtained.

In this study it was expected to detect 16S rRNA sequences from isolated strains such as *Acidiphilum sp.*, *Acidobacteria capsulatum*, *Terromicrobium acidophilum*, *Leptospirillum ferrooxidans*, and/or closely-related strains because these strains represent the primary constituents of acid mine drainage (AMD) habitats (Brofft et al. 2002). However, 16S rRNA sequences closely-related to common constituents of AMD habitats were no detected in this study. A plausible reason for the absence of the expected 16S rRNA sequences is that soil samples from *Bosque del Pueblo* are less acidic (pH ~3.7 to 4.69) than AMD habitats (pH ~ 0-3). The pH values of the four soil samples from which clone libraries were obtained in this study were mildly acidic which is typical for pine-dominated forest (Brofft et al. 2002). The mildly acidity of soil samples from *Bosque del Pueblo*, a tropical forest, may be due to the soil mineral composition. pH values of the four soil samples at *Bosque del Pueblo* were measured by the two methods as indicated in Table 1 of Results the Section. Interestingly, a higher acidity was observed in deforested soil samples (EMS and EMP) than in forested soil samples (EBS and EBP), regardless the method used to measure acidity. Previously studies have identified soil pH as one of the most important selection factors that directly affect the bacterial composition in soils (Kim et al., 2007; Lima et al. 2002)

Differences in water physicochemical properties of nearby streams at forested and deforested sites at *Bosque del Pueblo* were also observed. These differences include a higher temperature, conductivity, and acidity in the water samples taken from the deforested site than the ones taken from the forested site. Our results also showed that nickel and copper

concentrations (ppm) were two and three times higher, respectively, in water samples from deforested site in comparison with corresponding concentrations from forested site. As expected, the variations in water physicochemical properties as well as nickel and copper concentrations in water samples at two locations were statistically-significant according to ANOVA analysis.

Based on the, (i) physicochemical properties in soil and water samples, and (ii) bacterial community composition observed at the two sites in *Bosque del Pueblo*, can be inferred that removal of vegetation at *Bosque del Pueblo*, resulting in the exposure of subsurface soil layers. In turn, the subsurface soil layers interact with water and atmospheric oxygen which promote metal ions and hydrogen ions release into the environment, altering acidity and metal composition of nearby soils and streams. These events alter the niches altering available to bacteria within soils' environment.

Similar to clone libraries data, comparison of communities TRF profiles revealed compositional differences in the bacterial communities inhabiting at *Bosque del Pueblo*'s soil samples. The soil bacterial communities from secondary forest (EBS and EBP) clustered together and separately from bacterial communities from deforested soil (EMS and EMP), which indicated that they were similar in composition. Our findings were consistent with our expectations based on physical chemical properties of soil and water streams at these locations. The two secondary forest soil samples (EBS and EBP) were characterize by a higher pH, and higher organic matter, organic carbon, and total nitrogen content. This nutrient rich environment seemed to create a wide range of microhabitats within these soils that result in high values of phylotype richness and more diverse bacterial community than in the deforested soil samples (EMS and EMP).

The observed increase in acidity, and decrease in organic carbon, organic matter and total nitrogen content at Bosque del Pueblo's may be caused by a shift in cover vegetation from secondary forest to deforested soil and this effect may be accompanied by other changes in soil properties such as metal ions concentration and acidity. These changes in deforested soil samples (EMS and EMP) represent habitats with lower nutrients content where only bacteria with more sophisticated metabolisms could establish, such as *Chloroflexi*, *Acidobacteria* and *Proteobacteria*.

In general a change from forest vegetation to nude soil constitute a major shift in the soil environment at *Bosque del Pueblo*, especially in carbon sources for the microorganisms, and accompanying changes in soil properties as well as water properties of nearby streams. This vegetation change caused some changes in soil properties including an increase in soil acidity, and a decrease in nutrients availability, organic carbon, organic matter and total nitrogen content. The water streams nearby were also impacted by the change in vegetation at *Bosque del Pueblo*. The water streams suffered an increase in acidity and in nickel and copper trace concentrations. Hence, our study suggests that change from forest vegetation to nude soil, which in turn caused changes in soil and water properties is a determinant of the current microbial community composition observed at the two sites sampled. Furthermore, now the bacterial community at each sampling sites also are modeling the soil environment and determinates which bacterial groups could establish within the soil.

Considering that the unmined porphyry deposits at *Bosque del Pueblo* were abandoned decades ago, and differences in the microbial community composition at the two sampling sites can still be observed, it would be interesting to formulate the following questions. How long succession to a secondary forest will take? Will the forest return to its original state? As a result

of secondary succession which changes in carbon substrates, and soil and water properties will occur? And, will the microbial community at *Bosque del Pueblo* will return to its original state?

CONCLUSIONS

- In terms of the physico-chemical properties of soil and water streams at the sites, the organic matter, organic carbon, and total nitrogen content in the forested soil samples were higher than the corresponding content in deforested soil samples. The removal of vegetation cover caused a significantly decrease in the organic matter input to the soil and created a low nutrient environment. It was also observed that water samples from deforested site exhibited nickel and copper concentrations (ppm) two and three times higher than samples from forested site. As a results of vegetation cover change the soils were exposed to water and atmospheric oxygen, causing the released of metals ions and hydrogen ions into the environment. These differences in physico-chemical properties of water streams samples from Bosque del Pueblo may be caused by a shift in vegetation cover from secondary forest to deforested soil.
- Based on the phylogenetic analysis of 16S rDNA gene sequences, members of *Acidobacteria* phylum were the most dominant bacterial population at secondary forest soil samples (EBS and EBP). While in deforested soil samples (EMS and EMP) *Chloroflexi* members were the major constituents. These differences in bacterial community composition may be attributed to differences in soil samples' physico-chemical properties from two locations in Bosque del Pueblo.
- According to diversities indexes, the bacterial composition was more diverse in *Estacion Bosque* than in *Estacion Mina*. Apparently, the low nutrients environment in the deforested soil at Bosque del pueblo favors the establishment of microorganisms with more specialized metabolic mechanisms.

- PCR-amplified 16S rRNA clone libraries and Terminal Restriction Fragment Length Polymorphism (T-RFLP) analysis demonstrated to be useful tools to study the bacterial diversity from Bosque del Pueblo's soil samples.

LITERATURE CITED

Amaral J.A., T. Ren, and R. Knowles. 1998. Atmospheric Methane Consumption by Forest Soils and Extracted Bacteria at Different pH Values. *Appl. Environ. Microb.* 64: 2397-2402.

Anderson R. T., J. N. Rooney-Varga, C. V. Gaw, and D. R. Lovely. 1998. Anaerobic benzene oxidation in the Fe(III) reduction zone of petroleum-contaminated aquifers. *Environ. Sci. Technol.* 32:1222–1229.

Atlas R.M and R. Bartha. 1997. *Microbial Ecology: Fundamentals and Applications*. 4th Ed. pp 99, 108, 219, 290. Benjamin/Cummings, California.

Barabas A.H. 1982. Potassium-argon dating of magmatic events and hydrothermal activity associated with porphyry copper mineralization in west central Puerto Rico. *Econ. Geol.* 77: 109-126.

Blackwood C. B., T. Marsh, S. Kim, and E.A. Paul. 2003. Terminal Restriction Fragment Length Polymorphism Data Analysis for Quantitative Comparison of Microbial Communities. *Appl. Environ. Microb.* 69: 926-932.

Bond P.L., G.K. Druschell, J.F. Banfield. 2000. Comparison of acid mine drainage microbial communities in physically and geochemically distinct ecosystems. *App. Environ. Microb.* 66: 4962-4971.

Bond P.L., S.P. Smriga, J.F. Banfield. 2000. Phylogeny of microorganisms populating a thick, subaerial, predominantly lithotrophic biofilm at an extreme acid mine drainage site. *Appl. Environ. Microb.* 66: 3842-3849.

Bonkowski, M. 2004. Protozoa and plant growth: the microbial loop in soil revisited. *New Phytol.* 162: 617–631.

Borneman J., and E.W. Triplett. 1997. Molecular microbial diversity in soils from eastern amazonia: evidence for unusual microorganisms and microbial population shift associated with deforestation. *Appl. Environ. Microb.* 63: 2647-2653.

Bürgmann H., F. Widmer, W.V. Sigler, and J. Zeyer. 2004. New molecular screening tools for analysis of free-living diazotrophs in soil. *Appl. Environ. Microb.* 70: 240-247.

Burrell P.C., J. Keller, and L.L. Blackall. 1998. Microbiology of a nitrite-oxidizing bioreactor. *Appl. Environ. Microb.* 64: 1878-1883.

Brencic A., and S.C. Winans. 2005. Detection of and Response to Signals Involved in Host-Microbe Interactions by Plant-Associated Bacteria. *Microbiol. Mol. Biol.* 69: 155 - 194.

- Brofft J.E., J.V. McArthur, and L.J. Shimkets. 2002. Recovery of novel bacterial diversity from a forested wetland impacted by reject coal. *Environ. Microbiol.* 4:764-769.
- Bryant D.A., A.M. García Costas, J.A. Maresca, A.G. 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: 523-526.
- Byrne D.B., G. Suarez, and W.R. McCann. 1985. Muertos Trough subduction- Microplate tectonics in the northern Caribbean: *Nature.* 317: 420-421.
- Carney K.M., P.A. Matson, and B.J.M. Bohannan. 2004. Diversity and composition of tropical soil nitrifiers across a plant diversity gradient and among land-use types. *Ecol Lett.* 7: 684-694.
- Cavaletti L., P. Monciardini, R. Bamonte, P. Schumann, M. Rohde, M. Sosio, and S. Donadio. 2006. New lineage of filamentous, spore-forming, gram-positive bacteria from soil. *Appl Environ Microb.* 72: 4360-4369.
- Chan O.C.¹, X. Yang, Y. Fu, Z. Feng, L. Sha, P. Casper, and X. Zou. 2006. 16S rRNA gene analyses of bacterial community structures in the soils of evergreen broad-leaved forests in south-west China. *FEMS Microbiol Ecol.* 58: 247–259.
- Cheng-Hui X. and A. Yokota. 2007. *Paenibacillus terrigena* sp. nov., isolated from soil. *Int J Syst Evol Micr.* 57: 70-72.
- Chen W.M., S. M. de Faria, J.-H. Chou, E. K. James, G. N. Elliott, J. I. Spret, C. Bontemps, J. P. W. Young, and P. Vandamme. 2008. *Burkholderia sabiae* sp. nov., isolated from root nodules of *Mimosa caesalpinifolia*. *Int J Syst Evol Microb.* 58: 2174-2179.
- Chen, W.M., S. M.de Faria, R. Stralio, R.M. Pitard, J.L. Simoes-Araujo, J.H. Chou, Y. J. Chou, E. Barrios, A.R. Prescott, G. N. Elliott, J. I. Spret, J. P. W. Young, and E. K. James. 2005. Proof that *Burkholderia* strains form effective symbioses with legumes: a study of novel *Mimosa*-nodulating strains from South America. *Appl Environ Microb.* 71: 7461–7471.
- Cline J.S. 2003. How to Concentrate Copper. *Science.* 302:2075-2076.
- Coates J., D. J. Ellis, C. V. Gaw, and D. R. Lovley. 1999. *Geothrix fermentans* gen. nov., sp. nov., a novel Fe(III)-reducing bacterium from a hydrocarboncontaminated aquifer. *Int J Syst Evol Micr.* 49:1615–1622.
- Costello A.M. and M.E. Lidstrom. 1999. Molecular Characterization of Functional and Phylogenetic Genes from Natural Populations of Methanotrophs in Lake Sediments. *Appl Environ Microb.* 65: 5066-5074.

- Costello, E.K., S.R. Halloy, S.C. Reed, P. Sowell, and S.K. Schmidt. 2009. Fumarole-supported islands of biodiversity within a hyperarid, high-elevation landscape on Socompa Volcano, Puna de Atacama, Andes. *Appl Environ Microb.* 75: 735-747.
- Cottorell M.T., and D.L. Kirchman. 2000. Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence in situ hybridization. *Appl Environ Microb.* 66: 5116-5122.
- Cox D.P., I.P. González, J.T. Nash. 1975. Geology, geochemistry, and fluid-inclusion petrography of the Sapo Alegre porphyry copper prospect and its metavolcanic wallrocks, west-central Puerto Rico. *Jour Research U.S. Geol Survey.* 3: 313-327.
- Cruz-Martínez K., K. Suttle, E.L. Brodie, M.E. Power, G.L. Andersen, and J.F. Banfield. 2009. Despite strong seasonal responses, soil microbial consortia are more resilient to long-term changes in rainfall than overlying grassland. *ISME Journal.* 3: 738-744.
- Dedysh S. N., T. A. Pankratov, S. E. Belova, I. S. Kulichevskaya, and W. Liesack. 2006. Phylogenetic Analysis and In Situ Identification of *Bacteria* Community Composition in an Acidic *Sphagnum* Peat Bog. *Appl Environ Microb.* 72: 2110-2117.
- Degelmann D. M., W. Borken, H. L. Drake, and S. Kolb. 2010. Different Atmospheric Methane-Oxidizing Communities in European Beech and Norway Spruce Soils. *Appl Environ Microb.* 76: 3228-3235.
- DeSantis T.Z., P. Hugenholtz, N. Larsen, M. Rojas, E.L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G.L. Andersen. 2006. Greengenes, a chimera-checked 16SrRNA gene database and workbench compatible with ARB. *App Environ Microb.* 72: 5069- 5072.
- Dunbar J., S.M. Barns, L.O. Ticknor, and C.R. Kuske. 2002. Empirical and Theoretical Bacterial Diversity in Four Arizona Soils. *App Environ Microb.* 68: 3035- 3045.
- Dunbar J., L.O. Ticknor, and C.R. Kuske. 2000. Assessment of microbial diversity in four southwestern United States soils by 16S rRNA gene terminal restriction fragment analysis. *App Environ Microb.* 66: 2943-2950.
- Ellis R.J., P. Morgan, A.J. weightman, and J.C. Fry. 2003. Cultivation-dependent and independent approaches for determining bacterial diversity in heavy-metal-contaminated soil. *App. Environ. Microb.* 69: 3223-3230.
- Edwards K.J., M.O. Schrenk R.M. Goodman, R.J. Hamers, and J.F. Banfield. 1998. Distribution of *Thiobacillus ferrooxidans* and *Leptospirillum ferrooxidans*: Implications for generation of acid mine drainage. *Science.* 279:1519-1522.
- Edwards K.J., P.L. Bond, T. M. Gihring, and J.F. Banfield. 2000. An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. *Science.* 287: 1796-1799.

- Edwards, K.J., M.O. Schrenk, R. Hamers, and J.F. Banfield. 1998. Microbial oxidation of pyrite: experiments using microorganisms from an extreme acidic environment. *Am Mineral*. 83:1444–11453.
- Elshahed M.S., N.H. Youssef, A.M. Spain, C. Sheik, F.Z. Najjar, L.O. Sukharnikov, B.A. Roe, J.P. Davis, P.D. Schloss, V.L. Bailey, and L.R. Krumholz. 2008. Novelty and uniqueness patterns of rare members of the soil biosphere. *Appl Environ Microb*. 74: 5422 – 5428.
- Euzéby J. 2007. List of new names and new combinations previously effectively, but not validly, published. *Int J Syst Bacteriol*. 57: 433-434.
- Ewing B., L.D. Hillier, M.C. Wendl, and P. Green. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research* 8:175-185.
- Garau G., R.J. Yates, P. Deiana, and J.G. Howieson. 2009. Novel strains of nodulating *Burkholderia* have a role in nitrogen fixation with herbaceous legumes adapted to acid, infertile soils. *Soil Biol Biochem*. 41: 125–135.
- Gihring T.M. , D. P. Moser, L.H. Lin, M. Davidson, T.C. Onstott, L. Morgan, M. Milleson, T. L. Kieft, E. Trimarco, D. L. Balkwill, and M. E. Dollhopf. 2006. The Distribution of Microbial Taxa in the Subsurface Water of the Kalahari Shield, South Africa. *Geomicrobiol J*. 23: 415-430.
- Girvan M.S. J., Bullimore, J.N. Pretty, A.M. Osborn, and A.S. Ball. 2003. Soil type is the primary determinant of the composition of the total and active bacterial communities in arable soils. *Appl Environ Microb*. 69: 1800-1809.
- Gomez-Alvarez V., G. M. King, and K. Nüsslein. 2007. Comparative bacterial diversity in recent Hawaiian volcanic deposits of different ages. *FEMS Microbiol Ecol*. 60: 60-73.
- Hackl E., S. Zechmeister-Boltenstern, L. Bodrossy, and A. Sessitsch. 2004. Comparison of Diversities and Compositions of Bacterial Populations Inhabiting Natural Forest Soils. *Appl Environ Microb*. 70: 5057-5065.
- Hammer, Ø., Harper, D.A.T., and P. D. Ryan. 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica* 4(1): 9pp. http://palaeo-electronica.org/2001_1/past/issue1_01.htm
- Hansel C.M., S. Fendorf, P.M. Jardine, and C.A. Francis. 2008. Changes in Bacterial and Archaeal Community Structure and Functional Diversity along a Geochemically Variable Soil Profile. *Appl Environ Microb*. 74: 1620-1633.
- He,J., L. Zhang, S. Jin, Y. Zhu, and F. Liu. 2008. Bacterial communities inside and surrounding soil iron-manganese nodules. *Geomicrobiol J* . 25:14-24.
- Herlemann, D.P., O. Geissinger, W. Ikeda-Ohtsubo, V. Kunin, H. Sun, A. Lapidus, P. Hugenholtz, and A. Brune. 2009. Genomic analysis of 'Elusimicrobium minutum,' the first

cultivated representative of the phylum 'Elusimicrobia' (formerly termite group 1). *Appl Environ Microb.* 75: 2841-2849.

Heylen K., L. Lebbe, and P. De Vos. 2001. *Acidovorax caeni* sp. nov., a denitrifying species with genetically diverse isolates from activated sludge. *Int J Syst Evol Microb.* 58: 73-77.

Hongli A., L. Zhang, Y. Tang, X. Luo, T. Sun, Y. Li, Y. Wang, J. Dai, and C. Fang. 2009. *Skermanella xinjiangensis* sp. nov., isolated from the desert of Xinjiang, China. *Int J Syst Evol Microb.* 59:1531-1534.

Hugenholtz P., B.M. Goebel, N.R. Pace. 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Appl Environ. Microb.* 180: 4765-4774.

Janssen P.H. 2006. Identifying the dominant soil Bacteria taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microb.* 72: 1719-1728.

Janssen P.H., Sait M., and K. E.R. Davis. 2006. Effect of pH on isolation and distribution of members of subdivision 1 of the Phylum *Acidobacteria* occurring in soil. *Appl Environ Microb.* 72: 1852-1857.

Janssen P.H., P. Sangwan, X. Chen, and P. Hugenholtz. 2004. *Chthoniobacter flavus* gen. nov., sp. nov., the First Pure-Culture Representative of Subdivision Two, *Spartobacteria* classis nov., of the Phylum *Verrucomicrobia*. *Appl Environ Microb.* 70: 5875-5881.

Kemp P.F and J. Y. Aller. 2004. Estimating prokaryotic diversity: When are 16S rDNA libraries large enough? *Limnol Oceanogr Meth.* 2: 114-125

Kim J.S., G. Sparovek, R.M. Longo, and W.J. De Melo. 2007. Bacterial diversity of terra preta and pristine forest soil from the Western Amazon. *Soil Biol Biochem.* 39: 684-690.

Kishimoto, N., Y. Kosako, and T. Tano. 1991. *Acidobacterium capsulatum* gen. nov., sp. nov.: an acidophilic chemoorganotrophic bacterium containing menaquinone from acidic mineral environment. *Curr Microbiol.* 22:1-7.

Knief, C., S. Vanitchung, N.W. Harvey, R. Conrad, P.F. Dunfield, and A. Chidthaisong. 2005. Diversity of methanotrophic bacteria in tropical upland soils under different land uses. *Appl Environ Microb.* 71: 3826-3831.

Kuske C.R., J. Dunbar, S.M. Barns, L.O. Ticknor. 2002. Empirical and Theoretical Bacterial Diversity in Four Arizona Soils. *App. Environ. Microbiol.* 68: 3035-3045.

Kuske C.R., N. L. Ward, J. F. Challacombe, P. H. Janssen, B. Henrissat, P.M. Coutinho, M. Wu, G. Xie, D.H. Haft, M. Sait, J. Badger, R.D. Barabote, B. Bradley, T.S. Brettin, L.M. Brinkac, D. Bruce, T. Creasy, S.C. Daugherty, T.M. Davidsen, R.T. DeBoy, J.C. Detter, R.J. Dodson, A. Scott Durkin, A. Ganapathy, M. Gwinn-Giglio, C.S. Han, H. Khouri, H. Kiss, S.P. Kothari, R. Madupu, K.E. Nelson, W.C. Nelson, I. Paulsen, K. Penn, Q. Ren, M. J. Rosovitz, J. D. Selengut, S. Shrivastava, S. A. Sullivan, R. Tapia, L.S. Thompson, K.L. Watkins, Q. Yang, C.Yu, N.

- Zafar, and L. Zhou. 2009. Three genomes from the Phylum *Acidobacteria* provide insight into the lifestyles of these microorganisms in soils. *Appl Environ Microb.* 75: 2046-2056.
- Kuske C.R., S.M. Barns, E.C. Cain, and L. Sommerville. 2007. *Acidobacteria* Phylum Sequences in Uranium-Contaminated Subsurface Sediments Greatly Expand the Known Diversity within the Phylum. *Appl Environ. Microb.* 73: 3113-3116.
- Kuske C.R., S.M. Barns, and J.D. Bush. 1997. Diverse uncultivated bacterial groups from soils of the arid southwestern United States that is present in many geographic regions. *Appl Environ Microb.* 63: 3614-3621.
- Lauber, C.L., M.S. Strickland, M.A. Bradford, and N. Fierer. 2008. The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biol Biochem.* 40: 2407-2415.
- Le Mer, J.L., and P. Roger. 2001. Production, oxidation, emission, and consumption of methane by soils: A review. *Eur J Soil Biol.* 37: 25–50.
- Lesaulnier C, D. Papamichail, S. McCorkle, B. Ollivier, S. Skiena, and S. Taghavi S. 2008. Elevated atmospheric CO₂ affects soil microbial diversity associated with trembling aspen. *Environ Microbiol.* 10: 926–941.
- Liesack, W., F. Bak, J.U. Kreft, and E. Stackebrandt. 1994. *Holophaga foetida* gen. nov., sp. nov., a new, homoacetogenic bacterium degrading methoxylated aromatic compounds. *Arch Microbiol.* 162: 85–90.
- Liles M.R., B.F. Manske, S.B. Bintrim, J. Handelsman, and R.M. Goodman. 2003. A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Appl Environ Microb.* 69: 2684-2691.
- Lima, H.N., C.E.R. Schaefer, J.W.V. Mello, R. J. Gilkes, and J.C. Ker. 2002. Pedogenesis and pre-Colombian land use of “Terra Preta Anthrosols” (“Indian black earth”) of Western Amazonia. *Geoderma* 110: 1–17.
- Lindström E. S., K. Vrede, and E. Leskinen. 2004. Response of a member of the Verrucomicrobia, among the dominating bacteria in a hypolimnion, to increased phosphorus availability. *J Plankton Res.* 26: 241–246.
- Ludwig, W., S. H. Bauer, M. Bauer, I. Held, G. Kirchhof, R. Schulze, I. Huber, S. Spring, A. Hartmann, and K.-H. Schleifer. 1997. Detection and *in situ* identification of representatives of a widely distributed new bacterial phylum. *FEMS Microbiol. Lett.* 153:181–190.
- Marxsen J., S. Beier, K.P. Witzel. 2008. Bacterial Community Composition in Central European Running Waters Examined by Temperature Gradient Gel Electrophoresis and Sequence Analysis of 16S rRNA Genes. *Appl. Envir. Microb.* 74: 188-199.

- Massol-Deyá A., and E. Díaz. 2003. Trace elements composition in forage samples from various anthropogenically impacted areas in Puerto Rico. *Caribb J Sci.* 39: 215-220.
- McCaig, A.E., L.A. Glover, and J.I. Prosser. 1999. Molecular Analysis of Bacterial Community Structure and Diversity in Unimproved and Improved Upland Pastures. *Appl Environ Microb.* 9: 1800-1809.
- Moyer C. L., and J. J. Engebretson. 2003. Fidelity of Select Restriction Endonucleases in Determining Microbial Diversity by Terminal-Restriction Fragment Length Polymorphism. *Appl Environ Microb.* 69: 4823-4829.
- Nannipieri P., J. Ascher, M.T. Ceccherini, L. Landi, G. Pietramellara, and G. Renella. 2003. Microbial diversity and soil functions. *Eur J Soil Sci.* 54: 655-670.
- Nüsslein K., and J.M Tiedje. 1998. Characterization of the dominant and rare members of a young Hawaiian soil bacterial community with small-subunit ribosomal DNA amplified from DNA fractionated on the basis of its guanine and cytosine composition. *Appl. Environ. Microb.* 64: 1283-1289.
- Nüsslein K., and J.M Tiedje. 1999. Soil Bacteria Community Shift correlated with change from forest to pasture vegetation in tropical soil. *Appl. Environ. Microb.* 65: 3622-3626.
- Okabayashi A., S. Wakai, T. Kanao, T. Sugio, and K. Kamimura. 2005. Diversity of 16S Ribosomal DNA-Defined Bacterial Population in Acid Rock Drainage from Japanese Pyrite Mine. *J Biosci Bioeng.* 100:644-652.
- Osborn A.M., E.R.B. Moore, and K.N. Timmis. 2000. An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. *Environ Microbiol.* 1: 39-50.
- Pace N.R. 1997. A molecular view of microbial diversity and the biosphere. *Science.* 276: 734-740.
- Pace N.R., P. Hugenholtz, and B.M. Goebel. 1998. Impact of Culture-Independent Studies on the emerging Phylogenetic view of Bacterial Diversity. *J. Bacteriol.* 180: 4765-4774.
- Parker, M.A. 1999. Relationships of bradyrhizobia from the legumes *Apios americana* and *Desmodium glutinosum*. *Appl Environ Microb.* 65: 4914-4920.
- Peixoto R.S., H.L. da Costa Coutinho, N.G. rumjanek, A. Macrae, and A.S. Rosado. 2002. Use of *rpoB* and 16S rRNA genes to analyse bacteria diversity of a tropical soil using PCR and DGGE. *Lett Appl Microbiol.* 35: 316-320.

- Percent S.F., M.E. Frischer, P.A. Vescio, E. B. Duffy, V. Milano, M. McLellan, B. M. Stevens, C.W. Boylen, and S.A. Nierzwicki-Bauer. 2008. Bacterial Community Structure of Acid-Impacted Lakes: What Controls Diversity? *Appl Environ Microb.* 74: 1856-1868.
- Peterson G. and M. Heemskerk. 2001. Deforestation and forest regeneration following small-scale gold mining in the Amazon: the case of Suriname. *Environ Conserv.* 28: 117–126.
- Plotkin J.B., M.D. Potts, N. Leslie, N. Manokarane, J. LaFrankie and P.S. Ashton. 2000. Species-area Curves, Spatial Aggregation, and Habitat Specialization in Tropical Forests. *J Theor Biol.* 207: 81-99
- Poly F., I. Ranjard, S. Nazaret, F. Gourbière, L.J. Monrozier. 2001. Comparison of *nifH* gene pools in soils and soil microenvironments with contrasting properties. *Appl Environ Microb.* 67: 2255-2262.
- Polymenakou P.N., S. Bertilsson, A. Tselepides, and E.G. Stephanou. 2005. Bacterial community composition in different sediments from the Eastern Mediterranean Sea: a comparison of four 16S ribosomal DNA clone libraries. *Microb Ecol.* 50: 447–462.
- Quaiser A., T. Ochsenreiter, C. Lanz, S.C. Schuster, A.H. Treusch, J. Eck, and C. Schleper. 2003. Acidobacteria form a coherent but highly diverse group within the bacterial domain: evidence from environmental genomics. *Mol Microbiol.* 50: 563-575.
- Rawlings, D.E. 2005. Review: Characteristics and adaptability of iron- and sulfur-oxidizing microorganisms used for the recovery of metals from minerals and their concentrates. *Microb Cell Fact.* 4:13-28.
- Ritchie N.J., M.E. Schutter, R. P. Dick, and D.D. Myrold. 2000. Use of length heterogeneity PCR and fatty acid methyl ester profiles to characterize microbial communities in soil. *Appl Environ Microb.* 66: 1668-1675.
- Rosado A.S., R.S. Peixoto, H.L. da Costa Coutinho, N.G. Rumjanek and A. Macrae. 2002. Use of *rpoB* and 16S rRNA genes to analyse bacterial diversity of a tropical soil using PCR and DGGE. *Lett Appl Microbiol.* 35: 316–320.
- Rösch C, A. Mergel, and H. Bothe. 2002. Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil. *Appl Environ Microb.* 68: 3818–3829.
- Sandaa R.A., V. Torsvik, Ø. Enger, F. L. Daae, T. Castberg, and D. Hahn. 1999. Analysis of bacterial communities in heavy metal-contaminated soils at different levels of resolution. *FEMS Microbiol Ecol.* 30:237-251.
- Schloss P.D., and J. Handelsman. 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microb.* 71: 1501-1506.

Schmidt T.M. and D.H. Buckley. 2001. Environmental factors influencing the distribution of rRNA from Verrucomicrobia in soil. *FEMS Microbiol Ecol.* 35:105-112.

Schmidt T.M., S.A. Eichorst, and J. A. Breznak. 2007. Isolation and Characterization of Soil Bacteria That Define *Terriglobus* gen. nov. in the Phylum *Acidobacteria*. *Appl Environ Microb.* 73: 2708-2717.

Shimkets L. and C.R. Woese. 1992. A phylogenetic analysis of the myxobacteria: Basis for their classification. *Proc Natl Acad Sci.* 89: 9459-9463.

Singh B.K., L. Nazaries, S. Munro, I. C. Anderson, and C. D. Campbell. 2006. Use of Multiplex Terminal Restriction Fragment Length Polymorphism for Rapid and Simultaneous Analysis of Different Components of the Soil Microbial Community *Appl Environ Microb.* 72: 7278-7285.

Singleton, D. R., M. A. Furlong, S. L. Rathbun, and W. B. Whitman. 2001. Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. *Appl Environ. Microb.* 67:4374-6.

Smith J.L., and E. Paul. 1990. The significance of soil biomass estimations. *Soil Biol Biochem.* Vol. 6 (Eds j. M. Bollog and G. Stotzky). pp 357-96. Marcel Dekkar, New York.

Spröer c., H. Reichenbach, and E. Stackebrandt. 1999. The correlation between morphological and phylogenetic classification of myxobacteria. *Int J Syst Bacteriol* 49: 1255-1262.

Tamura K, J. Dudley, M. Nei and S. Kumar. 2007. *MEGA4*: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution.* 24: 1596-1599.

Tarlera S., K. Jangid, A. H. Ivester, W. B. Whitman, and M. A. Williams. 2008. Microbial community succession and bacterial diversity in soils during 77000 years of ecosystem development. *FEMS Microbiol Ecol* 64:129–140.

Torsvik, V.L., R. Sørheim, and J. Goksoyr. 1996. Total bacterial diversity in soil and sediment communities – a review. *J Ind Microbiol.* 17: 170–178.

Torsvik V., L. Ovreas, and T.F. Thingstad. 2002. Prokaryotic diversity-magnitude, dynamics, and controlling factors. *Science.* 296: 1064-1966.

Tsai S.-H., A. Selvam, Y.-P. Chang, and S.-S. Yang. 2009. Soil bacterial community composition across different topographic sites characterized by 16S rRNA gene clones in the Fushan Forest of Taiwan. *Bot Stud.* 50: 57-68.

Tyson G.W., J. Chapman, P. Hugenholtz, E.E. Allen, R.J. Ram, P.M. Richardson, V.V. Solovyev, E.M. Rubin, D.S. Rokhar, and J.F. Banfield. 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature.* 428: 37-43.

Upchurch, R., C.-Y. Chiu, K. Everett, G. Dyszynski, D.C. Coleman, and W.B. Whitman. 2008. Differences in the composition and diversity of bacterial communities from agricultural and forest soils. *Soil Biol Biochem.* 40: 1294-1305.

Vasilyeva L.V. 1985. *Stella*, a New Genus of Soil Prostheco bacteria, with Proposals for *Stella humosa* sp. nov. and *Stella vacuolata* sp. nov. *Int J Syst Bacteriol* 35: 518-521.

Vasilyeva, L. V. and G.A. Zavarzin. 1995. Dissipotrophs in the microbial community. *Microbiology* (English translation of *Mikrobiologiya*) 64: 198–202.

Wawrik B., L. Kerkhof, J. Kukor, G. Zylstra. 2005. Effect of different carbon sources on community composition of bacterial enrichments from soil. *Appl Environ Microb.* 71: 6776-6783.

White P.J., J.A.W. Morgan, and G.D. Bending. 2005. Biological costs and benefits to plant–microbe interactions in the rhizosphere. *J Exp Bot.* 56: 1729-1739.

Woese C.R. 1987. Bacterial evolution. *Microbiol Rev.* 51: 221- 271.

Zhou J., I. M. E. Davey, J. B. Figueras, E. Rivkina, D. Gilichinsky and J. M. Tiedje. 1997. Phylogenetic diversity of a bacterial community determined from Siberian tundra soil DNA. *Microbiology.* 143: 3913-3919.

Appendix I

Temperature values of water samples at Estación Bosque and Estación Mina in Bosque del Pueblo, Adjuntas, PR.

Site	Temperature measures in water samples (°C) ^a				
	S ₁	S ₂	S ₃	Mean	SD
Estación Bosque					
February	19.71	19.89	19.87	19.82	.09
March	20.82	20.87	20.99	20.89	.08
April	21.99	21.9	22.23	22.04	.17
May	22.6	22.74	22.51	22.61	.12
June	23.01	23.12	23.16	23.09	.08
July	22.55	22.63	22.78	22.65	.11
August	22.02	22.07	22.11	22.06	.04
September	21.79	21.88	21.90	21.85	.05
October	20.35	20.28	20.32	20.31	.04
November	19.98	19.92	19.96	19.95	.08
Estación Mina					
February	22.18	22.6	22.3	22.36	.22
March	23.51	23.25	23.47	23.41	.14
April	23.10	23.02	23.09	23.07	.04
May	23.19	23.12	23.17	23.16	.04
June	23.53	23.41	23.50	23.48	.06
July	23.54	23.45	23.48	23.49	.04
August	22.65	22.60	22.55	22.60	.05
September	22.10	22.31	22.15	22.18	.11
October	22.09	22.13	22.17	22.13	.04
November	21.30	21.36	21.27	21.31	.04

^a Water samples were taken in triplicates.

pH values of water samples at Estación Bosque and Estación Mina in Bosque del Pueblo, Adjuntas, PR.

Site	pH measures in water samples ^a				
	S ₁	S ₂	S ₃	Mean	SD
Estación Bosque					
February	5.59	5.49	5.53	5.54	0.05
March	5.43	5.38	5.41	5.41	0.02
April	5.4	5.42	5.35	5.39	0.04
May	5.38	5.41	5.33	5.37	0.04
June	4.9	4.83	4.88	4.87	0.04
July	4.95	4.85	4.93	4.91	0.05
August	5.51	5.43	5.46	5.47	0.05
September	5.63	5.57	5.61	5.6	0.03
October	5.19	5.2	5.27	5.22	0.04
November	5.3	5.27	5.34	5.3	0.04
Estación Mina					
February	4.49	4.5	4.51	4.50	0.01
March	3.99	4.2	4.10	4.10	0.10
April	4.2	4.3	4.01	4.17	0.15
May	3.9	3.92	3.99	3.94	0.05
June	3.95	4.09	4.02	4.02	0.07
July	4.05	3.98	4.04	4.02	0.04
August	3.88	3.89	3.92	3.90	0.02
September	3.98	4.31	3.99	4.10	0.19
October	3.86	4.03	4.27	4.05	0.21
November	3.92	3.93	4.11	3.99	0.10

^a Water samples were taken in triplicates.

Conductivity (mS/cm) values of water samples at Estación Bosque and Estación Mina in Bosque del Pueblo, Adjuntas, PR.

Site	Conductivity measures in water samples (mS/cm) ^a				
	S ₁	S ₂	S ₃	Mean	SD
Estación Bosque					
February	0.049	0.041	0.047	0.045	0.004
March	0.054	0.042	0.047	0.047	0.006
April	0.065	0.063	0.055	0.061	0.005
May	0.051	0.052	0.064	0.055	0.007
June	0.040	0.057	0.059	0.052	0.010
July	0.041	0.062	0.053	0.052	0.010
August	0.042	0.060	0.051	0.051	0.009
September	0.063	0.071	0.059	0.064	0.006
October	0.055	0.056	0.068	0.059	0.007
November	0.047	0.049	0.062	0.052	0.008
Estación Mina					
February	0.399	0.432	0.419	0.416	0.016
March	0.39	0.399	0.399	0.396	0.005
April	0.34	0.353	0.35	0.347	0.006
May	0.46	0.385	0.409	0.418	0.038
June	0.415	0.453	0.469	0.445	0.027
July	0.388	0.39	0.399	0.392	0.005
August	0.376	0.353	0.452	0.393	0.051
September	0.38	0.415	0.391	0.395	0.017
October	0.35	0.408	0.369	0.375	0.029
November	0.403	0.469	0.45	0.440	0.033

^a Water samples were taken in triplicates.

Appendix II

**Nickel elemental analysis of water samples at Estación Bosque and Estación Mina in
Bosque del Pueblo, Adjuntas, PR.**

Site	Niquel concentration in water samples, Ni (ppm) ^a				
	S ₁	S ₂	S ₃	Mean	SD
Estación Bosque					
February	0.0143	0.0133	.0135	0.0138	0.0005
March	0.2030	0.1940	0.1990	0.0199	0.0005
April	0.1690	0.1706	0.1620	0.0167	0.0005
May	0.1860	0.1913	0.1846	0.0187	0.0004
June	0.1720	0.1646	0.1656	0.0167	0.0004
July	0.0956	0.1133	0.1046	0.0105	0.0008
August	0.1206	0.1216	0.1200	0.0121	0.0008
September	0.1110	0.1113	0.1163	0.0113	0.0003
October	0.0843	0.0983	0.1046	0.0096	0.0010
November	0.1073	0.1093	0.1013	0.0106	0.0004
Estación Mina					
February	0.2126	0.2220	0.2196	0.0218	0.0005
March	0.2660	0.2550	0.2550	0.0259	0.0006
April	0.2243	0.2253	0.2240	0.0225	0.0007
May	0.2493	0.2513	0.2513	0.0251	0.0001
June	0.2590	0.2653	0.2540	0.0259	0.0005
July	0.2560	0.2593	0.2586	0.0258	0.0002
August	0.2743	0.2803	0.2630	0.0273	0.0008
September	0.2313	0.2370	0.2440	0.0237	0.0006
October	0.2126	0.2110	0.2166	0.0213	0.0003
November	0.2426	0.2506	0.2430	0.0245	0.0005

^a Water samples were taken in triplicates.

**Copper elemental analysis of water samples at Estación Bosque and Estación Mina in
Bosque del Pueblo, Adjuntas, PR.**

Site	Copper concentration in water samples, Cu (ppm) ^a				
	S ₁	S ₂	S ₃	Mean	SD
Estación Bosque					
February	1.8040	1.7746	1.7840	1.7875	0.0152
March	1.8593	1.8646	1.8620	1.8619	0.0027
April	1.8600	1.8333	1.8333	1.8422	0.0154
May	2.1713	2.1293	2.1266	2.1424	0.0251
June	2.5726	2.6160	2.6040	2.5975	0.0224
July	2.3226	2.3013	2.3006	2.3082	0.0125
August	1.8660	1.8726	1.8666	1.8684	0.0037
September	1.8206	1.8126	1.8166	1.8166	0.0040
October	2.3226	2.3200	2.2966	2.3131	0.0143
November	1.6673	1.6660	1.6740	1.6691	0.0043
Estación Mina					
February	5.3200	5.3200	5.2967	5.3122	0.0134
March	6.1200	6.2200	6.1800	6.1733	0.0503
April	6.1467	6.1733	6.2200	6.1800	0.0371
May	5.7467	5.7867	5.8267	5.7867	0.0044
June	6.3200	6.3333	6.5333	6.3955	0.1195
July	6.6667	6.4800	6.7333	6.6267	0.1313
August	8.0400	7.9533	7.8800	7.9578	0.0800
September	7.0133	7.0266	6.8933	6.9777	0.0734
October	5.7933	5.6800	5.7866	5.7533	0.0635
November	5.3233	5.3500	5.3633	5.3455	0.0204

^a Water samples were taken in triplicates.

Appendix III

The closest relative of the 5'end region representative bacterial clone sequences (OTUs) from clone libraries at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Acidobacteria</i>	EBP402		Uncultured bacterium clone (EF516455.1)	Grassland soil from California	97	Cruz-Martinez et al. 2009.
<i>Acidobacteria</i>	EMS414 (EMP424, EMP426)		Uncultured bacterium clone (EF492955.1)	Soil around iron-manganese nodule from China	95	He et al. 2008.
<i>Acidobacteria</i>	EMS109		Uncultured bacterium clone (AM749754.1)	Steam-affected geothermal soil from Tikitere, New Zealand	96	Stott et al. 2008.
<i>Acidobacteria</i>	EBP140 (EBP133)		Uncultured bacterium clone (AY326548.1)	Forest soil from western Amazon	98	Kim et al. 2007.
<i>Acidobacteria</i>	EMS415		Uncultured bacterium clone (EF018396.1)	Rhizosphere soil from ...	98	Lesaulnier et al. 2008.
<i>Acidobacteria</i>	EBS80		Uncultured bacterium clone (FJ475507.1)	Soil from pinus sylvestris forest	84	Yarwood et al. 2009.
<i>Acidobacteria</i>	EBS101		Uncultured bacterium clone (GU015877.1)	Coniferous forest soils in central Taiwan	98	Lin et al. unpublished.
<i>Acidobacteria</i>	EBS87		Uncultured bacterium clone (EU335381.1)	Soil agreggate from BaC	98	Hansel et al. 2008.
<i>Acidobacteria</i>	EBP162		Uncultured bacterium clone (EU150198.1)	Soil from Niwot Ridge LTER, Colorado USA	96	Jones et al. unpublished.
<i>Acidobacteria</i>	EMS146 (EMS2)		Uncultured bacterium clone (EF018237.1)	Rhizosphere soil from ...	96	Lesaulnier et al. 2008.
<i>Acidobacteria</i>	EMS112		Uncultured bacterium clone (EU680451.1)	Forest soil of Gaoligong Mountains, Yunnan, China	96	Zhou et al. unpublished.
<i>Acidobacteria</i>	EBP9		Uncultured bacterium clone (DQ984559.1)	Tropical forest soil from Taiwan	89	Tseng et al. unpublished.
<i>Acidobacteria</i>	EMP207		Uncultured bacterium clone (DQ984559.1)	Tropical forest soil from Taiwan	95	Tseng et al. unpublished.
<i>Acidobacteria</i>	EBP114 (EBS89,EBS90)		Uncultured bacterium clone (EU335375.1)	Soil agreggate from BaC	97	Hansel et al. 2008.
<i>Acidobacteria</i>	EBP412		Uncultured bacterium clone (DQ829645.1)	Soil from Donana National Park, Spain	98	Zimmermann et al. unpublished.

Continued. The closest relative of the 5'end region representative bacterial clone sequences (OTUs) from clone libraries at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Acidobacteria</i>	EBP139		Uncultured bacterium clone (AY326563.1)	Forest soil from western Amazon	98	Kim et al. 2007.
<i>Acidobacteria</i>	EBP125		Uncultured bacterium clone (EU881256.1)	Grazing and burned annual, field soil in Karst region from Guangxi, China	97	Chen et al. unpublished.
<i>Acidobacteria</i>	EMP55 (EMS6,EMS113)		Uncultured bacterium clone (AY963457.1)	Forest soil from south- west China	99	Chan et al. 2006.
<i>Acidobacteria</i>	EMS139		Uncultured bacterium clone (Y11632.1)	Zinc-polluted soil from Canada	97	Brim et al. 1999.
<i>Acidobacteria</i>	EBP111		Uncultured bacterium clone (EU881269.1)	Grazing and burned annual, field soil in Karst region from Guangxi, China	99	Chen et al. unpublished.
<i>Acidobacteria</i>	EBS94		Uncultured bacterium clone (GQ402669.1)	Tropical peat swamp forest from Thailand	97	Kanokratana et al. unpublished.
<i>Alphaproteobacteria</i>	EBP107 (EBP144)		Uncultured bacterium clone (FJ418920.1)	<i>Bradyrhizobium sp. str-PC28</i>	99	Bogino and Giordano unpublished.
<i>Alphaproteobacteria</i>	EBP150		Uncultured bacterium clone (DQ404891.1)	Nitrate and heavy metal contaminated soils	98	Abulencia et al. 2006.
<i>Alphaproteobacteria</i>	EBP124 (EBS76,EBS103, EBS125,EBS413, EBP156,EMP418)		Uncultured bacterium clone (EU881284.1)	Grazing and burned annual, field soil in Karst region from Guangxi, China	99	Chen et al. unpublished.
<i>Alphaproteobacteria</i>	EBP14		Uncultured bacterium clone (AM162444.1)	Sphagnum peat bog from West Siberia, Russia:	97	Dedysch et al. 2006
<i>Alphaproteobacteria</i>	EMS417		Uncultured bacterium clone (EU979069.1)	Rhizosphere of faba bean (<i>Vicia Faba</i>)	92	Zhang et al. unpublished
<i>Alphaproteobacteria</i>	EBS91		Uncultured bacterium clone (EU335285.1)	Soil agreggate from BaC	97	Hansel et al. 2008.
<i>Alphaproteobacteria</i>	EBS106		Uncultured bacterium clone (FJ178158.1)	Upland red soil under different land uses	94	He et al. unpublished.

Continued. The closest relative of the 5' end region representative bacterial clone sequences (OTUs) from clone libraries at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Alphaproteobacteria</i>	EBP169 (EBP129)		Uncultured bacterium clone (AB250621.1)	<i>Rhodomicrobium vannielii</i> str- TUT3402	90	Hisada et al. 2007.
<i>Alphaproteobacteria</i>	EMP438		Uncultured bacterium clone (EF018202.1)	Rhizosphere soil from ...	96	Lesaulnier et al. 2008.
<i>Alphaproteobacteria</i>	EBP105		Uncultured bacterium clone (EF494369.1)	Nunnock River granitic landscape from Australia	95	Macalady et al. unpublished.
<i>Betaproteobacteria</i>	EMP411		Acidovorax sp. str- R-24667 (AM084010.1)	Cultivation of denitrifying bacteria	98	Heylen et al. 2006.
<i>Betaproteobacteria</i>	EBP126		Uncultured bacterium clone (AM284970.1)	<i>Burkholderia nodosa</i> str- Br3461	97	Chen et al. 2008.
<i>Deltaproteobacteria</i>	EBP414		Uncultured bacterium clone (GU016086.1)	Secondary coniferous forest soils in central Taiwan	99	Lin et al. unpublished.
<i>Deltaproteobacteria</i>	EBP104 (EBP101)		Uncultured bacterium clone (EU134345.1)	Tallgrass prairie soil from Kessler farm Oklahoma, USA	97	Elshahed et al. 2008.
<i>Deltaproteobacteria</i>	EBS74		Uncultured bacterium clone (EU335298.1)	Soil agreggate from BaC	87	Hansel et al. 2008.
<i>Deltaproteobacteria</i>	EBP163		Uncultured bacterium clone (AM941477.1)	Soil from an acidic fen from Germany	96	Kuesel et al. unpublished.
<i>Gammaproteobacteria</i>	EMS106		Legionella waltersii strain 2074-AUS-E (AF122886.2)	Isolated from water in France	93	Lo Presti et al. 2001.
<i>Gammaproteobacteria</i>	EMP208 (EMP205, EMP210)		Uncultured bacterium clone (DQ223199.1)	Subsurface water from Au mines of the Kalahari Shield, South Africa	98	Gihring et al. 2008.
<i>Gammaproteobacteria</i>	EMS134		Uncultured bacterium clone (EF019160.1)	Rhizosphere soil	95	Lesaulnier et al. 2008.
<i>Actinobacteria</i>	EBP147		Uncultured bacterium clone (FM206309.1)	Rhizosphere soil of Pinus radiata from Rotorua, New Zealand	97	Lottmann et al. unpublished.
<i>Actinobacteria</i>	EBP143 (EBP142)		Actinomadura livida str- IMSNU 22191T (AJ293706.1)	Genomic DNA	90	Lee and Hah unpublished.

Continued. The closest relative of the 5' end region representative bacterial clone sequences (OTUs) from clone libraries at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Actinobacteria</i>	EBP145		Thermomonosporaceae bacterium PENDO-1810 (GQ924504.1)	Endophytic actinomycetes from tropical plants, Papua New Guinea	90	Janso and Carter. 2010.
<i>Actinobacteria</i>	EBS116		Microbispora sp. 212030 (FJ261983.1)	Mangrove soil from China	95	Xu and Hong. unpublished.
<i>Actinobacteria</i>	EBS175		Uncultured bacterium clone (EU080982.1)	Soil from Rajasthan, India	91	Purohit et al. unpublished.
<i>Firmicutes</i>	EBS100		<i>Paenibacillus terrigena</i> str- TDSAS2-38 (GQ284528.1)	Mangrove sediment from India	96	Bhattacharya and Chakrabarti unpublished.
<i>Firmicutes</i>	EMS108		<i>Paenibacillus terrigena</i> str- TDSAS2-38 (GQ284528.1)	Mangrove sediment from India	89	Bhattacharya and Chakrabarti unpublished.
<i>Firmicutes</i>	EMS107		Uncultured bacterium clone (GU172181.1)	Eucalyptus plantation soil from China	98	Duan unpublished.
<i>Firmicutes</i>	EMP435		Uncultured bacterium clone (EU044321.1)	Inland dune fields at Georgia, USA	97	Tarlera et al. 2008.
<i>Firmicutes</i>	EMP427 (EMP410)		Uncultured bacterium clone (EF018887.1)	Rhizosphere soil from ...	97	Lesaulnier et al. 2008.
<i>Firmicutes</i>	EMS13		Uncultured bacterium clone (GU016168.1)	Coniferous forest soils in central Taiwan	99	Lin et al. unpublished.
<i>Bacteroidetes</i>	EMS405		Uncultured bacterium clone (EF075839.1)	Forest soil from Georgia, USA	97	Jangid et al. 2008.
<i>Planctomycetes</i>	EBP119 (EBP405, EBP418)		Uncultured bacterium clone (EF663133.1)	Cropland soil on GASP KBS- LTER sampling site from Michigan, USA	96	Jangid et al. unpublished.
<i>Planctomycetes</i>	EMS145		Uncultured bacterium clone (EF074464.1)	Pasture soil on GASP Watkinsville sampling site from Georgia, USA	97	Jangid et al. 2008.
<i>Verrucomicrobia</i>	EBS85		Uncultured bacterium clone (EF665522.1)	Forest at the GASP KBS- LTER sampling from Michigan, USA	94	Jangid et al. unpublished.

Continued. The closest relative of the 5'end region representative bacterial clone sequences (OTUs) from clone libraries at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Verrucomicrobia</i>	EBS120		Uncultured bacterium clone (AY963387.1)	Forest soil from south- west China	97	Chan et al. 2006.
<i>Verrucomicrobia</i>	EBS107 (EBS170)		Uncultured bacterium clone (EF018903.1)	Rhizosphere soil	97	Lesaulnier et al. 2008.
<i>Verrucomicrobia</i>	EBP109		Uncultured bacterium clone (GQ402703.1)	Tropical peat swamp forest from Thailand	98	Kanokratana et al. unpublished.
<i>Verrucomicrobia</i>	EBS102		Uncultured bacterium clone (EU335328.1)	Soil agreggate from BaC	91	Hansel et al. 2008.
<i>Verrucomicrobia</i>	EBS108 (EBS92,EBS98, EBS112)		Uncultured bacterium clone (AB240266.1)	Rhizosphere soil	95	Nakamura et al. 2005.
<i>Verrucomicrobia</i>	EBP417 (EBS172)		Uncultured bacterium clone (EF492980.1)	Soil around iron-manganese nodule from China	96	He et al. 2008.
<i>Cyanobacteria</i>	EMS141 (EMS88)		Uncultured bacterium clone (FJ592879.1)	Soil from Socompa Volcano, Andes	96	Costello et al. 2009.
<i>Cyanobacteria</i>	EBP123		Uncultured bacterium clone (EF516266.1)	Grassland soil from California	96	Cruz-Martinez et al. 2009.
<i>Cyanobacteria</i>	EMS110		Cylindrospermum sp.A1345 (DQ897365.1)	Genomic DNA	92	Ganesan and Anand. Unpublished.
<i>Elusimicrobia</i>	EBP120		Uncultured bacterium clone (EU044559.1)	Inland dune fields at Georgia, USA	94	Tarlera et al. 2008.
<i>Nitrospira</i>	EMS80		Uncultured bacterium clone (AY326514.1)	Forest soil from western Amazon	99	Kim et al. 2007.
<i>Chloroflexi</i>	EBP403		Uncultured bacterium clone (DQ129211.2)	Coweeta forest soil from North Carolina-georgia border, USA	88	Upchurch et al. 2008.
<i>Chloroflexi</i>	EMS449		Uncultured bacterium clone (AY917525.1)	Volcanic deposit from 1790, Hawaii,USA	88	Gomez-Alvarez and Nuesslein. unpublished.
<i>Chloroflexi</i>	EBS79		Uncultured bacterium clone (EU335420.1)	Soil agreggate from BaC	87	Hansel et al. 2008.

Continued. The closest relative of the 5'end region representative bacterial clone sequences (OTUs) from clone libraries at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Chloroflexi</i>	EMP406		Uncultured bacterium clone (GU205279.1)	Sediment from Venezuelan Tepui (Orthoquartzite) Cave	94	Giarrizzo et al. unpublished.
<i>Chloroflexi</i>	EMS85		Uncultured bacterium clone (AY425781.1)	Volcanic deposit from 1921, Hawaii,USA	94	Gomez-Alvarez et al. 2007.
<i>Chloroflexi</i>	EMP442		Uncultured bacterium clone (AY425781.1)	Volcanic deposit from 1921, Hawaii,USA	94	Gomez-Alvarez et al. 2007.
<i>Chloroflexi</i>	EMP407		Uncultured bacterium clone (EU680443.1)	Forest soil of Gaoligong Mountains from Yunnan, China	92	Zhou et al. unpublished.
<i>Chloroflexi</i>	EMS12 (EMS135, EMS140, EMS402)		Uncultured bacterium clone (AY917586.1)	Volcanic deposit from 1894, Hawaii,USA	91	Gomez-Alvarez and Nuesslein. unpublished.
<i>Chloroflexi</i>	EMP434 (EMP425, EMP429, EMP444)		Uncultured bacterium clone (EU335423.1)	Soil agreggate from Oak Ridge, TN, USA	95	Hansel et al. 2008.
<i>Chloroflexi</i>	EMP421 (EMP422, EMP441, EMP439)		Uncultured bacterium clone (EU335420.1)	Soil agreggate from Oak Ridge, TN, USA	96	Hansel et al. 2008.
<i>Chloroflexi</i>	EMS111		Uncultured bacterium clone (HM469587.1)	Soil from uranium mine surrounding area, India	90	Dhal et al. unpublished.
<i>Chloroflexi</i>	EMS115 (EMS117)		Uncultured bacterium clone (HM439333.1)	Open cast uranium mine environment from Banduhurang, India	87	Islam et al. unpublished.
<i>Chloroflexi</i>	EMS7 (EMS144)		Uncultured bacterium clone (AY917962.1)	Volcanic deposit from 1982, Hawaii,USA	91	Gomez-Alvarez and Nuesslein. unpublished.
<i>Chloroflexi</i>	EMP415		AY646712.1	Forest soil from Brasil	88	Val-Moraes et al. unpublished.

Continued. The closest relative of the 5'end region representative bacterial clone sequences (OTUs) from clone libraries at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Chloroflexi</i>	EMP7		(EU651820.1)	Forest soils in central Taiwan	91	Lin et al. 2010.
<i>Chloroflexi</i>	EMP11		Bacterium SOSP1-142 (AM180162.1)	Genomic DNA	91	Cavaletti et al. 2006.
<i>Chloroflexi</i>	EMS133		Uncultured bacterium clone (AY425771.1)	Volcanic deposit from 1790, Hawaii,USA	90	Gomez-Alvarez et al. 2008.
<i>Chloroflexi</i>	EMP413		Uncultured bacterium clone (AY917334.1)	Volcanic deposit from 1790, Hawaii,USA	93	Gomez-Alvarez and Nusslein. 2007
Unclassified	EBP166		Uncultured bacterium clone (EU881239.1)	Grazing and burned annual, field soil in Karst region from Guangxi, China	98	Chen et al. unpublished.
Unclassified	EBS104		Uncultured bacterium clone (GQ402688.1)	Tropical peat swamp forest from Thailand	98	Kanokratana et al. unpublished.
Unclassified	EBP406		Uncultured bacterium clone (DQ984572.1)	Tropical forest soil from Taiwan	97	Tseng et al. unpublished.
Unclassified	EBP401		Uncultured bacterium clone (FJ466267.1)	1959 Kilauea volcanic deposit, Hawaii	98	Weber and King. 2010.
Unclassified	EMS143		Uncultured bacterium clone (GU016302.1)	Secondary coniferous forest soils in central Taiwan	98	Lin et al. Unpublished
Unclassified	EBP113		Uncultured bacterium clone (EU445222.1)	Tropical forest topsoil from Kuala Lumpur, Malaysia	96	Pang et al. Unpublished
Unclassified	EBP404		Uncultured bacterium clone (AY724006.2)	Forest soil from Cologne, Germany	98	Rosch and Bothe. 2005.
Unclassified	EMP401		Uncultured bacterium clone (GU598869.1)	Forest soil from Hubbard Brook, New Hampshire, USA	97	Ganapathi et al. unpublished.

The closest relative sequences of the 3'end region representative bacterial clone sequences from soil samples at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Acidobacteria</i>	EBP146 (EBP170,EMS137, EMS136)		Uncultured bacterium clone (DQ829638.1)	Soil from Donana National Park (Spain)	99	Zimmermann et al. Unpublished.
<i>Acidobacteria</i>	EBP108		Uncultured bacterium clone (AY963506.1)	Forest soil from south-west China	99	Chan et al.2006.
<i>Acidobacteria</i>	EBP115 (EBP413,EMS120)		Uncultured bacterium clone (AY963457.1)	Forest soil from south-west China	98	Chan et al.2006.
<i>Acidobacteria</i>	EBP409 (EBS78, EBP122, EBP408)		Uncultured bacterium clone (EF516884.1)	Grassland soil from California	98	Cruz et al. 2009.
<i>Acidobacteria</i>	EBP110 (EBS95,EBS174)		Uncultured bacterium clone (EF515877.1)	Grassland soil from California	99	Cruz et al. 2009.
<i>Acidobacteria</i>	EMP417		Uncultured bacterium clone (EF018458.1)	Soil rhizosphere (elevated)	99	Lesaulnier et al. 2008
<i>Acidobacteria</i>	EMS409		Uncultured bacterium clone (AB238786.1)	Peat soil in shallow subsurface of Sarobetsu Mire	96	Akiyama et al. unpublished.
<i>Acidobacteria</i>	EMP404		Uncultured bacterium clone (EF516877.1)	Grassland soil from California	95	Cruz et al. 2009.
<i>Acidobacteria</i>	EBS86		Uncultured bacterium clone (AY913380.1)	Forest soil from Cologne, German	95	Roesch et al. unpublished.
<i>Acidobacteria</i>	EBP416 (EBP161,EBP152, EMP414)		Uncultured bacterium clone (DQ984566.1)	Tropical forest soil from Taiwan	99	Tseng et al. unpublished.
<i>Acidobacteria</i>	EBS77		Uncultured bacterium clone (EU881239.1)	Soil in Karst region from China	96	Chen et al. unpublished.
<i>Acidobacteria</i>	EBP136 (EBP112)		Uncultured bacterium clone (AY913233.1)	Forest soil from Cologne, German	98	Roesch et al. unpublished.
<i>Acidobacteria</i>	EMP440		Uncultured bacterium clone (AY913350.1)	Forest soil from Cologne, German	97	Roesch et al unpublished.

Continued. The closest relative sequences of the 3'end region representative bacterial clone sequences from soil samples at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Acidobacteria</i>	EBP151 (EBP153)		Uncultured bacterium clone (DQ984565.1)	Tropical forest soil from Taiwan	99	Tseng et al. unpublished.
<i>Acidobacteria</i>	EBS71		Uncultured bacterium clone (AY963306.1)	Forest soil from south-west China	96	Chan et al.2006.
<i>Acidobacteria</i>	EMS5		Uncultured bacterium clone (AY963306.1)	Forest soil from south-west China	97	Chan et al.2006.
<i>Acidobacteria</i>	EBP164 (EBP141)		Uncultured bacterium clone (EF516094.1)	Grassland soil from California	99	Cruz et al. 2009.
<i>Acidobacteria</i>	EMS418 (EMS419)		Uncultured bacterium clone (EF494344.1)	Nunnock River granitic landscape from Australia	97	Green et al. 2007.
<i>Acidobacteria</i>	EBS81		Uncultured bacterium clone (AB179518.1)	Iron oxide nodules in sedimentary rock from Toyohashi, Japan	96	Yoshida et al. unpublished.
<i>Acidobacteria</i>	EBS105		Uncultured bacterium clone (DQ451501.1)	Fushan forest soil from Taiwan	99	Tsai et al. 2009.
<i>Acidobacteria</i>	EBP411		Uncultured bacterium clone (AY963462.1)	Forest soil from south-west China	99	Chan et al.2006.
<i>Acidobacteria</i>	EBS126		<i>Solibacter usitatus</i> str-Ellin6076 (CP000473.1)		98	Copeland et al. unpublished
<i>Acidobacteria</i>	EBS97		Uncultured bacterium clone (EF516102.1)	Grassland soil from California	97	Cruz et al. 2009
<i>Acidobacteria</i>	EBS93		Uncultured bacterium clone (FJ870384.1)	Loa Alamos National Laboratory NM	93	Eichorst et al. unpublished.
<i>Acidobacteria</i>	EMS412 (EMS407)		Uncultured bacterium clone (FJ465998.1)	1959 Kilauea volcanic deposit from Hawaii	95	Weber et al. 2010.
<i>Acidobacteria</i>	EBS117		Uncultured bacterium clone (AY326550.1)	Forest soil from western Amazon	92	Kim et al. 2007.
<i>Alphaproteobacteria</i>	EBP157 (EBP155)		Uncultured bacterium clone (DQ453803.1)	Red soil from China	99	Huang et al. unpublished.
<i>Alphaproteobacteria</i>	EBP407		Uncultured bacterium clone (DQ451522.1)	Fushan forest soil from Taiwan	99	Tsai et al. 2009.

Continued. The closest relative sequences of the 3'end region representative bacterial clone sequences from soil samples at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Alphaproteobacteria</i>	EBP167 (EBS72, BP128)		Uncultured bacterium clone (GQ402595.1)	Tropical peat swamp forest from Thailand	99	Kanokratana et al. unpublished.
<i>Alphaproteobacteria</i>	EBS83		Uncultured bacterium clone (GQ402595.1)	Tropical peat swamp forest from Thailand	93	Kanokratana et al. unpublished.
<i>Alphaproteobacteria</i>	EBS99		Uncultured bacterium clone (GQ402595.1)	Tropical peat swamp forest from Thailand	95	Kanokratana et al. unpublished.
<i>Alphaproteobacteria</i>	EMS450		Uncultured bacterium clone (AY913376.1)	Forest soil from Cologne, German	92	Roesch et al. unpublished.
<i>Alphaproteobacteria</i>	EMP420		Uncultured bacterium clone (AF465655.1)	Geothermally disturbed Soils from Yellow Stone National Park	97	Norris et al. 2002.
<i>Alphaproteobacteria</i>	EMS403 (EMS406)		Uncultured bacterium clone (AY963489.1)	Forest soil from south- West China	98	Chan et al.2006.
<i>Alphaproteobacteria</i>	EBP148		Uncultured bacterium clone (EF520409.1)	acid-impacted lakes from Adirondack lake New York, USA	97	Percent et al. 2008.
<i>Alphaproteobacteria</i>	EBP165		Uncultured bacterium clone (AY395444.1)	Pasture soil from Australia	98	Schoenborn et al. Unpublished.
<i>Deltaproteobacteria</i>	EBP116		Uncultured bacterium clone (AB240359.1)	Rhizosphere of <i>Phragmites</i>	97	Nakamura et al. 2005.
<i>Deltaproteobacteria</i>	EMS1		Uncultured bacterium clone (EF516730.1)	Grassland soil from California	94	Cruz et al. 2009.
<i>Deltaproteobacteria</i>	EBS115 (EBS84)		Uncultured bacterium clone (AY326616.1)	Forest soil from western Amazon	92	Kim et al. 2007.
<i>Deltaproteobacteria</i>	EMP62		Uncultured bacterium clone (EU202851.1)	Agricultural soil from Mexico	96	Ceja-Navarro et al. unpublished.
<i>Deltaproteobacteria</i>	EBS73		Uncultured bacterium clone (EF516111.1)	Grassland soil from California	96	Cruz et al. 2009.
<i>Deltaproteobacteria</i>	EMP437		Uncultured bacterium clone (EU335336.1)	Soil agreggate from Melton Branch Watershed, OakRidge, TN,USA	92	Hansel et al. 2008.

Continued. The closest relative sequences of the 3'end region representative bacterial clone sequences from soil samples at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Gammaproteobacteria</i>	EMP203 (EMP204,EMP206)		Uncultured bacterium clone (DQ223199.1)	subsurface water of the Kalahari Shield, South Africa	97	Gihring et al. 2006.
<i>Gammaproteobacteria</i>	EBP154		Uncultured bacterium clone (AJ536870.1)	Uranium mining waste pile, soil sample from Saxony, Germany	98	Satchanska et al. unpublished.
<i>Gammaproteobacteria</i>	EBP102 (EBP127,EBP158)		Uncultured bacterium clone (DQ984542.1)	Tropical forest soil from Taiwan	99	Tseng et al. unpublished.
<i>Gammaproteobacteria</i>	EBS173		Uncultured bacterium clone (DQ984567.1)	Tropical forest soil from Taiwan	94	Tseng et al. unpublished.
<i>Gammaproteobacteria</i>	EMS138		Uncultured bacterium clone (AY359282.1EF613012.1)	AMD at Yinshan mine from China	93	Yin et al. 2008.
<i>Actinobacteria</i>	EMS411		Uncultured bacterium clone (EU044086.1)	Inland dune fields at Georgia, USA	98	Tarlera et al. 2008.
<i>Firmicutes</i>	EBS82		<i>Propionispira arboris</i> str-12B4 (NR_029357.1)		91	Stackebrandt et al. 1999.
<i>Bacteroidetes</i>	EBS171		Uncultured bacterium clone (EF667468.1)	Soil sample of growing radish area from Jaunpur, India	89	Shivaji et al. unpublished.
<i>Planctomycetes</i>	EMP57		Uncultured bacterium clone (AY963397.1)	Forest soil from south- west China	97	Chan et al. 2006.
<i>Verrucomicrobia</i>	EBS123		Uncultured bacterium clone (DQ451528.1)	Fushan forest soil from Taiwan	96	Tsai et al. unpublished.
<i>Verrucomicrobia</i>	EBS111 (EBS70)		Uncultured bacterium clone (EF516170.1)	Grassland soil from California	99	Cruz et al. 2009.
<i>Chloroflexi</i>	EBS121		Uncultured bacterium clone (EU134183.1)	Tallgrass prairie soil from Kessler farm Oklahoma, USA	92	Elshahed et al. 2008.
<i>Chloroflexi</i>	EMS116		Uncultured bacterium clone (AY425774.1)	Volcanic deposit from 1790, Hawaii,USA	96	Gomez-Alvarez et al. 2003.
<i>Chloroflexi</i>	EMS65		Uncultured bacterium clone (AY425792.1)	Volcanic deposit from 1959, Hawaii,USA	95	Gomez-Alvarez et al. 2007.

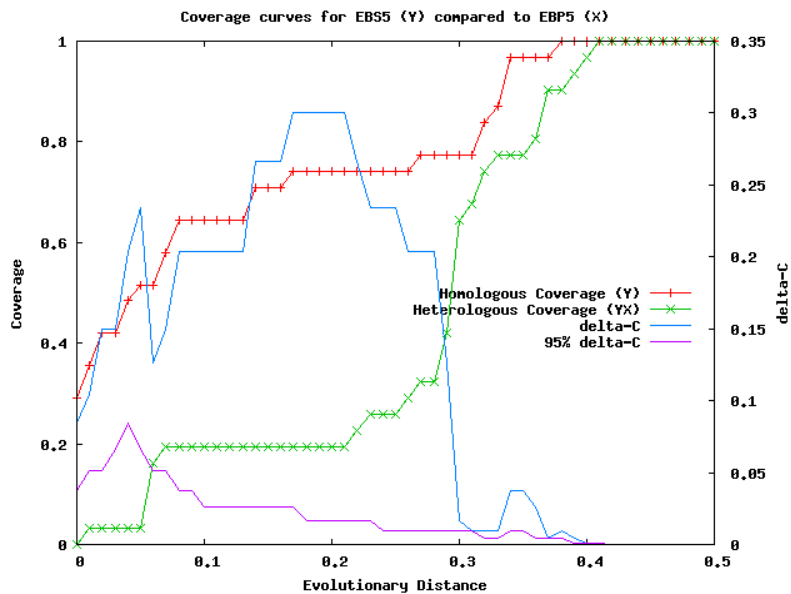
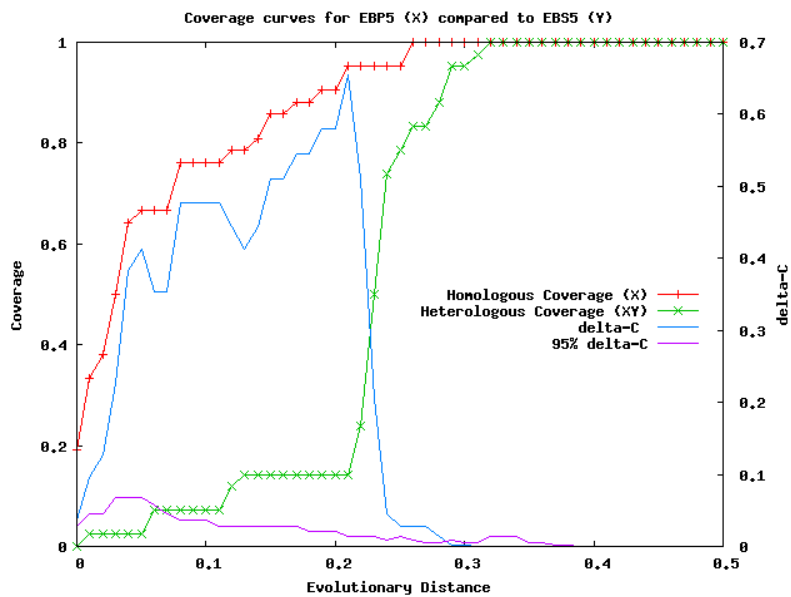
Continued. The closest relative sequences of the 3'end region representative bacterial clone sequences from soil samples at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
<i>Chloroflexi</i>	EMS451		Uncultured bacterium clone (EU335420.1)	Soil agreggate from BaC	96	Hansel et al. 2008.
<i>Chloroflexi</i>	EMP412 (EMP416)		Uncultured bacterium clone (EU043839.1)	Inland dune fields at Georgia, USA	91	Tarlera et al. 2008.
<i>Chloroflexi</i>	EMS413 (EMS3,EMS142, EMS147,EMS408)		Uncultured bacterium clone (AY689909.1)	AMD effluent stream sediment	96	Baeseman et al. unpublished.
<i>Chloroflexi</i>	EBS 119		Uncultured bacterium clone (DQ128926.2)	Coweeta forest soil from North Carolina-Georgia border, USA	91	Upchurch et al. 2008.
<i>Chloroflexi</i>	EMP405		Uncultured bacterium clone (AY690146.1)	Naturally acidic mountain stream sediment (pH 3.99)	94	Baeseman et al. Unpublished.
<i>Chloroflexi</i>	EMP445 (EMP443,EMP446)		Uncultured bacterium clone (EU335444.1)	Soil agreggate from BaC	98	Hansel et al. 2008.
<i>Chloroflexi</i>	EMP447		Uncultured bacterium clone (EU335444.1)	Soil agreggate from BaC	96	Hansel et al. 2008.
<i>Chloroflexi</i>	EMP448		Uncultured bacterium clone (DQ129114.2)	Coweeta forest soil from North Carolina-Georgia border, USA	94	Upchurch et al. 2008.
<i>Chloroflexi</i>	EMP430		Uncultured bacterium clone (EU335444.1)	Soil agreggate from BaC	93	Hansel et al. 2008.
<i>Chloroflexi</i>	EMP409		Uncultured bacterium clone (EU881288.1)	Grazing and burned annual, field soil in Karst region from Guangxi, China	98	Chen et al. 2005.
Unclassified	EMS131		Uncultured bacterium clone (AY917636.1)	Volcanic deposit from 1921, Hawaii,USA	95	Gomez-Alvarez et al. unpublished.
Unclassified	EMS91, EMS8		Uncultured bacterium clone (AY917636.1)	Volcanic deposit from 1921, Hawaii,USA	98	Gomez-Alvarez et al. unpublished.
Unclassified	EMS118		Uncultured bacterium clone (DQ660863.1)	AMD in DongChuan copper pyrites from China	96	Dai and Liu. Unpublished.

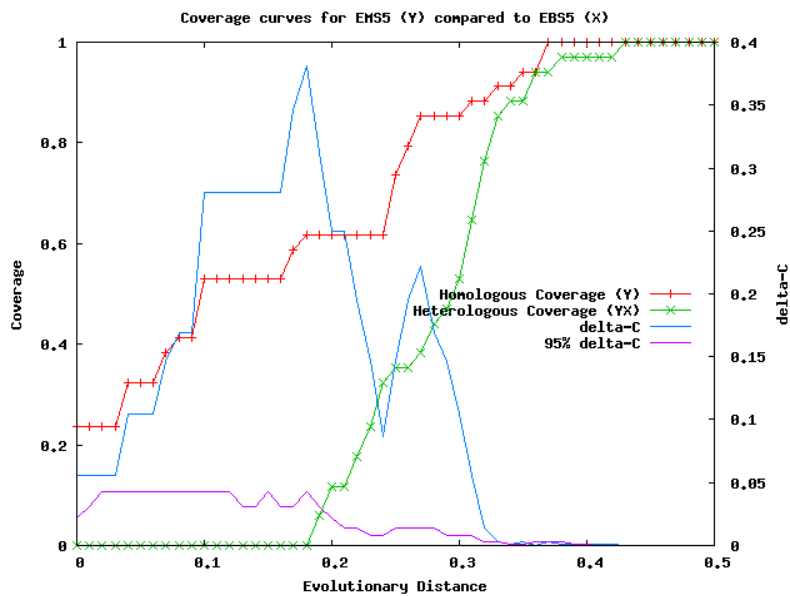
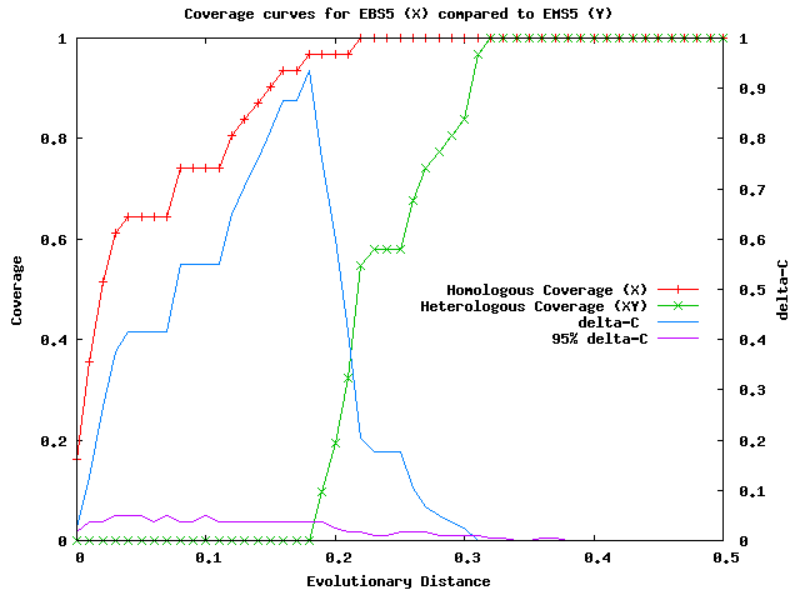
Continued. The closest relative sequences of the 3'end region representative bacterial clone sequences from soil samples at Bosque del Pueblo, Adjuntas PR

Phylogenetic group	OTU, (clone sequences within OTU)	Accession No.	Closesst identified relative in GenBank	Source or physiology of the closest relative	% Sequence identity	Reference
Unclassified	EBS114		Uncultured bacterium clone (AY661981.1)	Groundwater contaminated with high levels of nitric acid-bearing uranium waste from USA	97	Fields et al. unpublished.

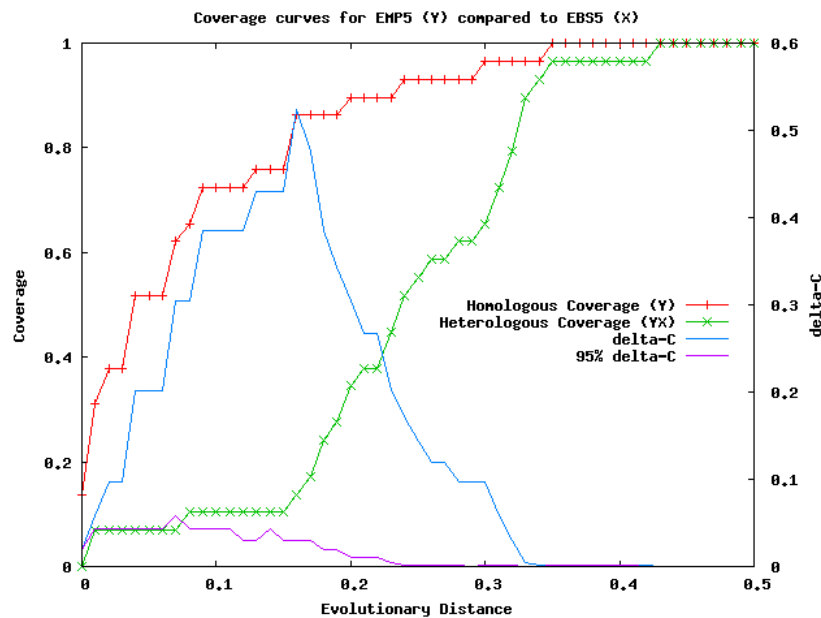
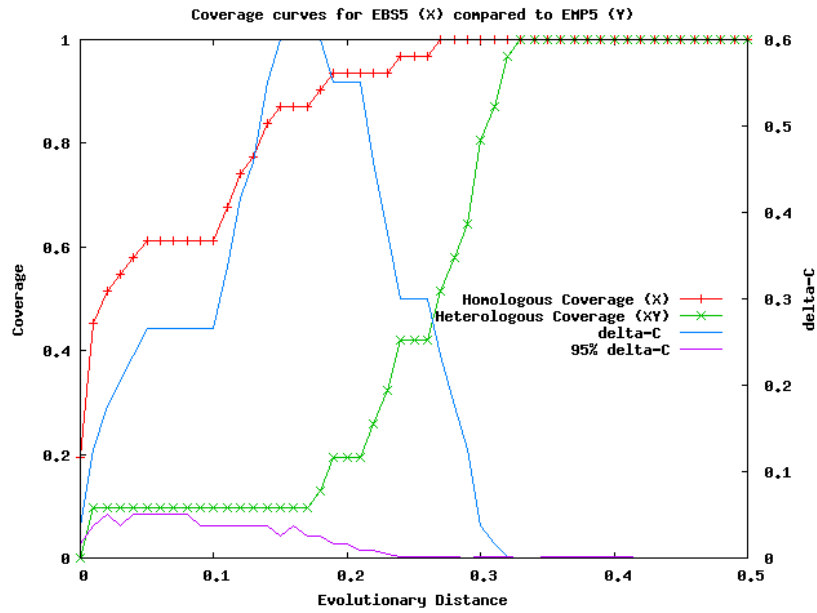
Appendix IV



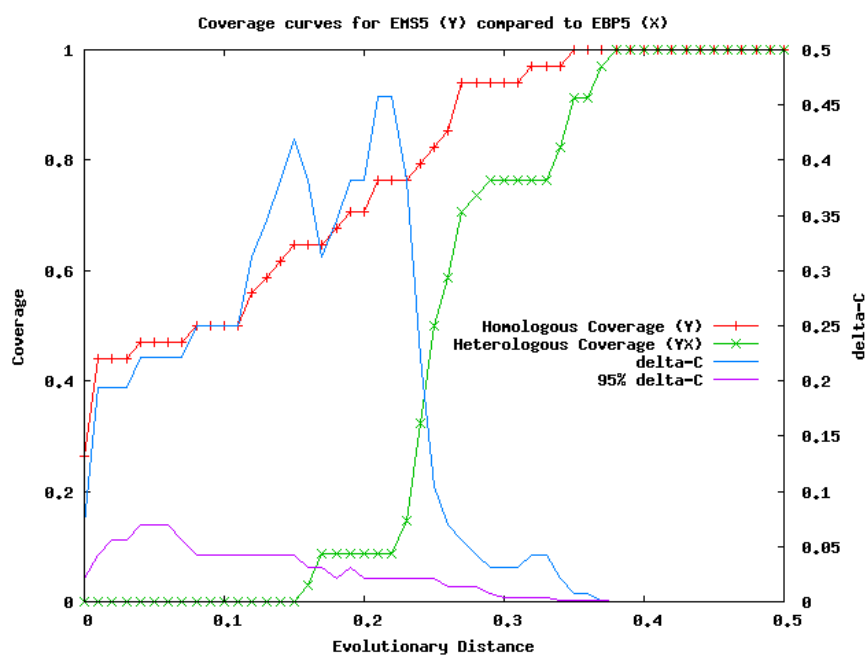
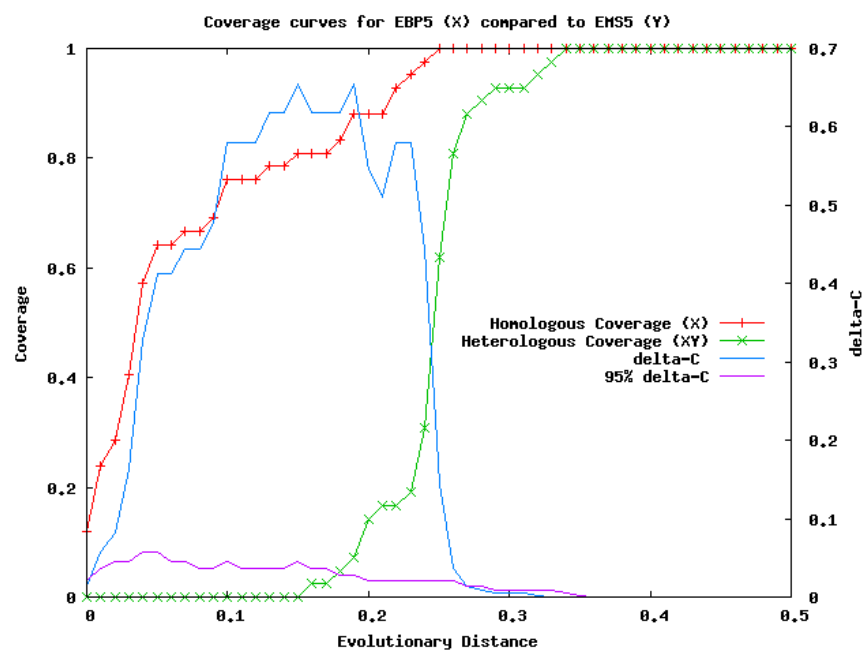
LIBSHUFF analysis of 16S rRNA gene sequences from EBP5 clone library compared to EBS5 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



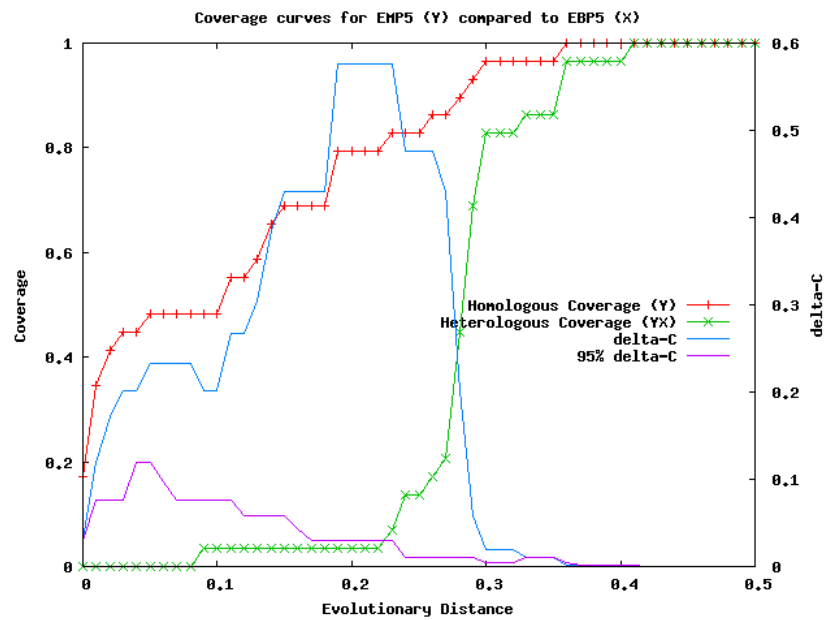
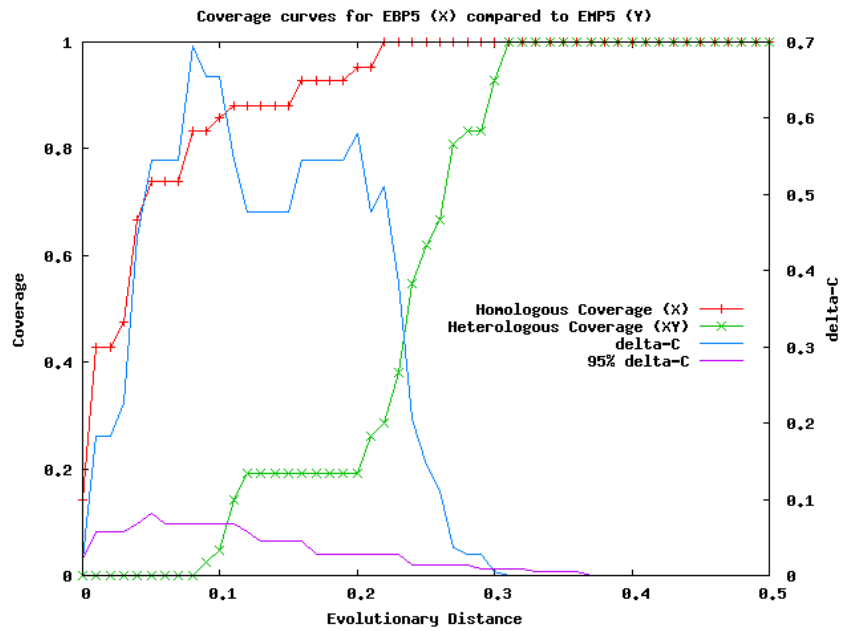
LIBSHUFF analysis of the 5'end portion-data-set of 16S rRNA gene sequences from EBS5 clone library compared to EMS5 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



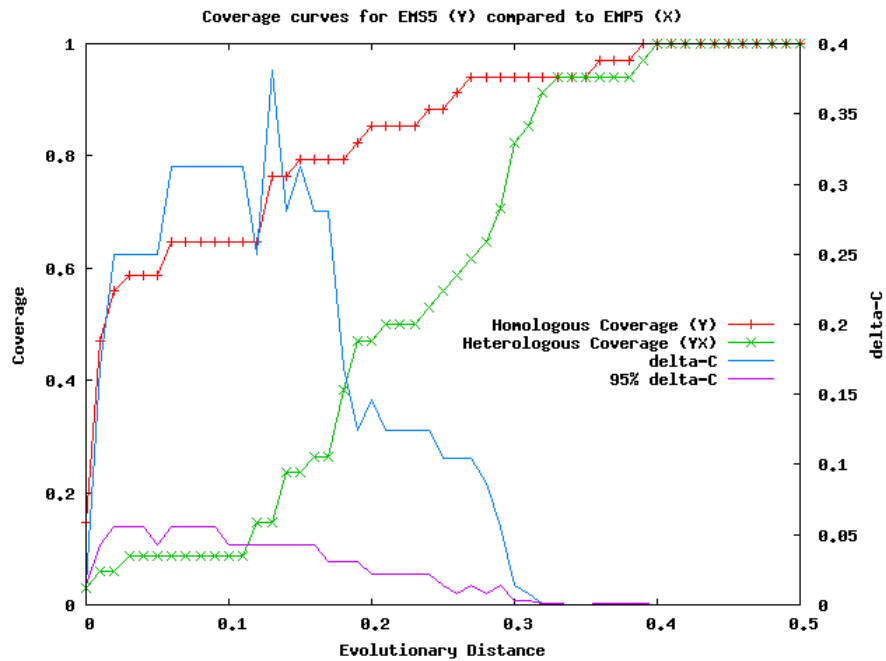
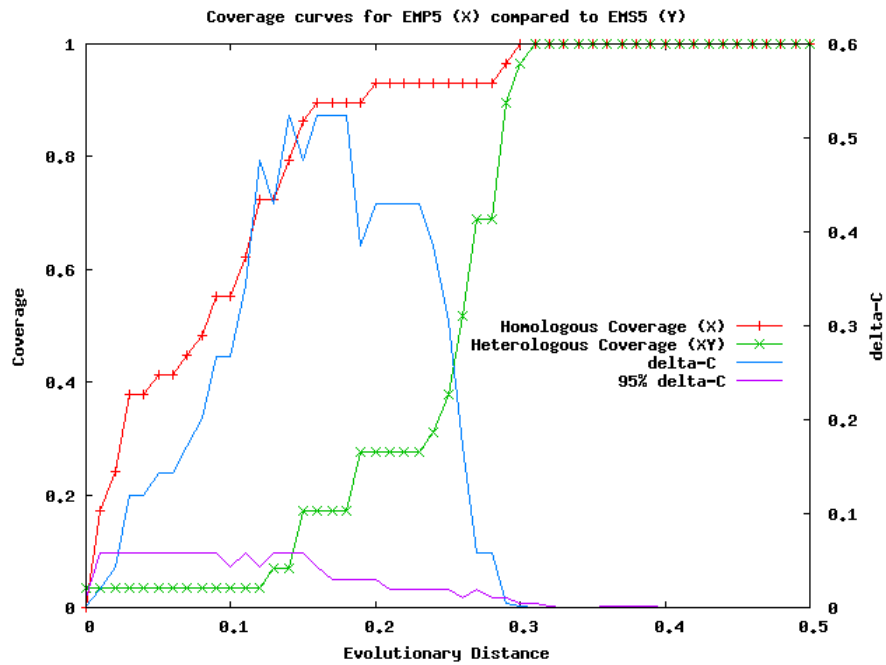
LIBSHUFF analysis of the 5'end portion-data-set of 16S rRNA gene sequences from EBS5 clone library compared to EMP5 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



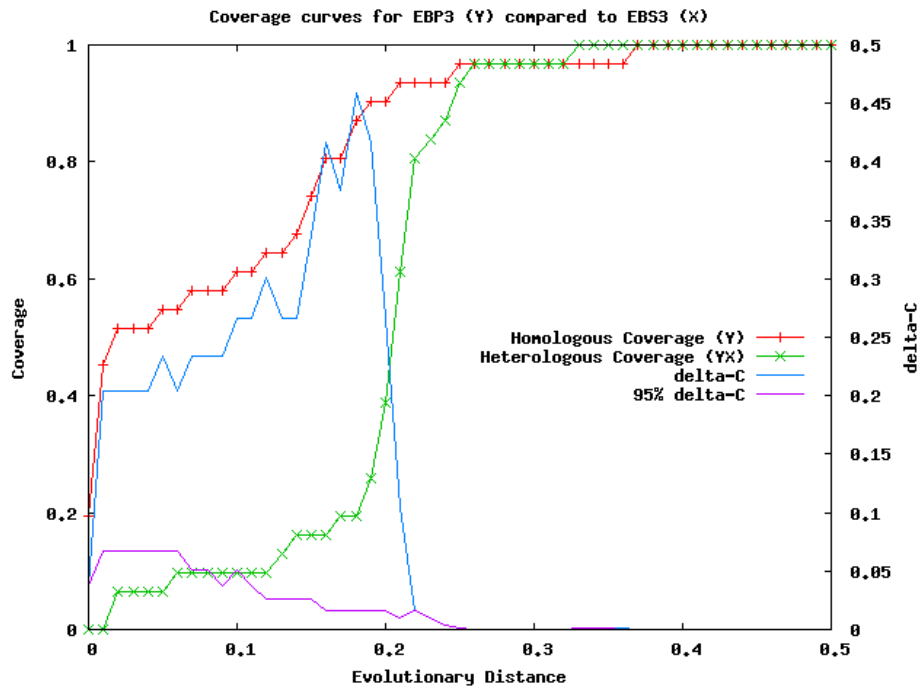
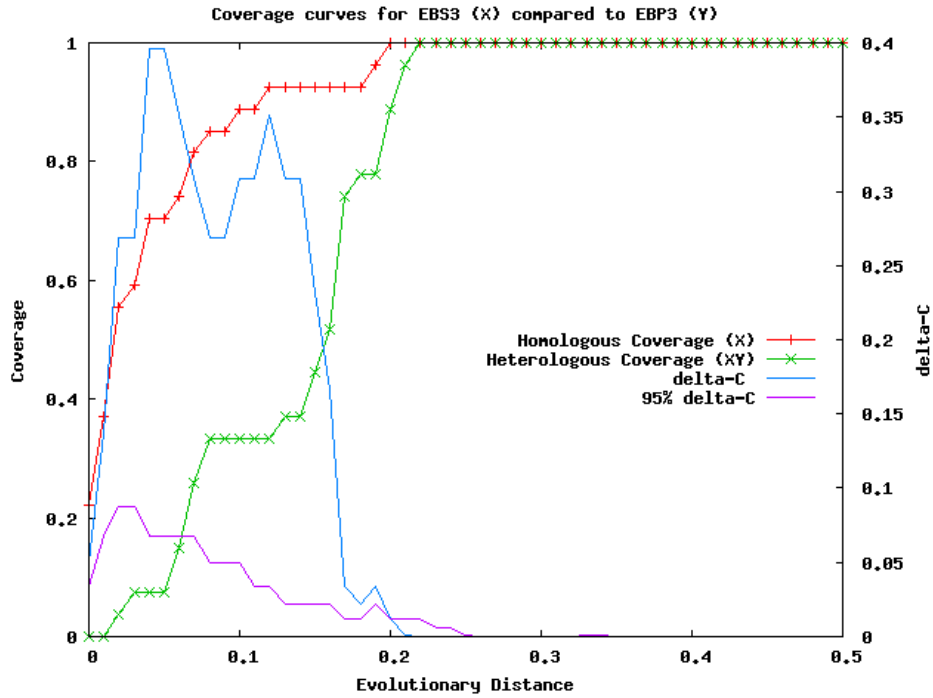
LIBSHUFF analysis of the 5'end portion-data-set of 16S rRNA gene sequences from EBP5 clone library compared to EMS5 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



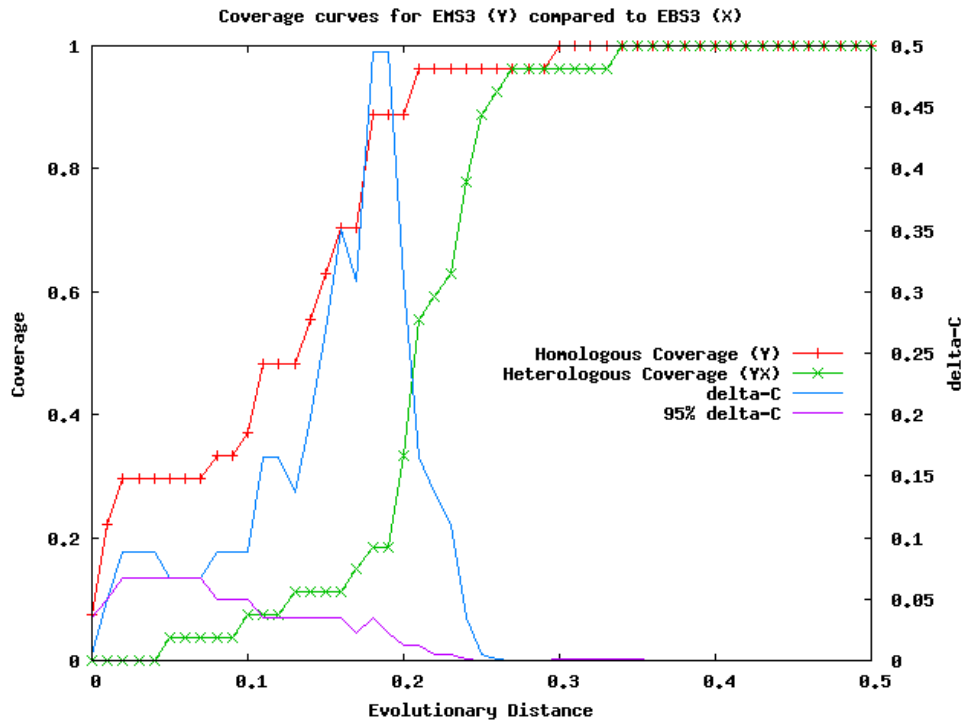
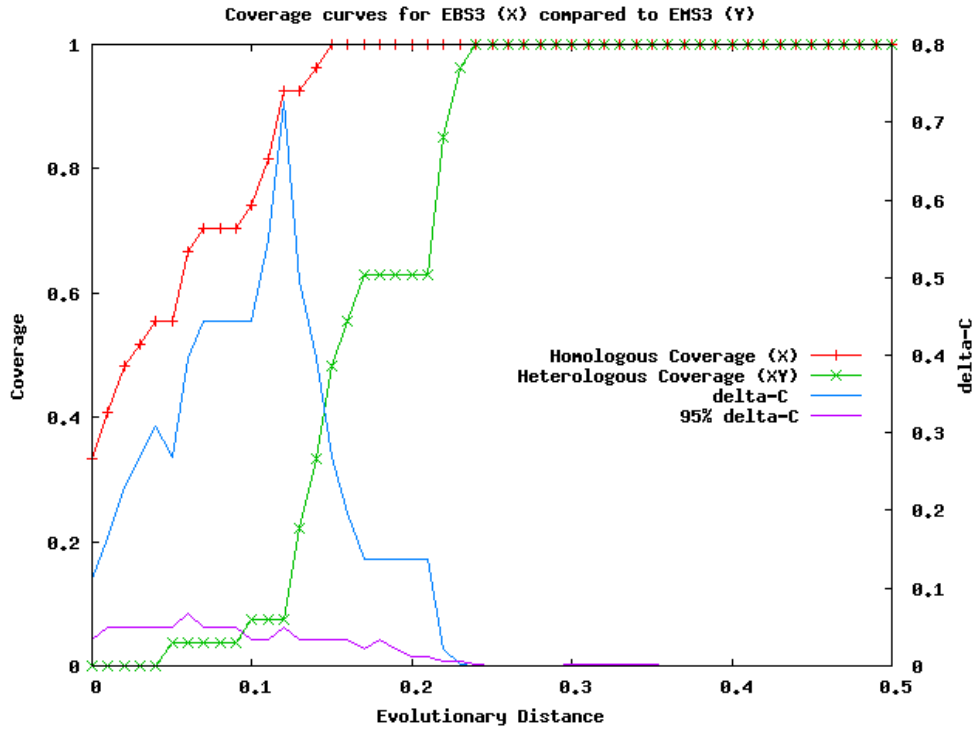
LIBSHUFF analysis of the 5'end portion-data-set of 16S rRNA gene sequences from EBP5 clone library compared to EMP5 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



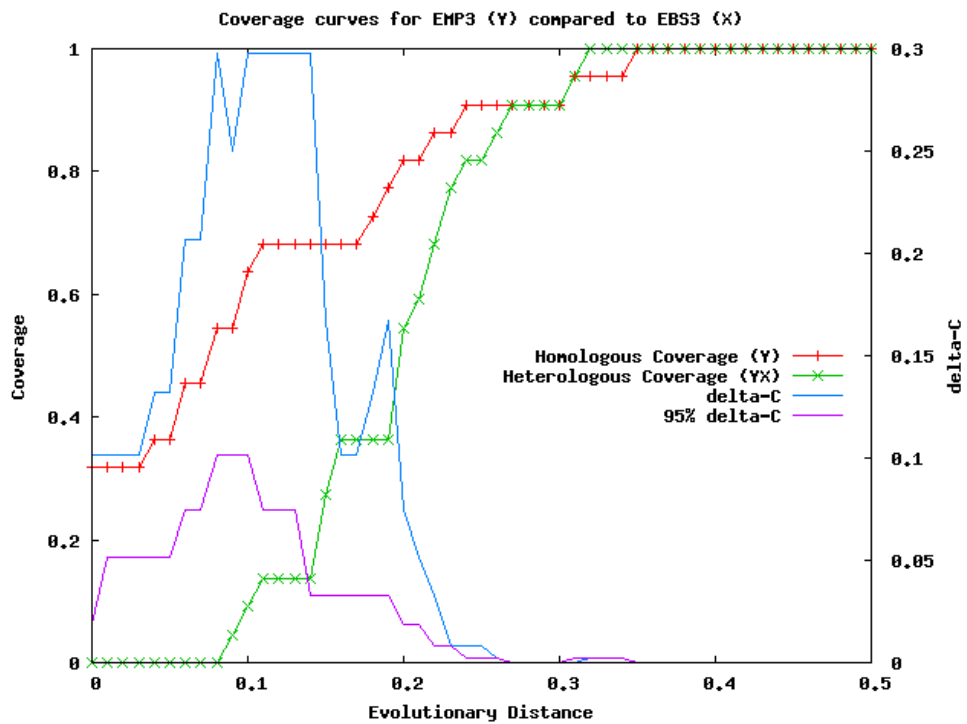
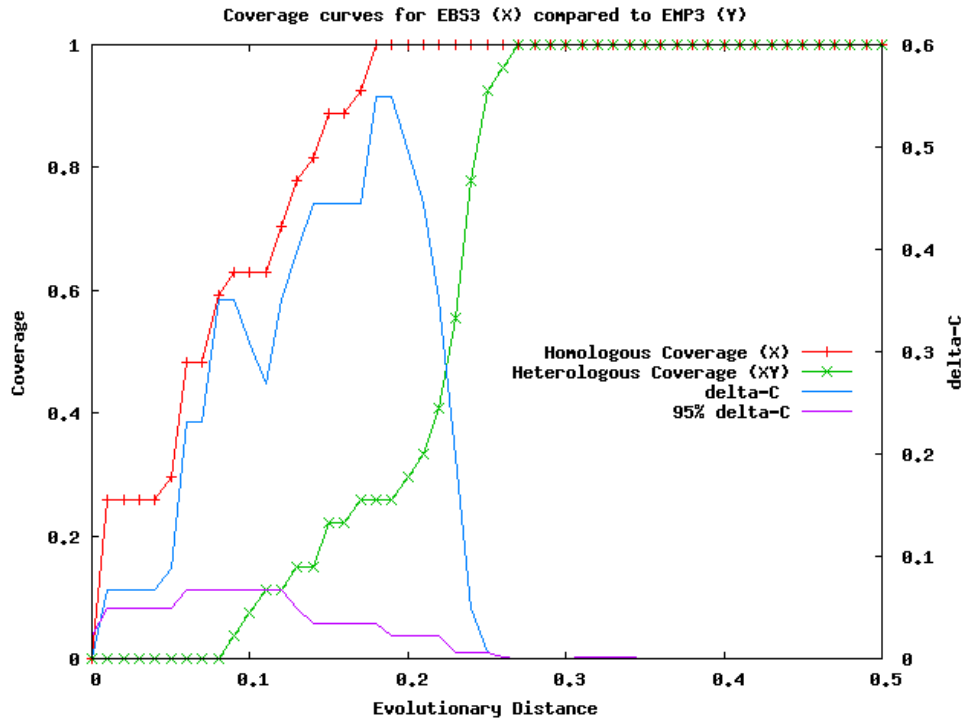
LIBSHUFF analysis of the 5'end portion-data-set of 16S rRNA gene sequences from EMP5 clone library compared to EMS5 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



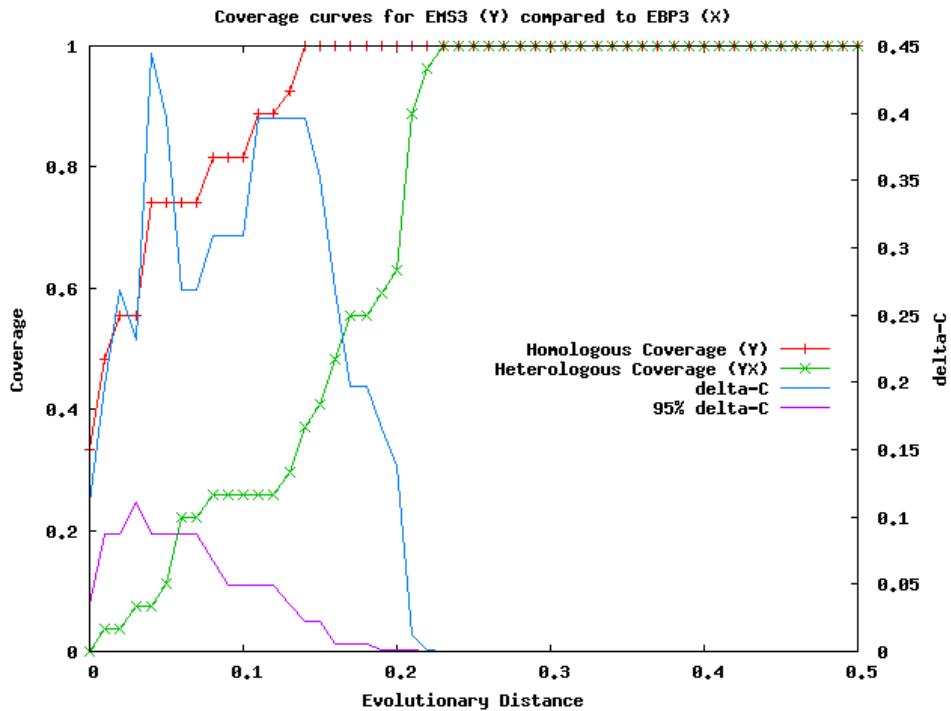
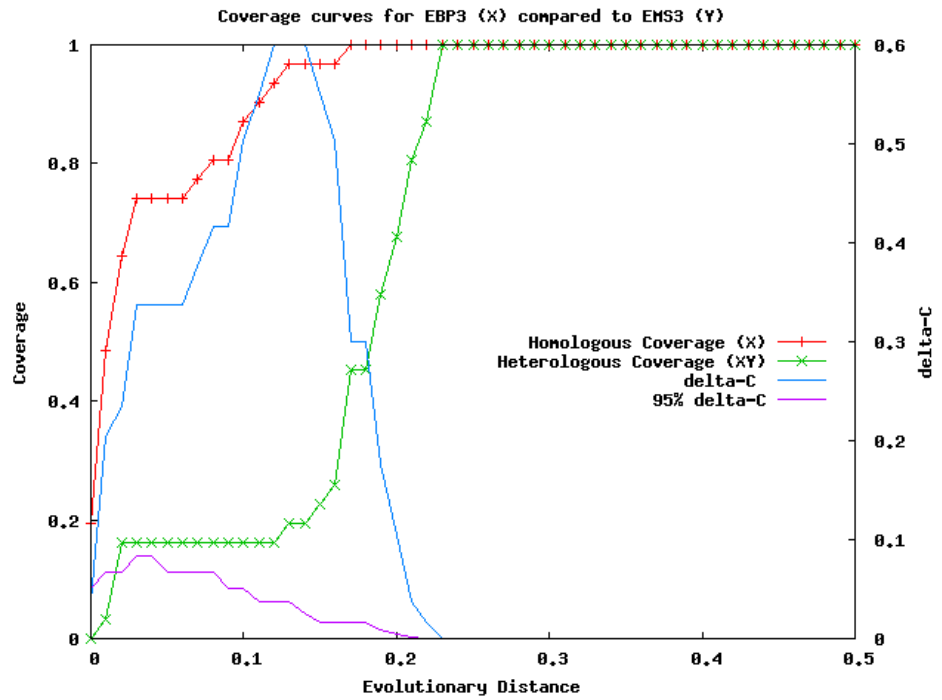
LIBSHUFF analysis of the 3'end portion-data-set of 16S rRNA gene sequences from EBS3 clone library compared to EBP3 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



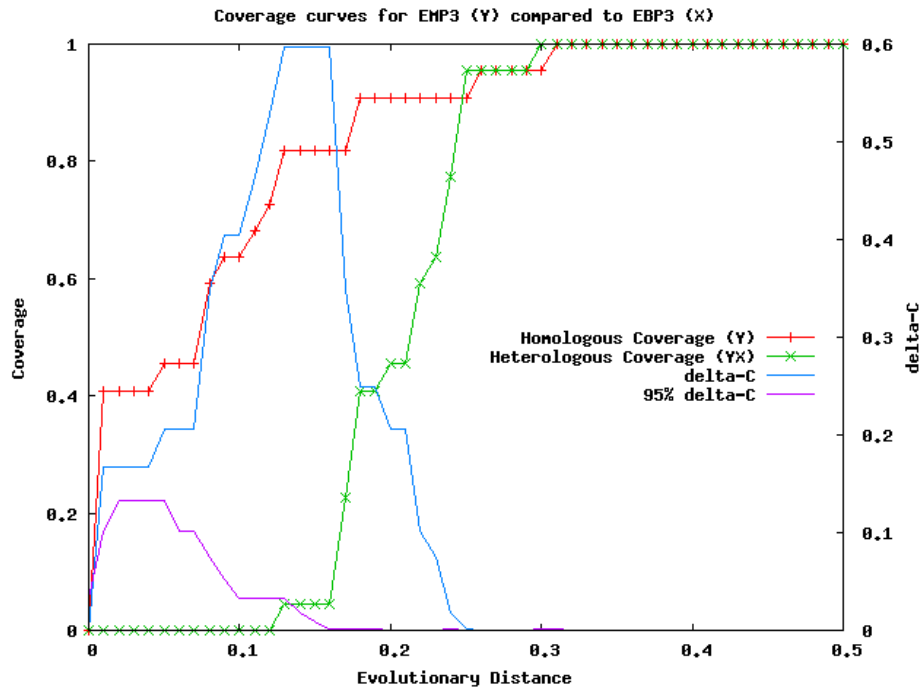
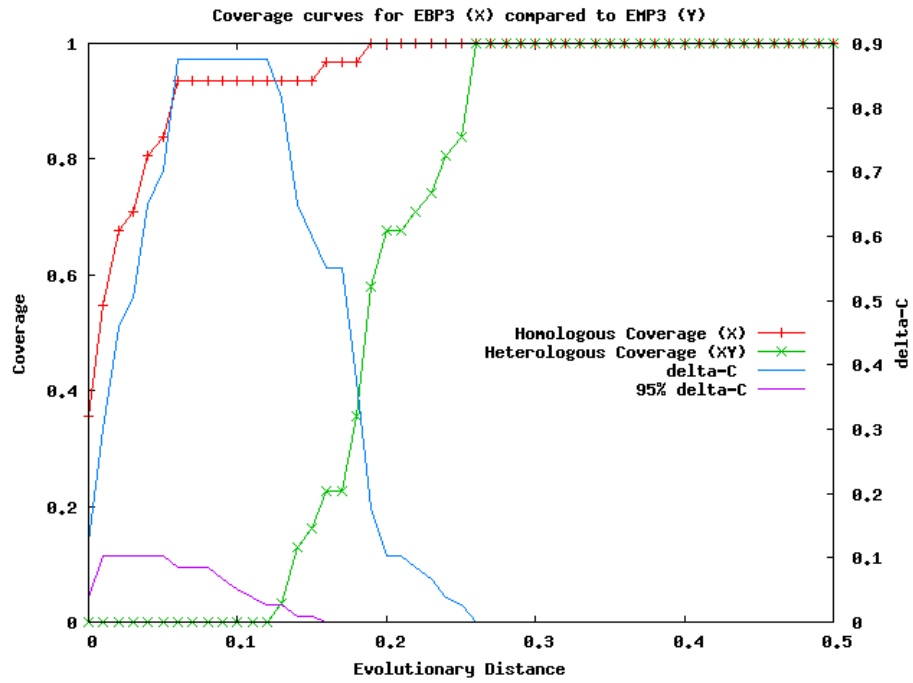
LIBSHUFF analysis of the 3'end portion-data-set of 16S rRNA gene sequences from EBS3 clone library compared to EMS3 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



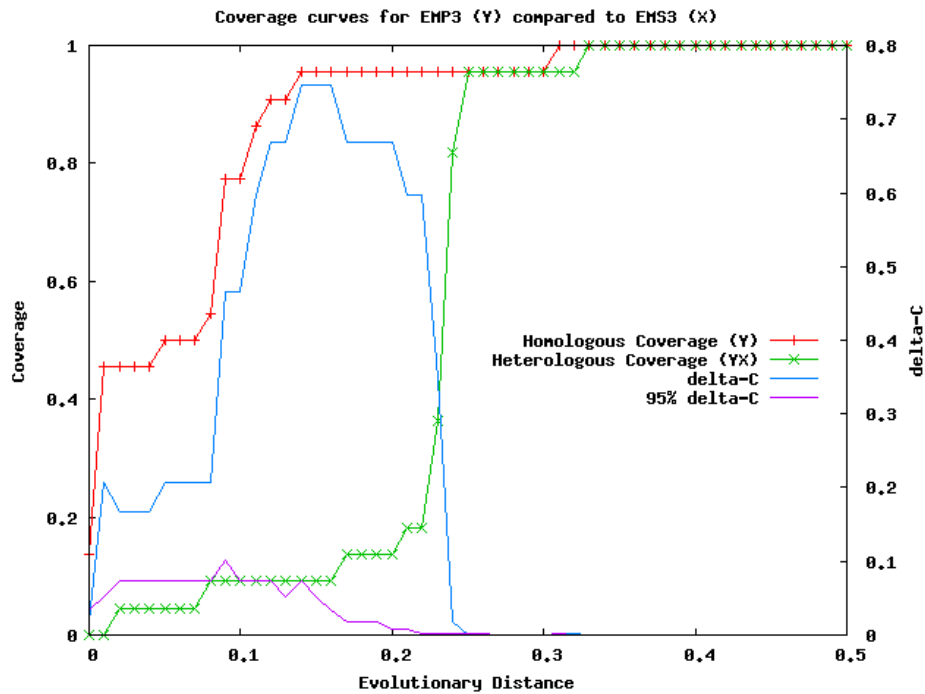
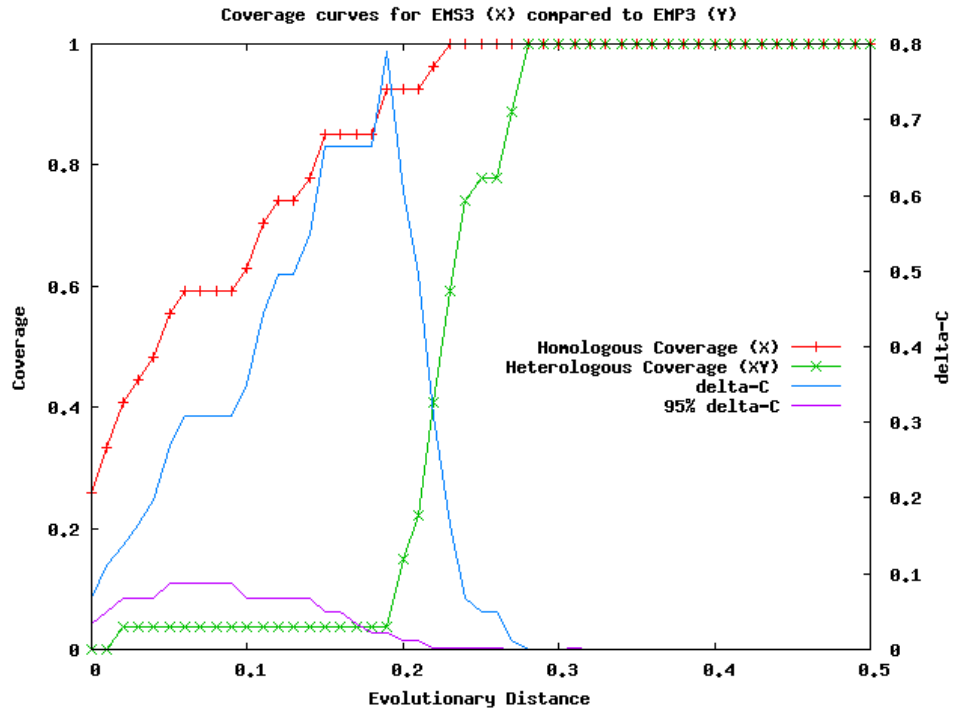
LIBSHUFF analysis of the 3' end portion-data-set of 16S rRNA gene sequences from EBS3 clone library compared to EMP3 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



LIBSHUFF analysis of the 3' end portion-data-set of 16S rRNA gene sequences from EMS3 clone library compared to EBP3 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



LIBSHUFF analysis of the 3' end portion-data-set of 16S rRNA gene sequences from EBP3 clone library compared to EMP3 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).



LIBSHUFF analysis of the 3' end portion-data-set of 16S rRNA gene sequences from EMS3 clone library compared to EMP3 clone library showing differences in composition of the bacterial communities (y-axis) over a range of evolutionary distances (D) (x-axis).

Appendix V

DNA Clone Sequences from Clone Libraries

Clone sequences recovered from the 5' end-portion-data-set:

>ebs9

agagtttgatcatggctcagaatgaacgctggcggcgtgcttaacacatgcaagtcgctgcgcagttccgtagcaatacgggatgggcgc
acggcgcacgggtgcgtaaacacgtgggcaatccgccctctgatggggaataaccccgcaaaagcggggctaataccgcataacattcctt
ggactacgggtccggggattcaaaagcagtgatgcgtcggaggaggagcccgcgccgattagctagttggtgaggtaatggctcaccaag
gcaatgatcggtagctggtctgagaggatgatcagccacactggaactgagacacgggtccagactcctacgggaggcagcagtggggaa
tcttgacaatggcgaaagcctgatgcagcgacgccgcgtgggggatgaagcttctcggagtgtaaaccccttcgactgggacgaatg
cccgcaaggaggatgacggtaccagtagaagaagccccggctaactacgtgccagcagccgcgtaagacgtagggggccagcgttgtt
cggaaactactgggtgtaaagggttcgtagcgggtggaagtcgggagtgaaatctttgggcttaaccagaggctgctccgaaactgct
gtgctcagtggtgggagagggcgcgtggaattgtaggtgtagcgggtgaaatgcgtagatatctgcaggaacaccggtggcgaaagcggcg
cgctggaccactactga

>ebs74

agagtttgatcctggctcagaacgaacgctggcggcgtgcctaacacatgcaagtcaggaggaaagccgcaaggtgagtaaacggcg
aacgggtgcgtaaacacgtggataatctacctccacactgggataactggccgaaaggcttgtaataaccagataagccacggggcttc
ggacctgagggaaaagggtgctctgtatacaagctatcgaggagatgagtcggccatcagctagttggcggggtaaaagcc
caccaaggctacgacgggtagctggtctgagaggatgatcagccacactggaactgagacacgggtccagactcctacgggaggcagca
gtggggaatcttgacaatggacgaaagtctgatgcagcaacgccgcgtgggggatgaggcttctcggagtgtaaaccccttcgaccg
gaagaatgcccgaagggtatgacgggtacgggtataagaagccccggctaactacgtgccagcagccgcgtaatacgtagggggcca
gctgtgctcgaactactggcgtaaagggtccgtagggcgggtgtagcaagtcgatagtgaatctctgcgcttaacgcagaggctgctaac
gaaactgctgtgc

>ebs76

agagtttgatcatggctcagagcgaacgctggcggcaggcctaacacatgcaagtcgagcggcgtagcaatacggagcggcagacggg
agagtaaacacgtgggaacgcgcccttcagttcgaacaactcagggaaactgggctaataccgaatacgtccgtaaggagaaagattat
cgctgaaggatcggcccgcgtctgattagctagttggtgggtaatggcccaccaaggcgacgatcagtagctggtctgagaggatgatc
agcctcactgggactgagacacggcccagactcctacgggaggcagcagttggggaatattggacaatgggcgcaagcctgatccagcc
atgccgcgtgggtgatgaaggccctaggggtgtaaagcccttcggcggggaagataatgacgggtaccgcagaagaagccccggctaa
ctctgtgccagcagccgcgtaatacgaagggggctagcgttctcggaaactcactgggcgtaaaagcgcacgtaggcggcttttaagtcag
gggtgaaatcctggagctcaactcagaactcccttgatactgagaggcttgatccgggagaggtgagtggaactgcgagtgtagagg
tgaatcgtagatattcgcaagaacaccagtggcgaaggcggctcactggcccggactgacgctgaggtgcgaaagcgtggggagca
aacaggat

>ebs79

agagtttgatcatggctcaggatgaacgctggcggcgtgcctaagacatgcaagtcgaggggggttagacccccggcgaacgggtgagta
acacgtgggaatttcattaagttagggatctctcccgaaggagagcaatactcagtggtcgcaagactaaagtcgcaagacgctta
atgcaaggcccgcggccatcagctagttggtagggtaatggcctaccaaggcgatgacgggtagggtagcgagagcttacaccccc
cactgggactgcgacacggcccagactcctacgggaggcagcagtttggaattttggacaatgggcgaaagcctgatccagcagccg
cgtgcacgaagaaggccttcgggtgtaaagtcttttggtgggaaaaattttgattgtaccacagaataagaggttgtaactctgtcca
gcagcagcggtaatacagagacctcaagcgttatccggaatcattgggcgtaaaagcgtaccgataggtggtcatgatagttttgggtaaaag
ccagaagctcaactctgcgtttgctcaaaatacttcagactaggggcaagagggaagctggaacgaacgggtgtagtagtgaatgctg
ttgatcgttcggaaacacgaatcgtaggcaggtcttgg

>ebs80

agagtttgatcctggctcaggatgaacgctggcggcgtgcctaatacatgcaagtcgagcggtaaggtcccttcggggatacacgagcgg
cggacggctgagtaaacacgtaggaacgtacccacagtgaggcataaagcaccagaaatggtgtctaataccgcatgtgctcttcggagtaa
agcttcggcgtatgggaacggcctacggatgattagcttgggtgaggtaatagctaccaaggcgaagatcgggtatccggcctgagag
ggcgacggacacactggaactgaaacacggctccagactcctacgggaggcagcagtgggggaatfttcgcaatgggggaaacctga
cgcagcaacgccgctggagatgaagcccttggggcgtaaactccttcgaccaagacgataatggcggfactggtggaagaagcac
cggctaactctgtgccagcagccgcggtataacgaagggtgcaaacgttgcggaattattggcgtaaagcgcacgtagcggcctgg
caagtcggatgtgaaatccctcggctcaaccaaggaagtgcacccgaaactgctgagcttgagacttaagaggatcgcggaattcccggt
gtagagtgaaattcgtagatcgggagggacaccagtgccgaaggcggcgtctgggaagatactgacgctgaggtgcgaaagcgt
ggggagcaaacaggat

>ebs85

agagtttgatcatggctcaggacgaacgctggcggcgtgcttaacacatgcaagtcgagcggaaaggcccttcggggfactcagcggc
gaacgggtgagtaaacacgtgagcaacctgcctcagctccgggatagctccgggaaaccggaattaataccgatatgaccacggccga
catcggctgtggtgaaaggattccggctgaggatgggctcgcggcctatcagctagtgtgagtaaacggctaccaaggcttgacgg
gtagctggtctgagaggacgactagccacactgggactgagacacggcccagatacctacgggtggcagcagtcgagaatfttgacaat
gggggaaacctgatcagcgcagccgcgtggaggatgaaggtcttcggattgaaactcctgtcattggggaacaatgcacagatattaa
ctgtattgtgtgatagtaccgaaagaggaagagacggctaactctgtccagcagccgcggtataacagagggtgcgagcgtgttcgg
aattattggcgtaaaggcgcgtagggcggcggtaagtcacctgtgaaatcctcggcttaactgggagctctcagggcgaactacct
gctgga

>ebs87

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtggagcaatccatgagtacgtg
gctgacgggtgagtaaacacgtgactaacctaccctcagtgggggaataaactcgggaaaccgagggcaataccgcataaacatctacggat
caaatgcgtaagcagctgaggagggggtcgcggctgattagctagtggcgggtaatggcccaccaaggcagtgatcagtagccggc
ctgagagggcgcacggccacactggaactgaaacacggctccagactcctacgggaggcagcagtgggggaatfttcgcaatgggggaa
acctgacgcagcaacgccgctggagatgaagcccttgggggtgaaactccttcgatcgggacgattatgacggtagccgatgaag
aagcaccggctaactctgtccagcagccgcggtataacagagggtgcaagcgtgttcggaattattggcgtaaagggtgcgtaggcg
gtcggtaagtctctgtgaaacctccggcctaactcggagcctgcagaggaaactgccgtgctggagtatgggagaggtgagtggaatt
cccgggtgtagcgggtgaaatcgtagatcgggaggaacacctgtggcgaaagcggctcactggaccattactgacgctgatgcagaaa
gctagg

>ebs89

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagagagtggagcaatccatgagtaaagt
gagaccgggtgagtaaacacgtgactaacctaccctcagtgggggataaactagggaaacctggcctaataaccgcataatgccttcgggtt
aaaggagcaatcgttgaagagggggtcgcggccgattaggtagtggcgggtaacggcccaccaagccgtatcggtagccggc
ctgagagggcgcacggccacactggaactgaaacacggctccagactcctacgggaggcagcagtgggggaatfttcgcaatggaggaa
acctgacgcagcaacgccgctggagatgaagtccttgggacgtaaaactccttcgaccgggacgattatgacggtagccggtggaag
aagccccggctaactcgtccagcagccgcggtataacgagggggcgaagcgtgttcggaattattggcgtaaagggtgcgtaggc
ggtgcggtaagtcacctgtgaaacctccggcctaactcggagcctgcagggcgaactaccgtgcttgagtgtgggagaggtgcgtgga
ttcccgggtgtagcgggtgaaatcgtagatcggg

>ebs90

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtggagcaatccatgagtaaagt
gagaccgggtgagtaaacacgtgactaacctaccctcagtgggggataaactagggaaacctgggctaataaccgcataatgccttcgggtt
aaaggagcaatcgttgaagagggggtcgcggccgattaggtagtggcgggtaacggcccaccaagccgtatcggtagccagc
ctgagagggcgcacggccacactggaactgaaacacggctccagactcctacgggaggcagcagtgggggaatfttcgcaatgggggaa
acctgacgcagcaacgccgctggagatgaagtccttgggacgtaaaactccttcgaccgggacgattatgacggtagccggtggaag
aagccccggctaactcgtgtagcagccgcggtataacgagggggcgaagcgtgttcggaattattggcgtaaagggtgcgtaggc
ggtgcggtaagtcacctgtgaaacctccggcctaactcggagcctgcagggcgaactgccgtgcttgagtgtgggagaggtgcgtgga
attcccgggtgtagcgggtgaaatcgtagatcgggaggaacacctgtggcgaaagcggcgcactggaccacaactgacgc

>ebs91

agagtttgatcctggctcagaacgaacgctggcggcaggcctcaaacatgcaagctagggggcgcagcaatgcatcaccggcggacg
ggtgcgtaacgcgtgggaatgtacctattggtcggaaactgagggaaacttcagctaataccggatgagcccgaagggggaaagatt
ctcgccaatggatcagcccgcgtccgattagctcgttgggtgggtaacggcctaccaaggcaacgacggtagctgttctgagaggaagat
cagccacactgggactgagacacggcccagactcctacgggaggcagcagttgggaatcttgacaatgggggaaacctgatccagc
catgccgcgtgagtgatgaaggccttcgggttgtaaaacttttcgacggggacgataatgacgggtaccgtagaagaagctccggctaac
tttgtccagcagccgcgtaatacgaagggggctagcgttggcgaacttctgacggggacgataatgacgggtaccgtagaagaagctccggctaac
ggtgaaaagcccggagctcaactccggaattgccattgaaactgttgactggagtacgagagaggtgagtggaaattcccagtgtagaggt
gaaattcgtagatattgggaagaacaccgggtggcgaaggcggctcactggctcgtaaactgacgctcagcgcgacagcgtgggatcaa
a

>ebs92

agagtttgatcatggctcaggatgaacgctagcggcaggcctaatacatgcaagtcggacgggactatttctagtagcaataccgggaata
gttttagtggcgaacgggtgcttaaacacgtgggcaacctgccaggaagtgggggatagctcggcgaaggcgaattaataccgcatgtgatt
agagaggacatccttcgaaattaaagtcggggcaacctgacgcttctgtggggcccggcctatcagctagttggtgaggttacggctc
accaaggctttgacgggtagctggtctgagaggacgaccagccacactgggactgagacacggcccagacacctacgggtggcagcag
tcgagaatthttgacaatgggggaaacctgatcgagcgacgccgcgtggaggatgaaggcttcggattgaaactcctgtcattggggaa
caatgcacagatattaactgtatttgttgatagtagccgaagaggaagagacgggtaactctgtccagcagccgcgtaatacagaggt
ctcaagcgttggcgttattgggcgtaaaagggtgcgtagggcggcgcggttaagtcgggtgtgaaatctcggagcttaactccgaaactgc
attcgatactgccgtgctgagga

>ebs94

agagtttgatctggctcaggatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaaagtgaggcaatccatgagtacagtg
cgaccgggtgagtaaacacgtgactacctaccctccgagtggggaataactccgggaaaccggggtaataccgcataacaccgcaaggtc
aaaggagcaatcgttggagagggggctcggcgtgattagctagttggttagggtaatggcctaccaaggcgaagatcggtatccggcct
gagagggcgcacggacacactggaactgaaacacggctccagactcctacgggaggcagcagttgggaatthttgcgaatgggggaaa
ccctgacgcagcaacgccgcgtggaggatgaaggcctttgggtcgtaaactccttcgaccgggacgatactgacgggtaccggtggaaga
agcaccggctaactctgtccagcagccgcggttaatacagaggggtcgcagcgttggcgaattattgggcgtaaaaggcgcgtaggcg
gtgcggttaagtcacctgtgaaatccccaggcttaacttgggcctgcagcgaactgccgtgctggagtgtgggagaggtgcgtggaatt
cccgtgtagcgggtgaaatcgtagatcgggaggaacacc

>ebs98

agagtttgatcatggctcaggatgaacgctagcggcaggcctaatacatgcaagtcggacgggactatttctagtagcaataccgggaata
gttttagtggcgaacgggtgcttaaacacgtgggcaacctgccaggaagtgggggatagctcggcgaaggcgaattaataccgcatgtgatt
agagaggacatccttcgaaattaaagtcggggcaacctgacgcttctgtggggcccggcctatcagctagttggtgaggttacggctc
accaaggctttgacgggtagctggtctgagaggacgaccagccacactgggactgagacacggcccagacacctacgggtggcagcag
tcgagaatthttgacaatgggggaaacctgatcgagcgacgccgcgtggaggatgaaggcttcggattgaaactcctgtcattggggaa
caatgcacagatattaactgtatttgttgatagtagccgaagaggaagagacgggtaactctgtccagcagccgcgtaatacagaggt
ctcaagcgttggcgttattgggcgtaaaagggtgcgtagggcggcgcggttaagtcgggtgtgaaatctcggagcttaactccgaaactgc
attcgat

>ebs100

agagtttgatcatggctcaggacggacgctggcggcgtgcctaatacatgcaagtcgagcggagttattccttcggggatagcttagcggc
ggacgggtgagtaaacgtagtaacctgcctgtaagactgggataactaccgaaacggtagctaagaccggatagcggcattctcgc
atgagagagtcgggaaaagcggagcaatctgctactatagatggacctgcggcgcattagctagttggtgggtaacggctcaccagg
cgacgatgcgtagccgacctgagaggggtgaacggccacactgggactgagacacggcccagactcctacgggaggcagcagtaggga
atctccgcaatggacgcaagtctgacggagcaacgccgcgtgagtatgaagtttcggatcgtaaagctctgttgcagggaagaacg
cttgggagagtaactgctctcaaggtgacggctacgtgagaagaaagccccgtaactacgtccagcagccgcgtaatacgtagggg
gcaagcgttgcgggaattattgggcgtaaaagcgcgcgagcggccttgaagtctggtgttaaaacttagggctcaacctgagtcgcat
ggaaaactgcaggcttgagtgcagaagaggagagtggaattccacgtgtagcgg

>ebs101

agagtttgatcctggctcaggacgaacgctggcggcgtgcctaaacatgcaagtcgaacgggtagcaatacctagtggcgaacggcg
gagtaatacgtgaacaacgtgccttaaagactgggatagccaggggaacacctgagtaataccggatgtggcaacgggatggcatcatc
cgttactaaatggttttcgctttaaagagcggttcacggcctattagctagttggcggggtaacggccccaccaaggccatgatcggtagccgg
cctgagagggcgcacggccacactggaactgaacacggcctcagactcctacgggagggcagcagtggggaatgttcgcaatggggga
aacctgacgcagcaacgccgcgtggaggatgaagccccttgggggtgaaactccttcgatcgggacgattatgacggtagccgatgaa
gaagcaccggcctaactctgtccagcagccgcggtaatacagaggggtgcaagcgtgttcggaattattgggcgtaaaaggcgcgtaggc
ggtgcggtaagtctattgtgaaatctccgggctcaactggagcctgcaacggaaactgccgtgctcagtggtggagaggtgagtggaat
tcccgggtgtagcgggtaaatgcgtagatcggg

>ebs102

agagtttgatcatggctcaggatgaacgctagcggcgtggataagacatgcaagtcggacgagatgtttcgttagcaatacgatgaaag
tctagtgccgaacgggtgctaacacgtgggcaacctgccgagaagtgggggacagctcggcgaaggcggattaataccgatgtgatt
agggaggcctcctcgaataaagtcggggcaacctgacgcttctgtggggcccgcggcctatcagctagttgtgaggtaatagct
caccaaggctttgacgggtagctggtctgagaggacgaccagccacactgggactgagacacggcccagacacttacgggtggcagca
gtcgagaatgttctcaatgggggaaacctgaaggagcgcgccgcgtggaggatgaaggcttccgggtgaaactcctgtcatttggga
acaacgcctcgcgttaactgcgtgaggttgatagtagcagaagaggagagacggcctaactctgtccagcagccgcgtaatacaga
gggtccgagcgtttcggattcattgggcgtaaaagggtgtgtaggaggtcaggcaagtcgggtgtgaaatctcggcgttaacggcgaaa
tggcactcgaactgcttggctcagggattggagggggagggcggaattcatggttagcgggtgaaatgcgtagatcgggaaggaaacacc
tgttcggaaggcggctcctggactcaact

>ebs103

agagtttgatcatggctcagagcgaacgctggcggcagggcctaacacatgcaagtcgagcggcgttagcaatacggagcggcagacggg
agagtaaacacgtgggaacgtgccctcagttcggacaacccagggaaacttgggctaataccggatacatccgtaaggagaaagattat
cgtgaaaggatcggcccgcgtctgattagctagttgggtgggtaaatggcccaccaaggcgcagatcagtagctggtctgagaggatgatc
agcctcactgggactgagacacggcccagactcctacgggagggcagcagtggggaatattggacaatgggcgcaagcctgatccagcc
atgccgcgtgggtgatgaaggccctaggggtgaaagcccttcggcggggaaagataatgacggtagccgcagaagaagccccggctag
ctcgtgccagcagccgcgtaatacgaagggggctagcgttctcggaaactcggcgtaaaagcgcacgtaggcggcttttaagtcag
gggtgaaatcctggagctcaactcagaactgccttgatactgagaagcctgagtcgggagaggtgagtggaactgcgagtgtagaggt
gaaattcgtagatattcgaagaacaccagtgggcgaaggcggctcactggcccgtactg

>ebs104

agagtttgatcctggctcagaacgaacgctggcggcgtgctaacacatgcaagtcgcgtgcgagttcgcgttagcaatacgtggatgggc
gcacggcgcacgggtgcgtaaacacgtgggcaacctgccctctgatggggaataaccccgcgaaagcggggctaataccgcataacattc
cttgattatggttaggaatcaaagcagtgatgcgtcggaggagggggcccgcgccgattaggtagttggtgaggtaacggctcaccaa
gccaatgatcggtagctggtctgagaggatgatcagccacactggaactgagacacggcctcagactcctacgggagggcagcagtgggga
atcttgacaatgggcgaaagcctgatgcggcgcacgccgcgtgggggatgaagcttttcggagtgtaaaccccttcgaccgggacgaat
gcccgcaaggagtgacggtagcgtatagaagccccggcctaactacgtgccagcagccgcgtaagacgtagggggccagcgttgt
tcggaattactgggtgtaaaagggtcgtagggcgggtggaagtcgggagtgaaatctctgg

>ebs106

agagtttgatcctggctcagaacgaacgctggcagcagggctaacacatgcaagtcgaacgcaccgcaaggtgagtgccagacgggtga
gtaacgcgtgggaacctcctctgtgttacggaataactcgggaaaccgaagctaataccgtatatctcctccgggagaaagatttatgcc
ataggatggggcccgcgttgattagctagttggtagggtaatggcctaccaaggcaacgatccatagctggtctgagaggatgatcagcca
cactgggactgagacacggcccagactcctacgggagggcagcagtggggaatattggacaatgggcgcaagcctgatccagccatgcc
gcgtgggtgatgaaggccctaggggtgaaagcccttcggcggggaaagataatgacggtagccgcagaagaagccccggcctaactcgt
gccagcagccgcgtaatacgaagggggctagcgttctcggaaactcggcgtgaaagcgcacgtaggcggcttttaagtcaggggt
gaaatcctggagctcaactccagaactgccttgatactgagaagcctgagtcgggagaggtgagtggaactgcgagtgtagaggtgaa
attcgtagatattcgaagaacaccagtg

>ebs108

agagtttgatcatggctcaggatgaacgctagcggcaggcctaatacatgcaagtcggacgggactatttctagtagcaataaccgggaata
gttttagtggcgaacgggtgcttaacacgtgggcaacctgccaggaagtgggggatagctcggcgaaggcgaattaataaccgcatgtgatt
agagaggacatccttfcgaaftaaagtcggggcaacctgacgcttctgtggggcccgcggcctatcagctagttggtgaggttacggctc
accaaggctttgacgggtgactgggtctgagaggacgaccagccactgggactgagacacggcccagacacctacgggtggcagcag
tcgagaatthttgacaatgggggaaacctgacgagcgcgcccgcgtggaggatgaaggtcttcggattgaaactcctgtcattggggaa
caatgcacagatattaactgtatttgttgatagtagccgaagaggaagagacgggtaactctgtccagcagccgcggaataacagaggt
ctcaagcgttggcgttattgggctgaaaggggtgcgtagggcggcgcggtaagtcgggtgtgaaatctcggagcttaactccgaaactgc
attcgatactgccgtgctgaggactggagaggagactggaacttacggtgtagcgggtgaaatgcgtagatat

>ebs109

agagtttgatcatggctcagaatcaacgctggcggcgtgcttaacacatgcaagtcgaacgcgcagtcacctgtagcaatacaggggatgg
gcgcgtggcgaacgggtgcgtaaacacgtggataaccttcttcaggtggggaataaccccgcgaaagcggggctaataaccgataacat
gacgatggcacatgctgttctcaaaagccgcaagggcggcagaagaggatccgcggccgattagctagttggcgaggtaacggctcac
caaggcaatgatcggtagccgggtctgagaggatgacggccacactggaactgagacacggtccagactcctacgggagggcagcagtg
gggaatcttgcgaatggcgaaagcctgacgcagcaacggcgcgtgggggatgaagcttctcggagtgtaaaccccttgcaccggga
cgaatggcggcaaggggttggcgtacgggtacaagaagccccggctaactacgtgccagcagccgcggtaaaacgtagggggccagc
gttgctcggaaactgggctgaaaggggtccgtagggcgggtgtggcaagtcggaaagtgaatctcgggcttaactcagaggctcctccgaa
actgctgtgtagagtgagagaggcagtggaattgcgggtgtagcgggtgaaatgcg

>ebs110

agagtttgatcctggctcagaacgaacgctggcggcgtgcttaacacatgcaagtcgcgtgcgcagttcgcgtagcaatacgtggatgggc
gcacggcgcacgggtgcgtaaacacgtgggcaacctgccctctgatggggaataaccccgcgaaagcggggctaataaccgataacattc
cttgattatggttaggaatcaaagcagtgatgcgtcggaggagggggcccgcggccgattaggtagttggtgaggtaacggctcaccaa
gccaatgatcggtagctggtctgagaggatgatcagccacactggaactgagacacggtccagactcctacgggagggcagcagtgggga
atcttgcacaatggcgaaagcctgatgcggcgcgctgggggatgaagcttctcggagtgtaaaccccttgcaccgggacgaat
gcccgcgaaggagtgacgggtaccggtataagaagccccggctaactacgtgccagcagccgcggtaagacgtagggggccagcgtgt
tcggaataactgggtgtaaaag

>ebs112

agagtttgatcatggctcaggatgaacgctagcggcaggcctaatacatgcaagtcggacgggactatttctagtagcaataaccgggaata
gttttagtggcgaacgggtgcttaacacgtgggcaacctgccaggaagtgggggatagctcggcgaaggcgaattaataaccgcatgtgatt
agagaggacatccttfcgaaftaaagtcggggcaacctgacgcttctgtggggcccgcggcctatcagctagttggtgaggttacggctc
accaaggctttgacgggtgactgggtctgagaggacgaccagccactgggactgagacacggcccagacacctacgggtggcagcag
tcgagaatthttgacaatgggggaaacctgacgagcgcgcccgcgtggaggatgaaggtcttcggattgaaactcctgtcattggggaa
caatgcacagatattaactgtatttgttgatagtagccgaagaggaagagacgggtaactctgtccagcagccgcggaataacagaggt
ctcaagcgttggcgttattgggctgaaaggggtgcgtagggcggcgcggtaagtcgggtgtgaaatctcggagcttaactccgaaactgc
attcgatactgccgtgctgaggactggagag

>ebs116

agagtttgatcatggctcaggacgaacgctggcggcgtgcttaacacatgcaagtcgagcggaaaggcccttcggggactcagcggc
gaacgggtgagtaaacacgtgagtaacctgcccccttctgggataagcctgggaaactgggtctaataaccggatgacatgttctcgc
ggggatgttgaaagttttcgggtgggggatgggctcggcggcctatcagcttgttgggggtgatggcctaccaaggcgcagcagggta
gcccggcctgagagggcgaccggccacactgggactgagacacggcccagactcctacgggagggcagcagtggggaatattgcacaat
gggcggaagcctgatgcagcgcgcccgcgtgaggatgacggccttcgggttgaacactcttgcaccgacgaacctgagcggtag
gtgcagaagaagcggcgttaactacgtccagcagccgcggaataacgtagggcgaagcgttgcggattattgggctgaaagag
ctcgtagtggttctgcgctcgtccgtgaaagccatggcttaactgtgggtctgcgggtgatacgggcaggctagagggcaggtagggg

>ebs120

agagtttgatcctggctcagaacgaacgctggcggcgtggataagacatgcaagtcgagcgcgtgattgcagggtagcaatattctgtagttg
gagcggcgaagggtgcgtaaacacgtgggtaactcgcctgaagttgggataactcgcgtaaggcgcagctaaatacaggatgtgaagatt
ggaaggcatctttgatctcaagctggggaccgtaaggcctggcgcctcatgatgagcccgcggcctatcagctagttggtgaggtaacg

gctcaccaaggctaagacgggtagctggctgagaggacgaccagccacactggaactgagacacggctccagacacctacgggtggca
gcagtcgagaattttcacaatgggggaaacctgatggagcgacgccgcgtgggggatgaatggctcggcccgtaaacctgtcattc
gcgaacaatgcctgtgtattaacaccatacaggtgatagtagcggagaggaaggacggctaactctgtccagcagccgcggtaata
cagaggtccaagcgttgctggattcactgggcgtaaagggtgcgtaggcggctgggtaagtctgatgtaaatcacagcttaactgtg
aaactgcattggatactattcggcttgaggctggag

>ebs125

agagttgatcatggctcagagcgaacgctggcggcaggcctaacacatgcaagtcgagcggcgtagcaatacggagcggcagacggg
agagtaacacgtgggaacgtgcccttcagttcgggaacaaccagggaactgggctaataccggatacatccgtaaggagaaagattat
cgctgaaggatcggcccgcgtctgattagctagttgggggtaatggcccaccaaggcgacgatcagtagctggctgagaggatgatc
agcctcactgggactgagacacggcccagactcctacgggagggcagcagtgggggaatattggacaatgggcgcaagcctgatccagcc
atgccgcgtgggtgatgaaggccctaggggtgtaaagcccttcggcggggaagataatgacggtaaccgcagaagaagccccggctag
cttcgtccagcagccgcgtaatacgaagggggctagcgttctcggaaactcactgggcgtaaagcgcacgtaggcggcttttaagtcag
gggtgaaatcctggagctcaactcagaactgccttgatactgagaagcctgagtcgggagaggtgagtggaactgcgagtgtagggt
gaaattcgtagatattcgaagaacaccagtgggcg

>ebs170

agagttgatcctggcttagaacgaacgctggcggcgtggataagacatgcaagtcgagcggcgtgatttcggtagcaatattcccagtcg
gagcggcgcaagggtgcgtaacacgtgggtaactgcgatgaagtcgggataactgtgaaaggcgagctaataccggatgtgataact
gaaaggcatcttttggtattaaagtggggaccgtaaggcctgacgctcttgatgaaccggcctatcagctagttggtgaggtaacgg
ctaccaaggctaagacgggtagctggctgagaggacgaccagccacactggatctgagacacggctccagacacctacgggtggcag
cagtcgagaattttcacaatgggcgaaagcctgatggagcgacgccgcgtgggggatgaatggctcggcctgtaaacctgtcattc
cgatcaaaccttatttttaaaagatgatgagctgatagtagcggaaagaggaaggacggctaactctgtccagcagccgcggtaataca
gaggtccaagcgttgctggattcactgggcgtaaaag

>ebs172

agagttgatcatggctcagaacgaacgctggcggcgtggataagacatgcaagtcggacgagattttcgttagcaatacgtgaaag
tctagtggcgaacgggtgcttaacacgtgggcaacctgccgagaggtgggggacagctcggcgaaggcggattaataccgatgtgatt
agggaggacatcctcgaatfaaagtcggggcaacctgacgctcttggtggggcccggcctatcagctagttggtgaggtaatagctc
accaaggctttgacgggtagctggctgagaggacgaccagccacactgggactgagacacggcccagacacctacgggtggcagcag
tcgagaattttgacaatgggggaaacctgatcgagcgacaccgcgtggaggatgaaggtcttcggattgtaaactcctgtcattggggaa
caatgccccgaatfaactgattcgggggtgatagctcgaagaggaagagacggctaactctgtccagcagccgcgtaatacagagg
tctcaagcgttgctggattcattgggcgtaaagggtgcgtaggcggcggtaagtcgggtgtgaaatcttgagcttaa

>ebs175

agagttgatcatggctcagggtgaacgttgtaacgtggataaccggcatgcaagtcgaagggggtccgacctggcggaccaccggcga
acggcttagtaacacgtggatacaccgaaagtggaggacagctcgtcgaagacgggtaatactccatagcttgtaaaagtaaagg
cgcaagccgctttgcgaatggtccggcctatcagcttggcggggtaaaagcccaccaaggcgatgacgggtagggggcgtgaga
gcgacccccaaatggcactgcgacacgggccatactcctacgggaggtcgagtcgagaatcttcgcaatggacgaaagtctga
cggagcgacgttacgtgaaggatgaagcccttcgggggtgaaactctttatctgggaggaatattgacagtaccagatgaataagggg
atcctagtctgtccagcaggagcgtaatacagaatccccgaatgtaccggattattgggcgtaaagggtgtgtaggtggtttgtaa
gttcaagattaaacctcgtcggcttaacctcagctgtttgaaactgccagactaagagggcgttagaggcaaacggaaccgatgaagta
gggggtgaaatccgttgatcatcgggaa

>ebs107

agagttgatcatggctcagaacgaacgctggcggcgtggataacatgcaagtcgagggcggattggtgcgtacacgtgggtaactgcc
atgaagtcgggataactgctgaaagcgagctaaaccggatgtgacagccgagagacatctttggtgtaaaagtggggaccgtaag
gcctgacgctcttgatgaaccggcctatcagctagttggtgaggtaacggctcaccaggcgaagcgggtagctggtctgagaggga
cgaccagccacactggaactgagacacggctccagacacctacgggtggcagcagtcgagaattttcacaacgggcgcaagcctgatag
agcgacccgcgtgggggatgaatggctcggcccgtaaacctgtcattcgcgatcaaacgttatttttaaaagatgataacctgatagt
agcggaaagaggaaggacggctaactctgtccagcagccgcgtaatacagaggtccaagcgttgctcggattcactgggcgtaaaag

ggcgtaggtggagaggtaagtcggatgtgaaatctcggagctcaactccgaaatggcattggaaactacctcgtcaggggtggagg
ggggactggaatacttggtgtagcagtgaaatgcgtagatatcaagtggaacaccagtggcgaaggcgagtcctgggacactcctgaca
ctgaggcacgaaagctaggggagcaaacaggat

>eps413

agagttgatcatggctcagagcgaacgctggcggcaggcctaacacatgcaagtcgagcgccgtagcaatacggagcggcagacggg
agagtaacacgtgggaacgtgcccttcagttcgggaacaaccagggaacttgggctaataccggatactccgtaaggagaaagattat
cgctgaaggatcggcccgctctgattagctagttggtggggtaatggcccaccaaggcgacgatcagtagctggctgagaggatgac
agcctcactgggactgagacacggcccagactcctacgggaggcagcagtggggaatattggacaatggcgcaagcctgatccagcc
atgccgcgtgggtgatgaaggccctaggggtgtaaagcccttcggcggggaagataatgacggtaccgcagaagaagccccggctaa
cttcgtgccagcagccgggtaatacgaagggggctagcgttgctcggaaactcactggcgtaaaagcgcacgtagggcggcttttaagtcag
gggtgaaatcctggagctcaactcagaactgcctttgatactgagaagcttgagtcgggagaggtagtggaactgcagtgtagagg
gaaatcgtagatattcgaagaacaccagtggcgaaggcggctcactggcccggtagctgaggtgcgaaa

>ebp9

agagttgatcatggctcaaaatctacgatggcggcgtgactaatacttgcagctagcgagaaaggtagcaatccatgagtacagcggc
gcaccggagaataacactagactgacctaccctccagtggggaataactccggaacctatgctaataccgcataataccctagagctcaa
aggagcaatccgcttgaggagggggctcgcggccgattagttagttggcggggtaatggcccaccaagacgatgacggtatccggcctg
agagggcgacggacacactggaactgaaacacggctccagactcctacgggaggcagcagtggggaatcttcgcaatgggggaaacc
ctgacgcagcaaccccgcgtggaggatgaagtccttcgggacgtaaaactccttcgacgggacgattatgacggtaccggaagatgaag
ccccggttaactcgtgccagcagccgcggtatacgaagggggcagcgttgctcggaaatattggcgaaaagggctccgatgcgatgt
gacaagtcggtagtgaaatctccgggcttaactcggaggtgctgccgaaactggtgctagagtgatagaggcagtgtaattacggg
ttagcgtgaaaatcgcataatataatccctaagaacatccgaagcgaagcggctcgtggaacacaaatgaccctgagggacgaaagcta
ggggagcaaacggg

>ebp14

agagttgatcctggctcagagcgaacgctggcggcaggcctaacacatgcaagtcgagcgctgtagcaatacacgagcggcagacg
ggagagtaaacacgtgggaacgtaccctcggctcggaaataaccagggaacttggactaataccggatactccgaaaggagaaagatt
taccgccgaaggatcggcccgcgctgattagctagttggtgaggtaacggctaccaaggcgacgatcggtagctggctgagaggatg
atcagcctcactggggctgagacacggcccagactcctacgggaggcagcagtggggaatattggacaatggcgcaagcctgatcca
gccatgccgcgtgggtgatgaaggccttaggggtgtaaagcccttcggcggggaagataatgacggtaccgcagaagaagccccggc
taactcgtgccagcagccgcggtatacgaagggggctggcgttgctcggaaactcactggcgtaaaagcgcacgtagggcggattgtaag
tcaggggtgaaatcctggagctcaactccagaactgcctttgatactggcgtccttgagttcgggagaggtagtggaactgcagtgtaga
ggtgaaatcgtagatattcgaagaacaccagtggcgaaggcggctcactggcccatactgacgctgaggtgcgaaaagcgtggggag
caaacaggattagataccctggtagtcacgccgtaaacgatggatgctagccggtggcgggttactcgtcagtgggcgacgtaacgcttt
aagca

>ebp101

agagttgatcctggctcagagcgaacgctggcggcgtgcttaacacatgcaagtcgaaagggagcaatcctgagtaaagtggc
gtacgggtgagtaacgcgtaggtaatctacccagagctcggaaataacctgccgaaaggcgggctaatactggatagagcagcgaatcc
gtaattgggttttcgggaaaggaggaagcaatcccttcgctctgggatgagcctgcgtcattagctagttgtagggtaacggcctac
caaggctatgatgggtagctggctgagaggatgatcagccacactggaactgaaacacggctccagactcctacgggaggcagcagtg
ggaatcctgcgcaatggacgaaagtctgacgcagccacgccgcgtgagtgaaaggccttcgggttgtaaagctctgtagcgggac
gaaaaatcttaggtgaaatagcctagagatctgacggtagcctaaaggaagcaccggcctaactccgtccagcagccgcggtatacggg
gggtgcaagcgtgctcggaaatattggcgtaaaaggtaggtgtaggtgctcatttgctgggtgaaagccttgggcttaactcaagaagt
gccccagaacggtagactggagtcctggagagggtcgtggaattcccgggtgtagcgggtaaatcgtagagatcg

>ebp104

agagttgatcctggctcagagcgaacgctggcggcgtgcttaacacatgcaagtcgaaagggagcaatcctgagtaaagtggc
gtacgggtgagtaacgcgtaggtaatctacccagagctcggaaataacctgccgaaaggcgggctaatactggatagagcagcgaatcc
gtaattgggttttcgggaaaggaggaagcaatcccttcgctctgggatgagcctgcgtcattagctagttgtagggtaacggcctac

caaggctatgatggtagctggtctgagaggatgatcagccacactggaactgaaacacggccagactcctacgggaggcagcagtg
ggaatcttgcgcaatggacgaaagtctgacgcagccacgccgcgtgagtgaagaaggccttcgggtttaaagctctgtcagcgggac
gaaaaatfttaggtgaatagcctagagatctgacgggtaccgctaaggaagcaccggctaactcctgccagcagccgcgtaatacggg
gggtgcaagcgttctcgggaattattggcgtaaaagggtaggtggtctcatttctggggtgaaagccttgggcttaactcaagaagt
gccccagaaacggtagactggagtcctggagagggcgtggaattcccggtgtagcgggtaaatgctagagatcgggaggaacacc
agaggcgaaagcggcgacctggacaggtactgacactcaacta

>ebp105

agagtttgatcatggctcagaacgaacgctggcggcatgcctaatacatgcaagtcgaacgcttttggggttccggctccaggggagtg
gcacgggtgagtagcgcgtgggaacctgccctgttgggggataaccgagggaacttcggctaataccgcatgcgcccgcggagttg
atccgcgggggaaagatttctcagcagcagggggccgcgtcggattagctggttggcggggtaacggcccaccaaggcagatgatccgt
agctggtctgagaggatgatcagccacactgggactgagacacggcccagactcctacgggaggcagcagtagggaatattggacaatg
ggcgcaagcctgatccagcaatgccgcgtgagtgaagggccttcgggtttaaagctcttcgcccgcgacgatgatgacggtagcgg
gagaagaagccccggctaactcgtgccagcagccgcgtaatacgaagggggcagcgttctcggattactgggcgtaaaagggcgc
gtaggcggctcgtcaagtcagggcgtgaaagccccgggctcaacctgggaggtgcgcttgggactggcgggcttgagttcgggagagga
gctggaattccagtgtagaggtgaaattcgtagatattgggaagaacaccgggtgccaagggcggcgtctggaccgagactgacgctg
aggcgcgaaag

>ebp107

agagtttgatcctggctcagagcgaacgctggcggcaggcctaacgcatgcaagtcgagcgggtgtagcaatacatcagcggcagacgg
gtgagtaacgcgtgggaactgacctttggtcgggaacaaccagggaacttgggctaataccggataagcccttacggggaaagattat
cgccgagagatcggcccgcgtctgattagctagttggtgaggtaatggctcaccaaggcagcagtagctggtctgagaggatgatca
gccacattgggactgagacacggcccagactcctacgggaggcagcagtggggaatattggacaatgggcgcaagcctgatccagccat
gcccgcgtgagtgatgaaggccctagggtttaaagctctttgtcgggaagataatgacggtaccgcaagaataagccccggctaactc
gtgccagcagccgcgtaatacgaagggggctagcgttctcggaaactcgggcgtaaaagggtgcgtagggcggctttaaagtcaggg
gtgaaatcctggagctcaactccagaactgcctttgatactgaagatcttgagttcgggagaggtgagtggaactgcgagtgtagaggtgaa
attcgtagatattcgaagaacaccagtgccgaaagcggctcactggcccagactgacgctgaggcagcaaaagcgtggggagcaaaaca
ggattagataccctgtagtccacgccgtaaacgatgaatgccagccgtt

>ebp109

agagtttgatcatggctcagaacgaacgctggcggcgtggataagacatgcaagtcgagcggctgatttgcgggtgtagcaatattcgaagtc
ggagtgccgcaagggtgtgtaacacgtgggtaacttgccatgaagtcggaataacttgctgaaaggcagcctaataccggatgtgaacgc
aagttcaaagttggggaccgtaaggcctgacgcttcatgatgagcccggccctatcagctagttggtgaggtaacggctaccaaggcca
agacgggtagctggtctgagaggacgaccagccacactggaactgagacacggctccagacacctacgggtggcagcagtcgagaatftt
tctcaatggcgaaagcctgaaggagcagcggcgtggggatgaatggcttcggcccgtaaaccctgtcatttgcgaacaaattgtt
cacctaacacgtgaagcattgatttagcggaaagaggaaggacggctaactctgtccagcagccgcgtaatacagaggtcccaagc
gttctcggattcactgggcgtaaaagggtgcgtagcggccaggtgaaattgatgtaaatcccgagctcaactgcggaactgcattgaat
actatttgctcaggaatggaggggagactggaattctcgggtgtagcagtgaaatgcgtagatcagaggaacaccagtgccgaaag
cgagctctggacatttctgacgct

>ebp111

agagtttgatcctggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtgagcaatccatgagtacagtg
gcgaccgggtgagtaacacgtgactacctacccttgattgggggataacccgggaaaccgggctaataccgataacatcagagat
caaaggagcaatccgctctggagggggcgcggctgattagctagttggtgaggtaacggctaccaaggcgaagatcggatccggcc
tgagaggcgcacggacacactggaactgaaacacggctccagactcctacgggaggcagcagtggggaatfttgcgcaatgggggaaa
ccctgacgcagcaacgccgcgtggaggatgaagcccttggggcgtaaactccttcgaccgggacgatactgacggtagcgggtggaag
aagcaccggctaactctgtccagcagccgcgtaatacagaggggtgcgagcgttctcggaaattattgggcgtaaaagggcgcgtagggc
gtcgggtaagtcacctgtgaaaccttgggcttaactcagagcctgcagggcaaaactgccgtgctcagtggtgggagaggtgcgtggaatt
cccgggtgtagcgggtaaatgcgtagatcgg

>ebp113

agagtttgatcatggctcagagcgaacgctggcggcgtgcctaacacatgcaagtcgaacgagggtagcaataccttcgtggcgaacggg
tgcgtaacacgtggataaccttcctcaggtggggaataaccccgcgaaagcggggctaataccgcataacaagacgagggcacatgcct
ttgtcttcaaagggcgaagtcgccggaagaggggatccgcggccgattagctagttggcgggtaaacggcccaccaagggcaatgatcgg
agccgggtctgagaggatgatcggccacactggaactgagacacgggtccagactcctacgggagggcagcagtggggaatcttgcgcaatg
ggcgaagcctgacgcagcaacgccgcgtgggggatgaagcttctcggagtgtaaaccccttcgacccggacgaatgcccgcaagg
cttgacgggtacgggtataagaagccccggctaactacgtgccagcagccgcggtaaaacgtagggggccagcgttgctcgggaattactgg
gcgtaaaagggtccgtaggcgggtggtgcaagtcggaagtgaatctctgggcttaactcagaggctgctccgaaactgctgtgctagagtgt
gagagaggcaagtgggaattgcgggtgtagcgggtgaaatgcgtagatatccgcaggaacatccgaggcgaaagcggccttctggtatcaca
actgacgct

>ebp114

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaaggggagcaatccctgagtaaagt
gagcaccgggtgagtaacacgtgactaacctaccctcagtgggggataacctcgggaaaccggggctaataccgcataatgccttaggg
taaaggagcaattcgttagggaggggggtcgcggccgattaggtagttggcgggtaaacggcccaccaagcccgtgatcggtagccggc
ctgagagggcgcacggccacactggaactgaaacacgggtccagactcctacgggagggcagcagtggggaatcttgcgcaatgggggaa
acctgacgcagcaacgccgcgtggaggatgaagctcttgggacgtaaacctcttcgaccgggacgataatgacggtaccggtggaag
aagccccggctaactctgtgccagcagccgcgtaatacagagggggcaagcgtgttcggaattattgggcgtaaaaggcgcgtaggc
ggtgcggtaagtcacctgtgaaacctctgggcttaactcagagcctgcagccgaaactccgtgctggagtgtgggagaggtgcgtgga
ttccgggtgtagcgggtgaaatgcgtagatatcgggaggaacacctgtggcgaaagcggcgcactggaccacaactgacgctgagggcgcg
aaagctag

>ebp417

agagtttgatcatggctcagaacgaacgctggcggcgtggataagacatgcaagtcggacgagatttttcgtttagcaatacagatggaaag
tctagtggcgaacgggtgcttaacacgtgggcaacctgccgagaagtggggacagctcgcgaaagggcggattaataccgatgtgatt
agggaggacatcctcgaattaaagtcggggcaacctgacgctcttctggggcccgcggcctaccagctagttggtgaggtaatagct
caccaagccttgacgggtagctggtctgagaggacgaccagccacactgggactgagacacggcccagacacctacgggtggcagca
gtcgagaatcttgacaatgggggaaacctgatcgagcagccgcgtggaggatgaaggtcttcggattgtaaactcctgctattgggga
acaatgccccgaatttaactgattcgggggtgatagctcgaagaggaagagacggctaactctgtgccagcagccgcgtaatacagag
gtctcaagcgttctcggattcattgggcgtaaaagggtgcgtagccggcggtaagtcgggtgtaaatcttgagacttaactccgaaact
gcattcgatactccgtgctcaggactggagaggagactggaattacggtgtagcgggtgaaatgcgtagatatcgaaggaagaccagt
ggcgaagggcgggtctctggacagttcctgacgctga

>ebp120

agagtttgatcctggctcagagttaacgctggcagcgtgcataacacatgcaagtcgagcgggctttggcggggatcgcctagagctagc
gcagacgagtgagtaacacgtaagaaatctacccaaaataggaacaacctgctgaaagggcgggctaataccctatataatcctggaatg
gcatcatccgggatgaaagggcgtgagccggcttgggagggatcttgcggcctatcagctagttggtgggtaaacggcccaccaaggtat
gacgggtagccggcctgaaagggcgtcggccacggaggcgtgagaaacggcctcactcctacgggagggcagcagtggggaat
tggacaatgggggaaacctgatccagcaacgctgcgtgaaaggacgaaggtcttcggattgtaaactctttgttcgggacgaaagccc
caagggttgacgggtaccggatgaataagccacggctaactacgtgccagcagccgcgtaagacgtaggtggcgaagcgttactcggaa
cactagggcgtaaagcgcgtgtaggggggtgc

>ebp123

agagtttgatcatggctcaggatgaacgctggcggcgtgcctaatacatgcaagtcgagcgcctggacagcgcacttcttgcctccgtctga
gacagggcgaacaccggcagggtgttctcatcaaaaaagaaggaggagcaagaaggagtgcgatgtccaggagcggcggacggc
agagtaacgcgtaggaacgtacccaaactgagggataagcccgggaaaccgggtctaataccgcataattatcttcgatcaaaactcgg
cggtttgggagcggcctgcgtatgattgcttgggtgaggtaatggctaccaaggcgcagatcattagctggtctgagaggatgatcag
cctgactgggactgagacacggcccagactcctacgggagggcagcagtaggggaatctccgcaatgggggcaacctgacggagcaac
gccgcgtgcaggaggaaggccttcgggtgtaaacgtttatctgcgacgattatgacggtagcagaagaataaggatcggctaattccg
tgccagcagccgcggctacaggaagatccgagcgttatccggaattactggcgtaaaagagtgcgtaggtggcgtgaaagcagtagt
gaaagatcccggtcaaccgggggacattatcgaactcacaagctagagcatagtagaggtacctggaattccaggt

>ebp125

agagtttgatcctggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaaaggaggcaatccatgagtaaagtgcgacgggtgagtaaacacgtgactacctaccctcgagtggggaataactccgggaaaccggggctaataaccgcataacatcgagagatcaaagcagcaatgcgcttgaggagggggcgcgctgattagctagtgggtgaggaacggctcaccaaggcgaagatcggtagccggcctgagagggcgcacggccacactggaactgaaacacgggtccagactcctacgggaggcagcagtggggaatfttgcgcaatgggggaaaccctgacgcagcaacgccgcgtggaggatgaagtccttgggacgtaaaactccttcgaccaagacgataatgacggtactggtggaagaagcaccggctaactctgtccagcagccgcggtgtaatacagagggtgcgagcgtgttcggaattattgggcgtaaaaggcgcgtagggcgtgcggtaaatcacctgtgaaatctcccggcttaactgggagcctgcagcgaactgccgtgctggagatgggagaggtgcgtggaattcccgtgtagcgggtgaaatgcgtagatatcgggaggaacacctgtggcgaaag

>ebp126

agagtttgatcatggctcagattgaacgctggcggcatgccttacacatgcaagtcgaacggcagcgcgggagcaatcctggcggcagtgggcgaacgggtgagtaataacatcggaacatgtccagtagtgggggatagcccggcgaaagccgggattaataaccgatacgtatcaggatgaaagcgggggatcttcggacctcgcgctattgggggtggccgatggcggattagctagtgggtgggtaaaaggcctaccaaggcgcagatccgtagctggtctgagaggacgaccagccacactgggactgagacagggccagactcctcgggaggcagcagtggggaatfttggacaatgggcgaaagcctgatccagcaatgccgcgtgtgtaagaaggccttcgggtgtgaaagcactfttgcgggaaagaaatcctggtggctaataaccgggggatgacggtaccggaagaataagcaccggctaactacgtgccagcagccgcggtgtaatacgtagggtgcgagcgttaatcggaaactggggcgtaaagcgtgcgagcgggtgatgtaagaccgatgaaatccccgggcttaacctgggaactgcattggtgactgcatcgtggagtagtgccagagggggtagaattccacgtgtagcagtgaaatgcgtagagatgtggaggaataccgatggcga

>ebp129

agagtttgatcatggctcagagtgaacgctggcggcatgcttaacacatgcaagtcgcacgggcaggggcaaccctgtcagtgggcggacgggtgagtaacgcgtaggtatctatcctcgggtgggggacaaccacgggaaactgtggctaataaccgatgacgactgaggtccaaaggcgcaagtcgcctggggaggagcctgcgtccgattagcttgggtggggtaaaaggcctaccaaggctccgatcggtagctggtctgagagatgatcagccacactgggactgagacagggccagactcctacgggaggcagcagtggggaatattggacaatgggcgaaagcctgataccagcaatgccgcgtgggtgaaagaaggtcttcggattgtaaaagcccttcgacggggacgatgatgacggtaccgtagaagaagccccggtaactcgtgccagcagccgcggtgtaatacgaagggggctagcgttactcggattgactgggcgtaaaaggcgcgtagggcggcttggtaagcagcagtgaaatcctgggctaacctgggggctgcgttgatagctgcgggcttgagttcggagagggcgtggaattcccagtgtaga

>ebp130

agagtttgatcctggctcagaacgaacgctggcggcgtgcttaacacatgcaagtcgagtcgcgctccgtagcaatacggatggcgcacggcgcacgggtgagtaaacacgtgggcaatctgcctccgaaggggaataactccgggaaaccgaaagctaatccgcataacatccctggactccgggtcggggattcaaagcagcaatgcgtcggaggaggagcccgcgccgattagctagtgggtgaggaacggctcaccaaggctatgatcggtagctggtctgagaggatggccagccacactggaactgagacacgggtccagactcctacgggaggcagcagtggggaatcttgcgcaatgggcgaaagcctgacgcagcagccgcgtgggggatgaagcttctcggagtgtaaacccttgcaccgggacgaatgcccgcaagggagtgacggtagcggtagaagaagccccggctaactacgtgccagcagccgcggtgaaagcgtgaaagggccagcgttgtcgaattactgggtgtaaaagggtctagcgggtggcaagtcgggagtgaaatfttgggcttaactcagaacctgctccgaaactgctgtgctcagtggtgggagaggcgcgtggaattgcaggtgtagcgg

>ebp133

agagtttgatcatggctcagaatcaacgctggcggcgcgcttaacacatgcaagtcgcacgagaaaaggaggcaatccatgagtacagtgctgtacgggtgagtaaacacgtggatcatctacctttagtgggggaataaccctgggaaaccggggctaataaccgcataagctcgagagagaaagcagcaatgcgctgagagaggagtccgcggccgattagctagtggcagggtaaaagcctaccaaggcaagatcggtagccggcctgagagggcacacggccacactggcactgaaacacgggcccagactcctacgggaggcagcagtggggaatcttgcacaatgggggaaacctgatgcagcagccgcgtgagcgtgaagcccttcgggggtgaaagccttccggcagggacgatcatgacggtacctgcagaagcagctgcggctaactacgtgccagcagccgcggtgtaatacgtaggcagcagcgttgcgagttactgggcgtaaaagggtgttagccggtgcttaagttggtgtaaatctcccgg

>ebp139

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtggagcaatccatgagtacagtg
gcgaccgggtgagtaaacacgtgaccaacctgccctcagtggggaataactctgggaaaccggagctaataaccgataacatcgaaga
tcaaagccgcaaggccttgaggagggggtcgcggctgattagctagttgtagggtaatggcctaccaaggcgaagatcggtagccgg
cctgagagggcgcacggccacactggaactgaaacacggccagactcctacgggagggcagcagtggggaatcttgccaatggggga
aacctgacgcagcaacgccgcgtggaggatgaagtccttgggacgtaaactccttcgaccgggacgatactgacggctactggaagaa
gaagcaccggctaactctgtgccagcagccgcggtaatacagaggggtgcgagcgtgttcggaattattgggcgtaaaaggcgcgtaggc
ggtgcggtaagtcaactgtgaaatcctggctcaactgggagcctgcaggcgaaactgccgtgctggagtgtgggagaggtgcgtggaa
ttcccggtagcgggtgaaatcgtagatcgggaggaacacctgtggcgaagcggcgcactggaccacaactgacgctgagggcgcg
aaagctag

>ebp140

agagtttgatcatggctcagaatcaacgctggcggcgcgctaacacatgcaagtcgcacgagaaagtggagcaatccatgagtacagtg
gcgtacgggtgagtaaacacgtggatcatctacctttagtggggaataaccctgggaaaccggggctaataaccgataagctcgagagag
gaaagcagcaatgcgctgagagaggagtccgcggccgattagctagttggcagggtaaaagcctaccaaggcgaagatcggtagccgg
cctgagagggcacacggccacactggcactgaaacacgggcccagactcctacgggagggcagcagtggggaatcttgacaatggggg
aacctgatgcagcgacgccgcgtgagcgatgaagccctcgggggtgtaaagctcttcggcagggacgatcatgacggctactgcaga
agcagctgcggctaactacgtgccagcagccgcggtaatacgtaggcagcgcgagcgtgttcggagtactgggcgtaaaagggtgttagg
cggttgcttaagtttggtgtgaaatccccggctcaactggg

>ebp142

agagtttgatcatggctcagagcgaacgctggcggcgtgctaacacatgcaagtcgagtgccgcctgtagcaatacaggtggcgcacg
gcgcacgggtgcgtaaacacgtggcaatctgccctctgatggggaataaccccgcgaaagcggggctaataaccgataacattccttgg
ctacggttcgaggactcaaaggcgtgaagtcgtcggaggaggagcccgcgcacgattagctagttgtagggtaacggcctaccaaggctt
tgatcgttagctggtctgagaggatggccagccacactggaactgagacacggccagactcctacgggagggcagcagtggggaatcttg
cacaatgggcgaaagcctgatgcagcgacgccgcgtgggggatgaagcttctcggagtgtaaaccccttcgaccggacgaatacctc
gcaagaggggtgacggtagcgggtataagaagccccggctaactacgtgccagcagccgcggtaagacgtagggggccagcgttgctcg
gaattactgggtgtaaagggtcgtagcgggtgtggcaagtcgggagtgaaatcctcggggcttactcagagactgctccgaaactgctgt
gcttgagt

>ebp143

agagtttgatcatggctcagagcgaacgctggcggcgtgctaacacatgcaagtcgagtgccgcctgtagcaatacaggtggcgcacg
gcgcacgggtgcgtaaacacgtggcaatctgccctctgatggggaataaccccgcgaaagcggggctaataaccgataacattccttgg
ctacggttcgaggactcaaaggcgtgaagtcgtcggaggaggagcccgcgcacgattagctagttgtagggtaacggcctaccaaggctt
tgatcgttagctggtctgagaggatggccagccacactggaactgagacacggccagactcctacgggagggcagcagtggggaatcttg
cacaatgggcgaaagcctgatgcagcgacgccgcgtgggggatgaagcttctcggagtgtaaaccccttcgaccggacgaatacctc
gcaagaggggtgacggtagcgggtataagaagccccggctaactacgtgccagcagccgcggtaagacgtagggggccagcgttgctcg
gaattactgagtgtaaagggtcgtagcgggtgtggcaagtcgggagtgaaatcctcggggcttactcagagactgctccgaaactgctgt
gcttgagt

>ebp144

agagtttgatcctggctcagagcgaacgctggcggcaggcctaacgcatgcaagtcgagcgggtgtagcaatacatcagcggcagacgg
gtgagtaacgcgtgggaacgtaccttttggttcggaacaaccagggaacttgggctaataaccggataagcccttacggggaaagattat
cgccgagagatcggcccgcgtctgattagctagttgtaggtaatggctcaccaaggcgacgatcagtagctggtctgagaggatgatca
gccacattgggactgagacacggcccagactcctacgggagggcagcagtggggaatattggacaatgggcgcaagcctgatccagccat
gccgcgtgagtgatgaaggccctaggggtgtaaagctcttttgctggggaagataatgacggtagcgaagaataagccccggctaactc
gtgccagcagccgcgtaatacgaagggggctagcgttctcggaaactactgggcgtaaaagggtgcgtaggcgggtctttaagtcaggg
gtgaaatcctggagctcaactccagaactgcctttgatactgaagatcttgagttcgggagaggtgagtggaactgcgagtgtagaggtgaa
attcgtagatattcgaagaacaccagtgccgaaggcggctcactggcccgatactgacgctgagggcagaaag

>ebp145

agagtttgatcatggctcaggatgaatgctggcggcgtggataaggcatgcaagtcgaacggctccgaaggacagtggcagacgaggtg
gtaacacgtaggtactccccgaagcctggtatagctcgtcgaagacggggtaatcccagatggctccgaaagggtaagatttatccg
ttcgggagaggcctgctcatatcagcttgttggtaggtaatggctcaccaggctacgacgtgtagggcgctgagaggcggacccc
accgatgggactgagacacggcccatactctacgggagggcgcagtcgagaattctccgaatggacgaaagtctgacggagcgcag
ccgcgtgctggaagaagtcttcggtagtaaacagctttatgagggacgaaattattgacggtagctcatgaataaggggctcctaactct
gtgccagcaggagcggtaatacagaggcccaagcattatccgggacactggcgtaaaagggtgcaagggcggctatattagctgttcg
aaaatccgtgggctcaactacggctccgagcgaacggatagctagaggttgaagaggtgcatggaactcacggtgtaggggtga
aatccgttgatcgtggggaacaccaaaggcgaaggcagtgactggccaatctgacgctcacaca

>ebp147

agagtttgatcctggctcaggacgaacgctggcggcgtgcttaacacatgcaagtcgaacggaaaggctctctcggagacactcagatgg
caaacgggtgagtaaacacgtgggtgatctgacctgcacttcgggataagcctgggaaactgggtctaataccggataggaccaagagatg
catgtcttttggtggaagcttttgcgggtgtagggatgggcccgcggcctatcagcttgttggtaggggtcacggcctaccaaggctccgacgg
gtagccggcctgagagggtgctccggccacactgggactgagatacggccagactcctacgggagggcagcagtggggaaatattgcaca
atgggcaagcctgatgcagcgcagccgctgggggatgacggccttcgggtcgtaaacctcttcaccatcgacgaaggtccgggttt
ctcggattgacggtaggtggagaagaagcaccggccaactacgtgccagcagccgcgtaatacgtagggtgagcgttgcgggaat
tactggcgtaaaagagctcgtaggtggttgcggttctgtaaaactcacggcctaactgtgagcgtgcccggcgatacgggcagactg
gagtactgcaggggagactggaattcctg

>ebp150

agagtttgatcctggcgcagaacgaacgctggcggcagggcttaacacatgcaggtcagcgcgcccgcaaggggagcggcagacgggt
gagtaaacacgtgggaacgtaccttttggctcggaaactccgggagactggagctaataccggataagccctaacggggaaagatttatc
gccagaagatcggcccgcgtctgattagctagtgtggggtaaacggcccaccaaggcgacgatcagtagctggctgagaggatgatca
ggcactgggactgagacacggcccagactcctacgggagggcagcagtggggaaatattggacaatggcgcaagcctgatccagcca
tgccgcgtgagtgatgaaggccttaggggtgtaaaactctttgtccgggaaagataatgactgtaccggaaagaataagccccggctaactc
gtgccagcagccgcggaatacgaagggggctagcgttctcggaaactactggcgtaaaagcgcacgtagggcgaccattaagtcaggg
gtgagagcctggagctcaactccagaactgcctttgatactgatggtctcagttcggaaagaggttgggtgaaactgcgagtgtagagtgaa
attcgtagatattcgaagaacaccagtggcgaaggcggccaactggtccgacactgacgctgaggtgcgaaagcgtggggagcaaca

>ebp156

agagtttgatcatggctcagagcgaacgctggcggcgtgcttaacacatgcaagtcgagcggcgtagcaatacggagcggcagacggga
gagtaaacacgtgggaacgtgccctcagttcggaaacaccagggaaacttgggctaataccggatacgtccgtaaggagaagatttatc
gctgaaggatcggcccgcgtctgattagctagtgtggggtaaacggcccaccaaggcgacgatcagtagctggctgagaggatgatca
gcctcactgggactgagacacggcccagactcctacgggagggcagcagtggggaaatattggacaatggcgcaagcctgatccagcca
ggcgcgtgggtgatgaaggccttaggggtgtaaaaccttccggcggggaagataatgacggtagccgcaagaagccccggctaact
tcgtgccagcagccgcggaatacgaagggggctagcgttctcggaaactactggcgtaaaagcgcacgtagggcggttttaagtcagg
gggtaaatcctggagctcaactccagaactgccctgatactgagaagcttgagtcgggagaggtgagtggaactgcgagtgtagaggt
gaaattcgtagatattcgaagaacaccagtggc

>ebp162

agagtttgatcatggctcagaatcaacgctggcggcgtgcttaacacatgcaagtcgaacgagaaagtggagcaatccacgagtaaagt
gagaccgggtgagtaaacacgtgactaacctaccctcagtggggaaataacctcgggaaaccggggctaataccgcataatgtcttcggat
caaaggagcaatccgctggaggaggggctcgcggccgattagctagtggcggggtaaacggcccaccaaggccatgatcggtagccg
gcctgagagggcgcacggccacactggaactgaaacacggctcagactcctacgggagggcagcagtggggaaattttgcgcaatggggg
aaacctgacgcagcaacgccgcgtggaggatgaagccccctggggtgtaaaactccttcgatcgggacgattatgacggtagccgatga
agaagcaccggctaactctgtgccagcagccgcgtaatacagagggtgcaagcgttctcggaaattgttggcgtaaaagcgtcgtaggt
gggttgacaagtctgcaagcaaacgggctcaactcgggaaactcgtggaactgctgatcttgaggtatgcagaggatgatggaa

c
>ebp163

agagtttgatcctggctcagaacgaacgctggcggcgtgcctaaacatgcaagtcgagcgagaaagggagcaatcctgagtaaagcgg
cgcacgggtgcgtaaacacgtgggtaatctgcctcgaagagagggataactagccgaaaggtagtaataacctataagcttacggcctctt
cggaggctgttagaaaagggtggcctctgtatacaagctatcctcagatgagcccgcggccatcagctagttggcgggtgaaaagac
caccaagcgatgacggtagctggtctgagaggacgatcagccacactggaactgagacacggtcagactcctacgggagggcagca
gttgggaatattgcgaatggcgaaagcctgacgcagcaacgccgcgtgagcgatgaaggctctcggatcgtaaagctcttcggcgg
gacgaaaactcaccgttcatacgcgggtgactgacgggtaccggaagaagaagcaccggctaactctgtgccagcagccgcggtaataca
gagggtgcaagcgttggtaattattgggcgtaaagcgcgtgtagggcgttgtgtgctgggtgtaaaagccctcggcttaaccgagg
aagtgcgccgaaactagcgagcttgagtgccggagaggagagtggaacacccgggtg

>ebp166

agagtttgatcctggctcagaacgaacgctggcggcgtgcttaacacatgcaagtcgagtcgagcggtgcccgtagcaatacggatggcgca
cggcgcacgggtgcgtaaacacgtgggcaatctgcctccgggtggggaataaccccgcgaaagcggggtaattccgcataacattctctg
gactctggttcggggattcaaagcagcaatgcgccgaaggaggagcccgcgcccattagctagttgggtgagtaacggctaccaagg
caatgatcggtagctggtctgagaggatggccagccacactggaactgagacacggtcagactcctacgggagggcagcagtggggaat
cttgcgcaatgggcgaaagcctgacgcagcgacgccgcgtgggggatgaagcttctcggagtgtaaaccccttcgaccgggacgaatg
cccgaaggagtgacggtaccggtagaagaagccccggctaactacgtgccagcagccgcggtaagacgtagggggccagcgttgtt
cggaaactgggtgtaaagggtccgtagggcgggtggtgcaagtcgggagtgaaattctgggcttaactcagaagctgctccgaaactgct
gtgctagagtgccggagagggcgcgtggaattgcaggtgtagcgggtgaaatgcgtagatatctgc

>ebp169

agagtttgatcatggctcagagtgaaacgctggcggcatgcttaacacatgcaagtcgcacgggcaggggcaaccctgtcagtgggcgac
gggtgagtaacgcgtaggtatctatcctcgggtgggggacaaccacgggaaactgtggctaataccgcatgacgactgaggtccaaagg
cgcaagtcgctggggaggagcctgcgtccgattagcttgttgggtggggtaaaggcctaccaaggctccgatcggtagctggtctgagag
gatgatcagccacactgggactgagacacggcccagactcctacgggagggcagcagtggggaatattggacaatgggcgaaagcctga
tccagcaatcccgcgtgggtgaaaggtctcggattgtaaaaccccttcgacggggacgatgatgacggtaaccgtagaagaagcccc
ggctaactcgtgccagcagccgcggtaatacgaagggggctagcgttactcggattgactgggcgtaaaagggcgcgtagggcgttga
cagtcagacgtgaaattcctgggtcaacctgggggctgcgttgatagctgcgggcttgagttcggaaaggggtcgtggaattcccagtg
agagtgaaattcgt

>ebp401

agagtttgatcatggctcagaacgaacgctggcggcgtgcttaacacatgcaagtcgcgtgagcagttccgtagcaatacggggatgggc
gcacggcgcacgggtgcgtaaacacgtgggcaatctgcctccgatggggaataaccccgcgaaagcggggtaataaccgcataacattc
cttgagcgtggttcgaggattcaaagcaagtaatgcgtcggaggaggagcccgcgcccgattagctagttggcgaggtaacggctacc
aaggcaatgatcggtagctggtctgagaggatgatcagccacactggaactgagacacggtcagactcctacgggagggcagcagtggg
gaatcttgcaaatggcgaaagcctgatgcagcgacgccgcgtgggggatgaagcttctcggagtgtaaaccccttcgctgggacga
atgcccgcaaggagtgacgggtaccagataagaagccccggctaactacgtgccagcagccgcggtaagacgtagggggccagcgtt
gttcggaattactgggtgtaaagggtcgtagggcgggtgcaagttgggagtgaaatcctcgggctcaactcagaggctgctccaaaact
gctgtgcttgagtcgagagaggctcgtggaattgcaggtgtagcgggtgaaatgcgtagatatctgcaggaacacccgtggcgaaagcgg
cgagctggatcactactgac

>ebp402

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaaacacatgcaagtcgaaacgcgagggtcccgaagggactagtagagtg
gcaaacgggtgagtaaacacgtgggtgacctgcctcagtgggggataacgtcccgaaaggagcgtataaccgcatgacatcctccgt
tgaaacgggtggagatcaaagctggggatcgcaagacctagcgttgaagagggggcccgcgtccgattagctagttggtgaggtaatggc
tcaccaaggcgacgatcggtagccggcctgagaggcggacggccacactgggactgagacacggcccagactcctacgggagggca
gcagtggggaattgtcacaatgggcgcaagcctgatgacgcaacgccgcgtggaggatgaagatcttcgggtcgtaaactccttcgatc
gagacgaacgactccgggtaataccggaaagagtgacgggtaccgagagaagaagccccggctaactcctgcccagcagccgcggta
atacggggggggcaagcgttgcggaattactgggcgtaaaagggctcgtagggcggccaactaagtcagacgtgaaatccctcggcttaa
ccggggaactgcgtctgatactggatggcttgaggttgggag

>ebp403

agagtttgatcctggctcaggacaaacgctggcggcgtgcgtaacacatgcaagtcggacgagggggcgggtcttttcggagagcggctt
tcgagtggcggacgggtgagtaaacgctggataacctaccctgggtggtgggggataccctgtagaaatcgggctaataccgcatgcggt
gatcgggggtcggcggatctaagaagccttcggcgcagcaggggggtccgcggccgattagctagtgggggggtaaaggcttcc
caaggcgacgatcggtagctggtctgagaggatgatcagccacactgggattgagaacggcccagactcctacggggggcagcagtg
ggaatcttcgtaatggcgcgaagcctgaacgagcaacgccgcgtggtgggaagacggccctcgggtgtaaaccactttcttgaacga
gtttggacgggtaccagaggaagaagtctcggtaactacgtccagcagccgcggtaatacgtaggagacgagcgttgcggagttact
gggcgtaaagggtgcgacggcgggtcgtttggtgggtgtaaatctcctggctcaactgggagggggccatccaacgggcgaactga
ggcgtggagaggacgtggaattccgggtggagcagtgaaatgcg

>ebp404

agagtttgatcatggctcagagcgaacgctggcggcgtgcttaacacatgcaagtcgaacgcgcagttctctgtagcaatacaggagaccg
gcgcgtggcgaacgggtgcgtaaaccttccctcaggtggggaataaccccgcgaaagcggggctaataccgcataacat
gacgatggcacatgctgttctcaagccgcaaggcggcagaagaggatccgcggccgattagctagtggcgaggaatggtcac
caaggcaatgatcggtagccggactgagaggtgaccggccacactggaactgagacacggtccagactcctacgggagggcagcagtg
gggaatcttcgcaatggcgaagcctgacgcagcaacgccgcgtgggggatgaagcttctggagtgtaaaccctttcaccggga
cgaatggcgcgaagggttgacggtagcgggtataagaagccccggtaactacgtccagcagccgcggtaaaacgtagggggccagc
gttgctcggaaactggcgtaaagggtccgtagggcgggtggtgcaagtcggaagtgaatctctgggcttaactcagaggctgctcccga
aactgctgtgctcagtgagagagggcagtggaattgcgggtgtagcggtgaaatgcgtagatatccgagggacatccgagggcga
agcggctcgtggtacgcaact

>ebp405

agagtttgatcatggctcagaacgaacgtggcggcgtggattagcatgcaagtcgaacgaatccctgggggcaactctggggggaagt
ggcgaaggggcgagtaatgctggcctaacctaccccggggacggggatagccttctaacggaagggtaataccggcgatgtggcga
gggggcatctccttgctatgaaatgagttcgcctcgggaggggggtcacgtgctattagcttgttggcggggtaacggccaccaaggcga
agatggctagcgggtgtgagagcacgaccgcgccactggcactgagacactggccagactcctacgggagggctgcagtcgaggatctt
cggcaatggcgcgaagcctgaccgagcagccgcgtgcgcatgaaggccttcgggtgtaaagcgtgtcaggggggaggaaagc
cgaagggtctgacctatccctgggggaagcacgggctaagttcgtgccagcagccgcgataagacgaaccgtgcgaacgttgttcggatt
cactgggcttaaagggcgcgtaggcggcttgccaagtcaggggtgaaatcttcggctcaaccgaaaacagcttctgatactggcgggct
ggagggaggtaggggcagatggaactccggtggagcgggtgaaatgcgtagatatcgaaggacgccgggtggcgaagcagatctgct
ggacctctctgacgctgatgc

>ebp119

agagtttgatcatggctcagaacgaacgtggcggcgtggattagcatgcaagtcgaacgaatccctgggggcaactctggggggaagt
ggcgaaggggcgagtaatgctggcctaacctaccccggggacggggatagccttctaacggaagggtaataccggcgatgtggcga
gggggcatctccttgctatgaaatgagttcgcctcgggaggggggtcacgtgctattagcttgttggcggggtaacggccaccaaggcga
agatggctagcgggtgtgagagcacgaccgcgccactggcactgagacactggccagactcctacgggagggctgcagtcgaggatctt
cggcaatggcgcgaagcctgaccgagcagccgcgtgcgcatgaaggccttcgggtgtaaagcgtgtcaggggggaggaaagc
cgaagggtctgacctatccctgggggaagcacgggctaagttcgtgccagcagccgcgataagacgaaccgtgcgaacgttgttcggatt
cactgggcttaaagggcgcgtaggcggcttgccaagtcaggggtgaaatcttcggctcaaccgaaaacagcttctgatactggcgggct
ggagggaggtaggggcagatggaactccggtggagcgggtgaaatgcgtagatatcgaaggacgccgggtggcgaagcagatctgct
ggacctctctgacgctgatgc

>ebp406

agagtttgatcctggctcagaatcaacgctggcggcgtgcttaacacatgcaagtcgcgtgcgcagttccgtagcaatacgggatggcgc
acggcgcacgggtgcgtaaacgctgggcaatctccctctgatggggaataaccccgcgaaagcggggctaataaccgcataacattcctt
ggactacggtccggggattcaagcagtgatgctcggaggaggagcccgcgccgattagctagtgtgaggaatggctaccaag
gcaatgatcggtagctggtctgagaggatgatcagccacactggaactgagacacggtccagactcctacgggagggcagcagtggggaa
tcttcacaatggcgaagcctgatgcagcagccgcgtgggggatgaagcttctggagtgtaaaccctttcactgggacgaatg
cccgaaggagtgacggtaccagtagaagaagccccggtaactacgtccagcagccgcggtaaagcgtagggggccagcgttgtt
cggaaacttggtgtaaagggtcgtagcgggtggtgcaagtcggagtgaaatctatgggcttaaccataggtgcttccgaaactgct

gtgcttgagtggtggagagggcgcgtggaattgcaggtgtagcgggtaaatgcgtagatactgcaggaacacccgtggcgaaagcggcg
cgctggaccacaactgacgctgaatgac

>ebp124

agagtttgatcatggctcagagcgaacgctggcggcaggcctaacacatgcaagtcgagcggcgtagcaatacggagcggcagacggg
agagtaaacacgtgggaacgtgcccttcagttcggacaacccagggaacttgggctaataccggatacggcctaaggagaaagattat
cgctgaaggatcggcccgcgtctgattagctagttggtgggtaacggcccaccaaggcgacgatcagtagctggtctgagaggatgatc
agcctcactgggactgagacacggcccagactcctacgggaggcagcagtggggaatattggacaatggcgcaagcctgatccagcc
atgccgatgggtgatgaaggccctagggtttaaagcccttcggcggggaagataatgacggtagccgcagaagaagccccgggcta
cttcgtgccagcagccgcgtaatacgaagggggctagcgttctcggaaactcggcgtaaaagcgcacgtagcggccttttaagtcag
gggtgaaatcctggagctcaactccagaactgccctgatactgagaagcttgagtcgggagaggtgagtggaactgcgagtgtagagg
tgaatcgtagatattcgaagaacaccagtgggcaagggcggctcactggcccggtagcgtgaggtgcgaaagcgtggggagca
aacaggattagataccctggtagt

>ebp412

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaaagtgaggcaatccatgagtatagtg
cgaccgggtgagtaaacacgtgactaacctgcccttgagtgggggataacctgggaaccggggctaataccgataacatcttcaggat
gagaagctggggatcaaggagcaatccgcttatggagggggcgcggctgattagctagttggcggggtaacggcccaccaaggcgat
gatcggatccggcctgagagggcgcacggacacactggaactgaaacacggctccagactcctacgggaggcagcagtggggaaatttg
cgcaatgggggaaacctgacgcagcaacggcgcgtggaggatgaagtccttgggacgtaaacctcttcgaccgggacgattatgac
ggtagccgggtggaagaagccccgggtaacttcgtgccagcagccgcgtaatacagggggggcagcgttctcggaaattattggcgta
aagggcgcgtagggcgtgcgtaagtcacctgtgaaacctctgggctcaactcagagcctgcaggcgaaactgccgtgctggagtgagg
gagaggtgctggaattcccgggtgtagcgggtaaatgcgtagatacgggaggaacacctgtggcgaaa

>ebp414

agagtttgatcctggctcagaacgaacgttagcggcgcctaacacatgcaagtcgcacgagaaaagggcttcggccccgtaaaagtg
cgcacgggtgagtaaacacgtgggtaatgtccctcggcgggggataacctccgaaaggaaaggctaataccgataagaccacggcct
cgaagaggctgaggtaaaagcaggcctctgatgcaagcttcacactgagatcagcccggcggccatcagctagttgtagggtaaatg
gcctaccaaggtaagacgggtgtagctggtatgagaggatgatcagccacactggaactgagacacggctccagactcctacgggaggcag
cagtggggaatcttgcaatggcgaaagcctgacgcagcagcgcgctgggtgatgaaggccttcgggttgaagccctgtgggg
agagacgaataagtcggagccaataacctccggcgatgacggtatttccttagcaagcaccgggtaactctgtccagcagccgcgtaag
acagaggggtgcaaacgttctcggaaactgagcgtgtaagcgtgtagcggcgcgcaagtcggatgtgaaagccccgggctcaac
ccgggaaagtcaccgaaactgcgtggccttgatcgggagaggttggtggaattctcgggtgtagaggtgaaattcgtagatac

>ebp418

agagtttgatcatggctcagaacgaacgttggcggcgtggattaggcatgcaagtcgaacgaatccctgggggcaactctggggggaagt
ggcgaagggggcagtaatgcgtggctaactacccggggacggggatagccttcctaacggaagggtataaccggcgatgtggcga
gggggcatctccttctatgaaatgagttcgcctcgggagggggtcacgtgctattagcttgttggcggggtaacggcccaccaaggcga
agatggctagcgggtgtagagcacgaccgcgcaactggcactgagacactggcagactcctacgggaggctgcagtcgaggatctt
cggcaatggcgcgaagcctgaccgagcagcgcgctgcgcgatgaaggccttcgggttgaagcgtctcaggggggaggaagc
cgaaggtctgacctacccctgggggaagcaggggtaagttcgtgccagcagccgcgataagacgaaccgtgcgaacgttctcggatt
cactgggcttaaaggcgcgtagggcggcttccaagtcaggggtgaaatcttcggctcaaccggaaaacagcttctgatactggcgggct
ggagggaggtaggggcagatggaactcgggtggagcgggtgaaatgcgtagatac

>ems2

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagtaagtggagcaatccatgagtacagtg
cgtacgggtgagtaaacacgtgactaacctacctccagtgggggaataacccgggaaccggggctaaccgcataaaaccgcaaggt
caaatgcgcaagcagcttggagagggggcgcggttgattagctagttggtgaggtaatggctaccaaggcgatgatcaatagccggcc
tgagagggcgcacggccactggaactgaaacacggctccagactcctacgggaggcagcagtggggaattttgcgaatgggggaaa
ccctgacgcagcaacggcgcgtggaggatgaagtccttgggatgtaaacctcttcgatcgggacgataatgacggtaccggaagaaga
agccccggtaactctgtccagcagccgcgtaatacagagggggggcagcgttctcgggattattggcgtaaaaggcgcg

>ems6

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaaaggagcaatccatgagtaaagtgcgaccgggtgagtaacacgtgactaacctaccttcgagtggggggataactccgggaaaccggagctaataaccgcataacatcgcgagatcaaagcagcaatgcgctggaagagggggcgcgggctgattagctagttggcggggtaacggcccaccaaggcgaagatcggtatccggcctgagagggcgcacggacacactggaactgaacacgggtccagactcctacgggagggcagcagtggggaatgtgcgcaatggggga aacctgacgcagcaacgccgcgtggaggatgaagcccttggggcgtaaactccttcgaccggacgatgatgacgggtacgggtgga agaagcaccggctaactctgtgccagcagccggtaatacagagggtgcaagcgttgttcggaattattgggcgtaagggggcgcgtagg cggcgcgtaagtacactgtgaaatctcccggcttaactgggagcctgcaggcgaactac

>ems7

agagtttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgtctctgggtgcttggtagccagaggagtgcggacgggtgaggaacacgtgggtgacctatctttagtgaggaataccggcgagaaatgccgacaagaccgatacggttctctcgtcggagagggaaacgaaagctctcgtgagaggggcgtggaagctgggctgcgtccgattagctggctgggggtaaaaggctacca gggcgacgatcggtagctggtctgagaggacgatcagccacactgggattgagaacggcccagactcctacggggggcagcagtgag gaatttctggcaatgggggcaacctgaccgagcaacgccgcgtgcaggaagacggccttcgggtgtaaaactgctttctgtgggacgag caaggacggtaccacaggaagaagccccggctaactacgtgccagcagcctcggtaatacgtagggggcaagcgttccggagtact gggcgtaaaaggcagcagggcgttccacaagagtggggtgacaggtgccggcttacctggcacaggccttcagactgtgagacttg agggcctcagagggacatggaactccaggtgtagtggtgaaa

>ems12

agagtttgatcatggctcaggatgaacgctggcggcgtgcctaacacatgcaagtcgaacgtctctggccgcaaggctggaggagtgccggacgggtgagtaacacatgggaacctacctccagtgaggaataccggcgagaaatgccgacaagaccgatgggcttctccttgg gggaaagcaaaagtttctcgcgagaggagacgctggaagatgggctgtggccgattagctagctgggtgggtaagagcctaccagggcg acgatcggtagctggtctgagaggacgagcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaatctt ccgcaatgggcgcaagcctgacggagcaacgccgcgtgcaggaagacggccttcgggtgtaaaactgctttctgagggacgagcaagg acggtacctcaggaacaagccccggctaactacgtgccagcagccgcgtaatacgtagggggcagcgttgcgggagtactggggc taaagggcatcagggcggagcgtcgcttggcatgacagctcctcgttcacagaggagagagtggaacaaacgagcgcacttgaggggacacagaggacacagaattcgggggtggagtgtgaaatgcgtagagatccggagggaataccea

>ems13

agagtttgatcatggctcaggacgaacgctggcggcgtgcctcaaacatgcaagtcgagcgtggcttcgttctcgcgaagagaggacga agtctagcggcgaacgggtgagtaaacacgtgagaaacctgccctccggacggggacaacctcgggaaaccggggtaataaccggattc gctcgggggtcgcagcctggtgaggaagagcgaagtcgccggaggagggtctcggcctatcagcttgggtggggtaaacgg cctaccaaggctatgacggtagctggcctgagaggggtggtcagccacactggaactgagacacgggtccagactcctacgggagggcag cagttgggaatctgggcaatgggcgaaagcctgaccagcagcgcgtgggggatgaagccttcgggtgtaaaccttgcagatga gggacgaagccgcaaggtgacggtacctcaggaggaagccccggctaactacgtgccagcagccgcgtaagacgtagggggcaag cgttgcggaatcactgggcgtaaaagcgcgtgtagggggttggccagtcctcgtgaaagcctccggctcaactggaggacgccggg gatactgccggactcaggacaggagagggaagtgaattcccgggtgta

>ems80

agagtttgatcatggctcagaacgaacgctggcggcgcgcctaatacatgcaagtcgaacgagaaatgtgtagcaatacacatgtacagtggcgaacgggtgaggactacatggtaacctaccctcagtgagggaataatagccgaaaggttagtaataaccgctacgcttcccggact gcggttcgggaaggaaagcagatacctgggtatcgcgctcttgatgggctcatgccctatcagcttgggtgaggtaacggctaccaaa ggcttcgacgggtgagctgatctgagaggatgatcagccacactggcactgcgacacggccagactcctacgggagggcagcagtaagg aatattgcgcaatgggcgaaagcctgacgcagcagcgcgtgggggatgaaggtcttcggattgtaaaccttctgggaggggaagat ggggatggggtaaccattcggacggtacctccagaagcagccacggctaactcgtgccagcagccgcgtaatacgaaggtggcaagc gttgttcggattactgggcgtacagggagcgtagggcgttgggtaagccctcgtgaaatctccaggttaactggaaagtgcagaggg gactgctcagctagaggatggggggag

>ems88

agagtttgatcatggctcaggatgtacgctggggcatgcttaacacatgcaagtgaaacgggtfataagtgaaactttaagtgaattcttgta
atcagtagcggacgggtgcgtaacgcgtaagaacctacctttaggtggagaataactactggaaacgggtgtaatactccatagctgag
gagtgaaagagaaaacaacttcttgcctagggatgggcttgcgctgattagcctgttgtagggtaacggctccaaggcgacgatca
gtagctggctgagaggacgatcagccactgggactgagacacggcccagactcctacgggagggcagcagtgaggaattttccgcaa
tgggcgaaagcctgacggagcaatgccgctggaggaagacagcctatgggctgtaaaccttttctcagagaagaattctgacggat
ctgaggaatcagcatcggtaactccgtgccagcagccgcgtaatacggggatgcaagcgttatccggatgattgggcgtaagggt
ctgtaggtgctggtaaaagtctaccgtcaataccaaggcttaacctggggcagcgggtgaaactcatcggtgagtagcgttagg
>ems106

agagtttgatcatggctcagattgaacgctggcggcatgcttaacacatgcaagtcgaacggcagcatttatttagttactagatgagatggc
gagtggcgaacgggtgagtaacgcgtaggaatatgcttagagactgggacaacttgggaaaccaagctaataccggatgccatctat
ggatgaaagctggggaccttgggctggcgtttaaagaatagcctgctgacgattagctagttggtgggtaaggccctaccaaggcga
cgatcgtagctggtctgagaggatggcagccactggaactgagacacggctccagactcctacgggagggcagcagtggggaatattg
gacaatggggggaacctgatccagcaatccgctgtgtgaagaaggcctgagggtgtaaagcactttcgtagggaggaggactgt
gggtaagagctatcagttggacgttacctacagaagaagcaccggtaactccgtgccagcagccgcgtaatacggagggtgcgag
cgtaat
>ems107

agagtttgatcctggctcaggacaaacgttggcggctgcttagacatgcaagtcgagcgtgaaagcccgaagggtgagtaaagcggc
gaacgggtgagtaatacttgggtaacgtccctcctagattgggataacgcgccgaaaggcgtgtaataccggatagaccacggcttct
ggaagttgggtaaaagatggcgcaagctatcactgggagatcggtcaagcaccattagctagttggcgtgtaacggcccaccaaggc
gacgatggttagccggcctgagagggtgatcggccactgggactgagatacggcccagactcctacgggagggcagcagttggaata
ttccacaatggcgcaagcctgatggagcagaccgcgtgaaggatgaaggtcttcggattgtaaactctttagacctagatgaaggctgc
agggtgaataatcctgcagttgacagtatagtagaataagccacggtaactctgtgccagcagccgcgtaatacagaggtggcgaac
gttgcctggatttattgggtgtaaagggcagtaggtggtt
>ems108

agagtttgatcctggctcagaactaacgctggcggcgtgcctaacacatgcaagtcgagcggacaccaattttagcaatacagatgaagt
gttagcggcagacgggtgagtaacacgttggcaatctatcttccgaccgggataactagccgaaagggttagctaatacctgatacgatcatc
gagcggcatcgcgatgatgaaagcagcaatgcgcggattgaggagcctgcggtccattagctagttggtgaggtaacggcccaccaa
ggcaacgatggatagctggtctgagaggacgaccagccactgggactgagacacggcccagactcctacgggagggcagcagtgag
gaattttccgcaatggcgaaagcctgacggagcagccgcgtgcacgaggaaggtcttcggatcgtaaagtctgctgaggggggacg
attatgacggtaccctggaggaaagccccggtaactacgtccagcagccgcgtaagacgtagggggcagcgttcttcggaattact
gggcgtaaaagcgtgcgtagggcggattcgtaaagtcgaggttgaata
>ems109

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagcgggggcaactcctgagtaaa
gtggcgtacgggtgagtaacacgtagctaacctaccctcagcgggggataactggggaaactcgggtaataaccgataacctcttagtt
gcacaagcagcagagagcaaaaggtccgaaggatcgtgaaggagggggctgcggccgattagctagttggcgggtaacggcccac
caaggcaatgatcggtagccggcctgagaggcgcacggccactggcactgaaacacggccagactcctacgggagggcagcagt
ggggaattttgcacaatcccgaagggtgatgcagcaacgccgctgagggatgaaggccttcgggctgtaaacctcttgcaccggga
aaaatgggctggtgaatagccagctctgatggtaccgatggaagaagccccggccaactacgtccagcagccgcgtaatacgtaggg
ggcagcgttctcgaattactgggcgtaaaaggcgtgtaggcggtgcgtaagtggggcgtgaaatcccggagcctaactcggggcg
gtcgttcagactccgtgctagggt
>ems110

agagtttgatcatggctcagaatcaacgctggcggatgcttaacacatgcaagtcgaacgaaccattcgtggtgagtgccggacgggtg
agtaacgcgtgagaatcgtgctcaggtcggggacaaccattgaaacgatggctaataccggataagccgcaaggtgaaagattaattgc
ctgaagatgagctcgcgtctgttagctagttggtgggtaagagcctaccaaggcgacgatcagtagctggtctgagaggacgatcagcc
aactgggactgagacacggcccagactcctacgggagggcagcagtggggaattttccgcaatggcgcaagcctgacggagcaatac
cgcgtgggggaggaaggctcttgggtgtaaaccttttctcagggaagaagaatgacggtacctgaggaatcagcatcggctaactcc
>ems111

gtccagcagccgcggaatacaggagatgcaagcgttatccggaatgattgggcgtaaacgctccgaggtgtagtcaagtctgctg
ttaaaggtttggcctaaccaagatcggcgggtgaaactgcacaactagagtgaggtaggggtagaggggattcctggtgtagcg
>ems111

agagtttgatcctggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcctttctctcgtaaaggaaggaggagtg
gaggacgggtgagtaacgcgtgggtacctgcactttggtgaggaatacaggcgagaaatcgccgacaataccgcatgagctctcgtgag
agagcaaagcttcgcatcggcggggccagagaatgggcctgctccgattagctagtgggtgaggtaaaggcttaccaggcgacga
tcggtagctggtctgagaggacgatcagccacactgggattgagaacggccagactcctacggggggcagcagtgaggaatttcggc
aatgggcgcaagcctgaccgagcaacgccgctgcaggaggaaggtcttcgattgtaaactgctttccagggaagagtaaggacgg
tacctggggaataagtcccggctaactacgtgccagcagccgcggaatacgtagggggcgagcgttatccggagttactggcgtaaaag
cgacgtaggcggctgctaagttcgagtgacagtcaccagcttaactggtgaagtactttgaaaactggcggacttgagagc
>ems112

agagtttgatcctggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtggagcaatccatgagtacagtg
gctacgggtgagtaacacgtgactaacctacctccgagtggggaataaacccgggaaaccggggctaaccgcataaaaaccgcaag
gtcaaatgcgcaagcagcttgagaggggggtcggggtgattagctagtgggtgaggtaatggctaccaaggcgatgatcaatagccgg
cctgagagggcgacggccacactggaactgaaacacggtccagactcctacgggagggcagcagtggggaattttgcgcaatggggga
aacctgacgcaacgccgctggaggatgaaaccctttggggcgtaaaactccttcgaccggacgataatgacggtacgggtgga
agaagcaccggctaactctgtgccagcagccgcggaatacagaggggtgcaagcgttcttcggaattattggcgtaaaaggcgctagg
cggcgcaagtaagtcacctgtgaaatctctggcttaactgggagcctgcagggcaaacctaccgtgctggagtggggagaggtgctggga
attcccgggtgtagcgggtgaaatgcgtagatcggg
>ems145

agagtttgatcatggctcaggacaaacgctggcggcgtggattaggcatgcaagtcgaacgaatcccagccttgtagcaataccgggctg
ggggaagtggcgaaaggggagtaaggcgtgggtaacctaccggggacgcggatagccattctaacgagtgaggtaatacgcggcg
attcagcgtgatggcatatcgtgctgggaaatgaatttcgctcgggaggggctcacgtgatattagcttgttggcggggtaacggcccac
caaggcgaagatgtctagcgggtgtgagagcagaccgcgccactggcactgagacactggccagacacctacgggtggctgcagtc
gaggatcttcggcaatggcgcaagcctgaccgagcagccgcgctgtgcgatgaagccttcggggtgtaaagcactgctgagggggga
taaagcccgaagggttgattgatccctggaggaagcacgggctaagttcgtccagcagccgcgtaagacgaacctgcaacgttct
tcggaatcactgggctaaagggcgctagggcgcttcgcaagtcgggggtgaaagccccagctcaactgggggaagtgccttgatact
gaggagctcgag
>ems133

agagtttgatcatggctcaggataaacgctggcagcgtgcctaacacatgcaagtcgaacgctctagctcgaagaggtagggagtgccg
gacgggtgaggaacacgtgggaatcgacctgaggtgaggaataccaccgggaaaccggtgacaataccgcatagccttggcagcaag
gaaagctcgaagagcggcacaggacgagcctgcgcccattagttagctggtgaggtaatggctcaccaggacgacgatcggtagctg
gtctgagaggacgatcagccacactgggattgagaacggccagactcctacggggggcagcagtgaggaatttcgtcaatgggggca
acctgaacgagcaacgccgctgcaggaagaaggttttcggatcgtaaactgctttctctgggacgagtagggacgggtaccagaggaa
caagccccggctaactacgtgccagcagccgcggaatacgtagggggcgagcgttctccggagttactggcgtaaaagggtccgaag
gcggtgtgcagcgttccacgtgacagctcggctcaactggagaaggtggtgaaaacgagcagacttgaggcgctgagaggaaagcg
gaattccgggtggagtggtgaaa
>ems134

agagtttgatcctggctcagattgaacgctggcggcacgcctaacacatgcaagtcgagcggtaacatgattagtagcaataactagttgatg
acgagcggcggaggggtgaggaacacgtaggaacctaccttcagagggggactaccaagggaactttggctaataccgcataatctc
gagagaggaagtgtcgaaggggtgacactgaaagagggcctgctccgattagctagtgggtgaggtaaaggctcaccaaagcga
tgatcggtaactggtctgagaggacgaccagtcacactgggactgagacacggccagactcctacgggagggcagcagtggggaactt
ggacaatgggggcaacctgatccagcagatccgcgctgggtgagaagggccttcgggtgtaaagccctttaggttgggacgaagtgtgt
agaggaaatgctataacattgac
>ems135

agagtttgatcatggctcaggatgaacgctggcggcgtgcctaacacatgcaagtcgaacgtcctctggccgcaaggctggaggagtg
cggacgggtgagtaacacatgggaacctaccctcagtgaggaataccggcgagaaatgccgacaagaccgatgggcttctccttgg
gggaagcaaaagtttctcgcgagaggagacgctggaagatgggcctgtggccgattagctagctggtgggtaagagcctaccaggcg
acgatcggtagctggtctgagaggacgagcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaat
ccgcaatgggcgcaagcctgacggagcaacgccgcgtgcaggaagacggccttcgggttgtaaactgctttctgagggacgagcaagg
acggtacctcaggaacaagccccggtaactacgtgccagcagccgcgtaatacgtagggggcgagcgttccggagtactggg
taaagggcatgcaggcggagcgtcgtttgcatgacagctcctcgttcacgaggagagagtggaacaaacgagcgcac

>ems139

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtggagcaatccatgagtacgtg
gcgacgggtgagtaacacgtgactaacctacccttgatgggggataacttcgggaaaccgaggtaataaccgataacaccgagagg
caaaggagcaatcgtcagggagggggctcggctgattagctagttggtgaggaatggctcaccaggcgatgatcagtatccggcc
tgagaaggcgcacggacacactggaactgaaacacggctccagactcctacgggaggcagcagtgagggaatfttgcgcaatgggggaaa
ccctgacgcagcaacgccgcgtggtggatgaagtcccttgggacgtaaactccttcgatcgggacgatgatgacggtaccgagagaaga
agccccggtaactcgtgccagcagccgcggtaatacagggggggcgagcgttctcgaattattggcgtaaagggcgcgtagggc
gttcggttaagtctccagtgaaatctatgggctcaaccatagctgctgggggaaactgccgggcttgatgtgggagaggtgagtggaat
cccgtgt

>ems140

agagtttgatcatggctcaggatgaacgctggcggcgtgcctaacacatgcaagtcgaacgtcctctggccgcaaggctggaggagtg
cggacgggtgagtaacacatgggaacctaccctcagtgaggaataccggcgagaaatgccgacaagaccgatgggcttctccttgg
gggaagcaaaagtttctcgcgagaggagacgctggaagatgggcctgtggccgattagctagctggtgggtaagagcctaccaggcg
acgatcggtagctggtctgagaggacgagcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaat
ccgcaatgggcgcaagcctgacggagcaacgccgcgtgcaggaagacggccttcgggttgtaaactgctttctgagggacgagcaagg
acggtacctcaggaacaagccccggtaactacgtgccagcagccgcgtaatacgtagggggcgagcgttccggagtactggg
taaagggcatgcaggcggagcgtcgtttgcatgacagctcctcgttcacgaggagagagtggaacaaacgagcgcacttgaggga
cacagaggacaca

>ems141

agagtttgatcctggctcaggatgaacgctgggggcatgctaacacatgcaagttgaacgggtataagttgaaactttaagtgaattctgta
atcagtagcggacgggtgcgtaacgcgtaagaacctacccttagtgaggagaataactactggaaacggtggtaatactccatgctgag
gagtgaaggggaaagacaacttctgcctagggatgggcttgcgtctgatcagcttgttgtagggtaacggctcccaaggcgacgatca
gtagctggtctgagaggacgatcagccacactgggactgagacacggcccagactcctacgggaggcagcagtgaggaatfttccgca
tggcgaaagcctgacggagcaatgccgcgtggaggaagacagcctatgggtcgtaaactcctttctcagagaagaattctgacggtat
ctgaggaatcagcatcggctaactccgtgccagcagccgcgtaatacgggggatgcaagcgttatccggaatgattgggcgtaaaggg
ctgtaggtggtcggtaaggtctaccgtcaatacgaagcctaacttggggcagcgggtgaaactatcggcttgatcggtaggggca
gaggggaattccggcgtagcgggtg

>ems143

agagtttgatcctggctcagagcgaacgctggcggcgtgctaacacatgcaagtcgaatgcacagcattctgtagcaatacagagcgtg
gcgatggcgaacgggtgcgtaacacgtgggcaaccttccctcgggtggggataccccgcgaaagcgggactaataaccgataaac
tcccttggttagttctgagggattcaaacgagcaatgcgccggaagagggggcccggccgattagctagttggtgaggaacggctcacc
aaggcaaatcgtgtagccggtctgagaggatgatcggccacactggaactgagacacggctccagactcctacgggaggcagcagtg
ggaatctgcacaatgggcaagcctgacgcagcagccgcgtgggggatgaagcttttcggagcgtaaaccttctgacccggac
gaatccccgcaaggcctgacggtacgggtagaagaagccccggtaactacgtgccagcagccgcgtaaacgtagggggccagc
gttctcggaaactgaggcgtaaaggtctgtaggcgggtcggcaagtcggaagtgaatctccgggcttaactcgggggctgctctcga
aactgccgtgctcagtgtagagagggcagtggaattgcaggtgtagcgggt

>ems144

agagtttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgtcctctgggtgcttggtagccaggggagtg
cggacgggtgaggaacacgtgggtgacttatcttccagtgaggaataccggcgagaaatgccgacaagaccgatacgggttctcctcgt

cggagaggggaacgaaagctcttcgtgagaggggctggaagctggcctcgtccgattagctggctggtgggtaaaagctcacca
gggcgacgatcggtagctggtctgagaggacgatcagccacactgggattgagaacggccctgactcctacggggggcagcagtgagg
aatttcggcaatgggggcaacctgactgagcaacgccgctgcaggaagacggcctcgggtgtaaactgctttctgtgggacgagc
aaggacggtaccacaggaagaagccccggctaactacgtccagcagccgcggtaatacgtagggggcaagcgttccggagttact
gggcgtaaaagggcacgcaggcggttcataagagtggggtgacaggtgccggcttacctggcacagggccttgcaactgtgagactg
agggcc

>ems146

agagtttgatcatggctcagaatcaacgctggcggcgcgctaacacatgcaagtcgaacgagaaagtggagcaatccatgagtacgtg
gctacgggtgagtaaacacgtgactaacctacctccgagtggggaataacccccgggaaaccggggtaaacaccgcataaaaccgcaag
gtcaaatgccaagcagcttgagagggggctcggctgattagctagttggtgaggtaatggctaccaaggcgtgatcaatagccgg
cctgagagggcgcacggcctaactggaactgaaacacggctccagactcctacgggaggcagcagtggggaatttcgcaatggggga
aacctgacgcagcaacgccgcgtggaggatgaagtcctttgggatgtaaactccttcgatcgggacgataatgacggtaccggaagaa
gaagccccggctaactctgtgccagcagccgcgtaatacagagggggcaagcgttggcgaattattggcgtaaaagggcgcgtagg
cggctgcgtaagtcacctgtgaaatctctgggctcaactcagagctgcagggcgaactgccgtgcttgagtgtgggagaggtatgtggaat
tccgggtgtagcgg

>ems405

agagttgacctggctcaggattaacgctagcggcaggcctaatacatgcaagtcgaggggcagcaggggtgtagcaatacatcgtggcg
accggcgcacgggtgcgtaaacacgtatgcaatcttcttaactggggaatagctttgagaaatcagaattaatcccataacattgcgaaa
tggcatctttgcaattaaagctccggcggftaagggtgagcatgcgtccattagctagttggcagagtaacggctcaccaaggcgacga
tgggtaggggttctaagaggatttaccacacgggaactgagacacggctccgactcctacgggaggcagcagtaaggaatattggtc
aatggacgcaagtctgaaccagccatgccgcgtgcaggatgaaggtcctatggattgtaaactgctttgtacgggaaaaaacctcgggttt
ccaccgagttgatggtactgtaagaataaggatcggctaactcctgccagcagccgcgtaatacggaggatccaagcgttatccggat
tactgggttaaaagggtgcgtaggtggaatgacaagtcagttggtaaaacctgcagcttaactgtagaactgccattgatactgttattctga
gtacatttgatgtgagtggaatgtgtcgtgtagcggggaaatg

>ems414

agagttgatcctggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtctccgcaaggggatgagtaaag
tggcgacgggtgagtaaacacgtagctaactacctcagagtgggggataacctggggaaactcgggtaataaccgcataacgtcctggc
ggcgcgagtcgctgggatcaaaagaccgcaagggtcgtcgtcagggggctcggccgattagctagttggtgaggtaacggcccac
caaggcgtatgctgtagccggcctgagagggcgcacggccacactggcactgaaacacgggccagactcctacgggaggcagcagt
ggggaatttgcacaatgcccgaagggtgatgcagcaacgccgctgagggatgaaggtcttcggatcgtaaacctcttcgaccgaga
aaaacggcgggtgaagagtcctgctctgatggtaccgggtggaagaagccccggccaactacgtccagcagccgcgtaatacgtagg
gggcgagcgttctcggaaactgggcgtaaagggcgtgtaggcggctcggtaagtcgggcgtgaaatccccgggcttaactcggggc
ggtcgttcgaaactgccgtgctagagtgacggagaggggaagtggaattccc

>ems415

agagttgatcatggctcagaatcaacgctggcggcgcgctaacacatgcaagtcgcacgagaaagtggagcaatccatgagtacgtg
gctacgggtgagtaaacacgtggatcatctgcctcttgggtggggaataacctgggaaaccggggtaataaccgcataagcctgtaaagg
gaaagcagcaatgccgagagaggagtcggcggcggattagctagttggtagggtaatggcctaccaaggcaatgatcggtagccggc
ctgagagggcagcagggcactggcactgaaacacgggccagactcctacgggaggcagcagtggggaattctgcacaatggggga
aacctgatgcagcagccgcgtgagcgtgaagccctcgggggtgaaagctcttcggcaggaacgataatgacggtacctgaagaa
gaagctcgggtaactacgtgccagcagccgcgtaatacgtaggcaggtt

>ems417

agagttgatcctggctcagaacgaacgctggcggcatgcctaacacatgcaagtcgaacgcctgtagcaatacaggagtggcgcacgg
gtgagtaacgcgtgggaacctgcctggagttgggataacccggcgaaagccgggtaataaccggataccgccccgaggggggaagat
ttatcgtctaggtgagcccgcgtcggattagctagttggtagggtaaaagcctaccaaggcgacgatccgtagctggtctgagaggatg
atcagccacactgggactgagacacggcccagactcctacgggaggcagcagtggggaatattggacaatggggcgaagcctgatcca
gcaatgccgcgtgagtgatgaagccttaggggtgaaagctctttaccgggatgataataacagtaccgggagaataagccccggccta

actccgtgccagcagccgcggaatacggaggggtagcgttgctcgaattactgggcgtaaagcgcacgtagggcgtttgtaagtta
gaggtgaaagcccagggttaaccctggaattgccttaagactgcatcgctagaatcgtggagaggtgagtggaattccgagtgtagagg
tgaattcgtagatattcggaagaaca

>ems449

agagtttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcccttcgcaagaagggagtgccggacgg
gtgagtaacacgtggataacctgccccgaggtgaggaataccggcgagaaatcgccgacaagaccgcataggctcagcaatgagcaaa
gcagcgtaacgagcggccacgggatgggtccgcgtccgattagctcgttgtagggtaatggcctaccaaggcgacgatcggtagctggc
ctgagaggacggctcagccacactgggattgagaacggcccagactcctacgggagggcagcagtgaggaatttcggcaatggcgcgca
gcctgaccgagcaacgccgcgtgcaggaagaagctcttcggagtgtaaactgctttaccagggacgagcaaggacgtacctggtgaat
cagccccggctaactacgtgccagcagccgcggaatacgtagggggcaagcgttgctccggaattactgggcgtaaagcgcacgagg
cggtctgtcaagttcagataacagtcaccagcttaactggcgaagggcctcgaaaactgatggacttgagagccagagagggacctgga
attccgggtggagcagtgaaatgcgtagagatccggaggaatacaaaaggcgaaaggcagggctctggctggtatctga

>ems402

agagtttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgtcctctggccgagggctggagggagtg
cggacgggtgagtaacacatgggaacctacttccagtgaggaataccggcgagaaatcgccgacaagaccgcatggccttctctttgg
ggaagcaagtttctcgcgagaggagacgctggaagatgggcctgtggccgattagctagctggtgggtaagagcctaccagggcga
cgatcggtagctggtctgagaggacgagcagccacactgggattgagaacggctcagactcctacggggggcagcagtgaggaatttc
cgcaatggcgcaagcctgacggagcaacgccgcgtgcaggaagacggccttcgggtgtaaactgctttctgagggacgagcaagga
cggtaacctcaggaacaagccccggctaactacgtgccagcagccgcggaatacgtgtt

>ems115

agagtttgatcatggctcaggatgaacgctggcggcgtgccttaggcatgcaagtcgaacgggctgcggtcgcaagaccaaagtcatgg
cagacgggtgagtaacgcgtaggaatctaccccagagtgagcatagccccgcgaaagcgggggtaattctcatggtcccgaagggt
aaaggccttaactggccgctttgggagggagcctgcgtcctatcagctagttggtgaggaatggcttaccaggcaatgacgggtagctgg
actgagaggtcagatcagccacaatcgactgagacacgggcagtacacctacgggtggcagcaaccaggaatattgggcaatggacga
aagtctgaccagcagcggcgcgtggaggaagacggccttcgggtgtaaactcctttgccaagctcttcacagcttgagaataagcctc
tactaactacgtccagaagtctcggtagcgtagggaggaagcgttatccggatacattgggcgtaaagtgtgcgtaggcgtcttaccaa
gtcttactta

>ems117

agagtttgatcatggctcaggatgaacgctggcggcgtgccttaggcatgcaagtcgaacgggctgcggtcgcaagaccaaagtcatgg
cagacgggtgagtaacgcgtaggaatctaccccagagtgagcatagccccgcgaaagcgggggtaattctcatggtcccgaagggt
aaaggccttaactggccgctttgggagggagcctgcgtcctatcagctagttggtgaggaatggcttaccaggcaatgacgggtagctgg
actgagaggtcagatcagccacaatcgactgagacacgggcagtacacctacgggtggcagcaaccaggaatattgggcaatggacga
aagtctgaccagcagcggcgcgtggaggaagacggccttcgggtgtaaactcctttgccaagctcttcacagcttgagaataagcctc
tactaactacgtccagaagtctcggtagcgtagggaggaagcgttatccggatacattgggcgtaaagtgtgcgtaggcgtcttaccaa

>ems113

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtggagcaatccatgagtaaagt
gcgaccgggtgagtaacacgtgaccaacctacctcaagtgggggaatagctccgggaaaccggggctaataccgataacaccgcaag
gtcaaagcagcaatgcgctgggagagggggcgcggctgattagctagttggcggggtaacggcccaccaaggcgaagatcggatcc
ggcctgagagggcgcacggacacactggaactgaaacacggctcagactcctacgggagggcagcagtggggaatttgcgcaatgggg
gaaacctgacgcagcaacgccgcgtggaggatgaagccctttggggcgtaaactccttcgaccggacgataatgacggtaggggtg
gaagaagcaccggctaactctgtgccagcagccgcggaatacagagggtgcaagcgttgctcgaattattggcgtaaagggcgcgta
ggcgggtgcgtaagtcacctgtgaaatctccaggcttaacttgagcttgaggcgaactcccgtgctggagtgaggagaggtgcgtg
gaattcccgggtgtagcgggtgaaatgcgtagatattcgggaggaacacctgtggcgaaagcggcgc

>ems85

agagtttgattatggctcaggacaacgctggcggcgtgcgtaacacatgcaagtcgaacgcccttcacgcgagtgaggaggagtgccgg
acgggtgaggaattcatgggtatctacctcgggtgggggataccggcgagaaatcgccgacaatccagcatagggtcacggagacgtg

agcaaagccgagagggcggggagatgagcctgtggccgattagctagctggtgggtaagggtcaccagggcgacgatcggtagct
ggtctgagaggatgatcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaatctccgcaatggacgc
aagtctgacggagcaacgccgctgacaggagcgttttcggagtgtaaactgcttttctccggacgagtacggacgggtacggcagga
agaagccccggctaactacgtgccagcagccgcgtaaacgtagggggcagcgttgcgggagtactggggtaaagggcgta
ggcgggtgtctggcgtagcggcggacagccgggggcttaacctccggagcggcgtcgtaaacggggagactggagggtccagaggaac
tgggtagctccgggtggagcgggtgaaatgcgtagagaccggaggaacaccaagggcgaaggcgaggttctgggggagtcactgacgct
gaggcgcgaaagcc

>emp11

agagttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcccttagtaataaggagtgccggacggg
tgaggaaacagtgaggatctacccaaaggtgaggaataccggcgagaaatcgccgacaataccgcataggttcttcggaacaaagctc
gcaagagcgcattggatgagcctgcggccgattagttgttgggtgggtaaatggcctaccaagggcagcagcgtgtagctggtctgagag
gacgatcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaatttcgtcaatgggggaaacctgaac
gagcaacgccgctgacggaagaaggttttcggatcgtaaactgcttttctgggacgagaacggacgggtaccagaggaatcagccccg
gctaactacgtgccagcagccgcgtaatacgtagggggcaagcgttgcgggattactggggtaaagcgcacgcagcgggtctcatg
cgtcttgagtacaggtcccggctcaactggggcagttctgaaacgggtgagactgtagcttcagaggggacgcgaatttcgggtgga
gcggtgaaatgcgtaaagatctggaagaacaccgatggcgaaagccacgttctggaaagcttctgacgctcaa

>emp418

agagttgatcatggctcagagcgaacgctggcggcaggcctaacacatgcaagtcgagcggcgtagcaatacggagcggcagacggg
agagtaaacagtgaggaaacgtgccctcagttcggaaacactcagggaactgggctaataccgaatacgtccgtaaggagaaagattat
cgctgaaggatcggcccgctctgattagctagttggtgggtaaatggcccaccaagggcagcagcagtagctggtctgagaggatgatc
agcctcactgggactgagacacggcccagactcctacggggaggcagcagtgggggaatattggacaatgggcgcaagcctgatccagcc
atgccgcgtgggtgatgaaggccctaggggtgaaagcccttcggcgggggaaagataatgacggtaccgcagaagaagccccgggta
cttcgtgccagcagccgcgtaatacgaagggggctagcgttgcctggaatcactggggtaaagcgcacgtagggcggcttttaagtcag
gggtgaaatcctggagctcaactcagaactgccctgatactgagagggctgagtcggggagaggtgagtggaactgcgagtgta

>emp205

agagttgatcatggctcagattgaacgctggcggcatgcctaacacatgcaagtcgaacggcagcagcagcagcagcagcagcagcagc
gagagtgaggacgggtgagtaatgcataggaatctgccagtagcgggggacaaccggagaaatccgggctaataccgatacagcc
ctacgggggaaagcgggggaccgaaaggcctcgcaactattggatgagcctatgctggattagctagttgtagggtaacggcctaccaag
gagcagatccgtagctggtctgagaggacgatcagccacactgggactgagacacggcccggactcctacgggaggcagcagtgggg
aatattggacaatgggggcaacctgatccagcaatgccgctgtgtgaagaaggcctgcgggtgtaaaagcacttccgggggacaaa
aagcgtcgtccaataccagatgtttgatttaaccatagaagaagcaccggctaactctgtccagcagccgcgtaatacagagggtg
cgagcgttaatcggaaactactggggtaaagcgcacgtagggcggcgcgtaagttggatgtgaaatccccgggcttaactgggaattgca
ttcaagactggcgtgctggagtgtgggagagagaagtggaaattccaggagta

>emp207

agagttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgcacgagaaagtggagcaatccatgagtaaaagt
gagcaggggtgagtaaacagtgactaacctaccttcgagtggggaataactcgggaaaccgaggctaataccgataacactacgggt
caaaggagcaatccgttgaggagggggcgcggccgattagttgtagggtaaatggcctaccaagacgatgatcggatccggcc
tgagaggcgcacggacacactggaactgaaacacggcctcagactcctacgggaggcagcagtggggaattttgcgcaatgggggaaa
ccctgacgcagcaacgccgctggaggatgaagtccttcgggacgtaaaactccttcgatcgggacgattatgacggtaccggaagaaga
agccccggtaacttcgtccagcagccgcgtaatacagggggggccgagccgttctcgaattattggggtaaaggggtcgtagggc
ggttcgtaagtctcgtgtgaaatcctcaggctcaactggagctcgcagaaaaactgcaggcttgattgtggga

>emp208

agagttgatcctggctcagattgaacgctggcggcatgcctaacacatgcaagtcgaacggcagcagcagcagcagcagcagcagcagc
gagagtgaggacgggtgagtaatgcataggaatctgccagtagcgggggacaaccggagaaatccgggctaataccgatacagcc
ctacgggggaaagcgggggaccgaaaggcctcgcaactattggatgagcctatgctggattagctagttgtagggtaacggcctaccaag
gagcagatccgtagctggtctgagaggacgatcagccacactgggactgagacacggcccggactcctacgggaggcagcagtgagg

aatattggacaatgggggcaaccctgatccagcaatgccgctgtgtgaagaaggcctgcgggtgtaaaagcactttcgggggacaaa
aagcgtcgtccaataccacgatgtttgatttaacccatagaagaagcaccggctaactctgtccagcagccgcggttaatacagagggtg
cgagcgttaatcggaaactgggcgtaaaagcgcacgtaggcggcgcgctaagttggatgtgaaatccccgggcttaactgggaattgca
ttcaagactggcgtgctggagtgtgggagagagaagtggaaatccaggagtagcgggtgaaatgcgtagatacctggaggaacatcaatg
gcaaggcggcttcttgccccacactgacgctgagggggcggaaagcctggggatcaaacaggattagatacctgggaatcc
>emp401

agagtttgatcatggctcagagcgaacgctggcggcgtgcttaacacatgcaagtcgagtcgagtcgctgtagcaatacagcagggc
ggcgcacggcgcacgggtgctaacacgtgggtaacctgcctgcaatgggggataaccgccgaaagcgggctaataccgcatgac
attccgggtggctcggacgccggaatcaagcccgaaggcgtggagggggcccgccgattagctagttggtgaggtaacg
gctaccaaggctgtatcggtagccggctgagaggatggccggccacactggaactgagacacggtcagactcctacgggagga
gcaagtggggaatctgcgcaatggcgaaaagcctgacgcagcagccgcgtgggggatgaagcttctcgagtgtaaacctttcga
cccggacgaaacggcgcaaggcgtgacggtacgggtataagaagccccggctaactacgtccagcagccgcggtaaagcgtagg
gggccagcgttctcggaactgaggcgtaaaagggtcgtaggcgggtgagcggtaagtcggaagtgaaatctctggg
>emp406

agagtttgatcatggctcagataaacgctggcggcgtgcttaacacatgcaagtcgaacgccctggtccttctggagggagcagggga
gtggcggacgggtgaggaacacgtgggtgacctacccccgggtgaggaataccgccgagaaatcggcgacaagaccgcataggcttcc
ctttagcagggtagcaaaagctgctttcgaagaagagcggcggccaggggatgggctgccccgattagctggttggtggggaacgg
ctaccaaggcgcacgatcggttagctggtctgagaggatgagcagccacactgggattgagaacggcccagactcctacgggagggcagc
agtgaggaatttccacagtgggggcaaccctgatggagcaacgccgcgtgcaggaggacgctttcggagtgtaaactgctttccagg
gacgagcaaggacggtacctggggaacaagtcccggctaactacgtccagcagccgcggtaaatcgtaggggacgagcgttccgg
agttactggcgtaaaaggccccgagggcgtctgccgcgttctgctgacagctcctgcttaactgggagagcgggtcgaagacgggc
ggactgagaggaccagagaggacatggaactccgggtgtagtggtg
>emp407

agagtttgatcatggctcaggataaacgctggcggcgtgcttaacacatgcaagtcgacgctcctgtgtcttctggagaggcagggagt
gaggacgggtgagtaacacatgggtgacctacctccagtgaggaataccggcgagaaatcggcacaagaccgcataggctctgctg
ctcaggcgagcagagcaaaagtcgcaagacgcttgggggatggcctgtggccgattagctggttggtgaggtaaaggctcacaaggcga
cgatcggttagctggtctgagaggacgatcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaatctt
cgcaatggacgcaagtctgacggagcaacgccgcgtgcaggaggacgcttctcgagtgtaaaactgctttctgggggacgagcaagga
cggtaccccaggaagaagccccggctaactacgtccagcagccgcggtaaatcgtagggggcagcgttctccggagttactggcgt
aaagggcacgcagcggctcctgcgcgtgctgacagcgggtggcttaactcccaggggtgcgacagacggtgggact
>emp410

agagtttgatcatggctcaggacgaacgctggcggcgtgcttaacacatgcaagtcgagcgtgggtgcacgctctcgaagagagggtg
cactctagcggcgaacgggtgagtaacacgtggggaacctgccctcagacggggacaacctcgggaaaccggggctaataccggata
cgttcggcgggtcgcacgtgctggcggagaaagggtcgaagatcgtggaggaggggccccggcctatcagctcgttggtggggtgta
acggcctaccaaggcgcacgggtgagctggcctgagagggtggtcagccacactggaactgagatacgtccagactcctacgggag
gcagcagttgggaaatctgggcaatggcgcaagcctgaccagcagccgcgtgggggatgaaggcctcgggttgaaacctgt
caggaggacgaagccgcaaggtgacggtacctcagaggaagccccggctaactacgtccagcagccgcggtaaagcgtagggg
gcgagcgttctccggaatcactggcgtaaaagcgcgtgcaggcggcctggttaagtcctggtgaaagcctacggctcaactgtagaaggt
cggggagaccgggggctcag
>emp411

agagtttgatcatggctcagattgaacgccggcggcgtgcttaacacatgcaagtcgaacgggtgacaggtctcggatgctgacgagtgcc
gaacgggtgagtaatacatcggaacgtgccgatcgtgggggataacgaagcgaagccttgctaataccgcataagatctacggatgaa
agcagggaaaccgcaaggccttgcgcgaacggagcggccgatggcagattaggtggtgggataaaaagcttaccagccgacgatct
gtagctggtctgagaggacgaccagccacactggggctgagacacggcccagactcctacgggagggcagcagtggggaaatctggaca
atggcgaaaagcctgatccagccatccgcgtgcaggatgaaggccttccgggtgtaaaactgctttgtacggaacgaaaagactctggtta
atacctggggtccatgacggtaccgtaagaataagcaccggctaactacgtccagcagccgcggtaaatcgtaggggtgcaagcgtta

cggaattactgggcgtaaagcgtgcgcaggcgggtatataagacagatgtgaaatccccgggctcaacctgggaactgcatttggactgta
tagctagagtacggcagagggggatgggaattccgcgtgtagcagtga

>emp415

agagtttgatcctggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgtccttctcttcgggagaagggagtgccgg
acgggtgagtaaacacgtgggtcatctgccccaaagtgaggaataccggcgagaaatcggcgacaataccccatacgttccgtgaggaa
gcaaagctcgtaaagggcgtttgggagggagcctgcggcagattagttggtggggttatggcctaccaagacgacgatctgtagctgg
tctgagaggacgatcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaatttcggcaatgggggcaa
ccctgaccgagcaacgccgcgtgcaggacgaaggtcttcggatcgtaaactgctttctgcaggacgagtagggactgaggaag
aagcctcggctaactacgtgccagcagccgcggtaatacgtaggaggcgagcgttgcggagttactggcgtaaagcgtccgcaggc
ggtaaccaagttcggagtgacaggtgggtgctcaactgccacagccctcgaaaactggtgacttgagagcttcagagggaggcggaa
ttccgggtggagtggtgaaatgcgtagagat

>emp421

agagtttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcgctgagtgatcagcgagtgccggacgg
gtgagtaaacacataggtcacctaccccgaagtgaggaataccggcgagaaatcggcgacaataccgcataaggcttcgcgagaagcaaa
cgcaagacgctttgggatgggcctgtggccgattagctggttgggggtaatggcctaccaagacgacgatcggtagctggtctgagag
gatgatcagccacactgggattgagaacggcccagactcctacgggaggcagcagtgaggaatttcggcaatggcgagagcctgacc
gagcaacgccgcgtgcaggacgaagatcttcggattgtaaactgctttctcggggacgaatttatgacggtagccgaggaagaagcctc
gtaactacgtgccagcagccgcggtaagacgtaggagcaagtggttccggagttactggcgtaaagcggccgagggcggctgta
agttcagggtgacagtcacagcttaactggtgaagtgtctgaaaactgacggacttgaggaccagagagggacctggaatttcgggtg
agcagtgaaatgcgtagagatc

>emp422

agagtttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcgctgagcaatcagcgagtgccggacgg
gtgagtaaacacataggtcacctaccccgaagtgaggaataccggcgagaaatcggcgacaagaccgcataaggcttcgcgagaagcaaa
gtcgcaagacgctttgggatgggcctgtggccgattagctggttgggggtaatggcctaccaagacgacgatcggtagctggtctgaga
ggatgatcagccacactgggattgagaacggcccagactcctacgggaggcagcagtgaggaatttcggcaatgggcaaaagcctgac
cgagcaacgccgcgtgcaggacgaaggtcttcggatcgtaaactgctttctcggggaaagagaatggacggtagccgaggaagaagcct
cggctagctacgtgccagcagccgcggtaagacgtaggagcaagtggttccggagttactggcgtaaagcggccgagggcggct
gttaagttcagggtgacagtcacagcttaactggtgaagtgtctgaaaactgacggacttga

>emp425

agagtttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcctgttcttcggcagcagggagtgccg
gacgggtgagtaaacacatgggccacctaccccctggtgaggaataccggcgagaaatcgtcgacaagaccgatgcgttccgacaacg
aagcaaaagccggcaacggcgccagaggatgggcctgtggccgattagctagttggggaggtaaaagcccccaagcgacgatc
gtagctggtctgagaggatgatcagccacactgggattgagaacggcccagactcctacgggaggcagcagtgaggaatttcggcaatg
ggcgcaagcctgaccgagcaacgccgcgtgcaggacgaaggtcttcggatcgtaaaccgctttctcggggatgaatttatgacagtacc
gaggaagaagccccggctgactacgtgccagcagccgcggtaagacgtagggggcaagcgttgcggagttactggcgtaaagcgc
ccgagggcggtcggtcgcgttgaggtaacagctcctggcttacctgggagagcgtctcaagacgggctgacttgagggccagagagg
gacacggaat

>emp426

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtctccgaaggggatgagtaa
tggcgacgggtgagtaaacgtagctaactacctcagtggtggggataacctggggaaactcgggctaataccgcataacgtcctggc
ggcgcgagtcgctgggatcaaaagaccgcaagggctgctcgtcagggggctcggccgattagctagttggtgaggtaacggcccac
caagcgatgatcggtagccggcctgagaggggcgacggccactggcactgaaacacgggcccagactcctacgggaggcagcagtg
ggggaatttgcacaatgcccgaaggggtgatgcagcaacgccgcgtgaggatgaaggtcttcggatcgtaaacctcttcgaccgaga
aaaacggcggggtgaagatccgctctgatggtaccgggtggaagaagccccggccaactacgtgccagcagccgcggtaatacgtagg
ggcgagcgttctcgaattactggcgtaaagggcgtgtagggcgtgcgtaagtcggcggtgaaatccccgggcttaactcggggc
ggtcgttcgaaactgc

>emp427

agagtttgatcctggctcaggacgaacgctggcggcgtgcctcaaacatgcaagtcgagcgtgggtgcacgctctcgaagagaggggtg
cactctagcggcgaacgggtgagtaacacgtggggaaacctgccctccagacggggacaacctcgggaaaccgggggctaataccggata
cgttcgccgggtcgcacggcctggcagggaaaggggtcgaagatcgtggaggagggccccggcctatcagcttgttgggtgggtaa
cggcctaccaaggcgcacgggtgagctggcctgagagggtggtcagccacactggaactgagatacggccagactcctacgggagg
cagcagttgggaatcttgggcaatgggcgcaagcctgaccagcgcgcccgcgtgggggatgaaggcctcgggtttaaaccctgtc
aggagggacgaagccgcaaggtgacggtacctccagaggaagccccggtaactacgtgccagcagccggtgaagacgtaggggg
cgagcgtgtccggaatcactgggcgtaaagcgcgtgcagggcggcctgtaagtcccgtgtgaaagcctacggctcaactgtaggaggtc
gcgggagaccgccgggctcagggcagggagagggaagtaga

>emp429

agagtttgatcctggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcctgttgcctcggcagcagggagcggcg
ggcgggtgagtaacacatgggtcacctacccccctgggtgaggaataaccggcgagaaatcgtcgacaagaccgcatgcctccgacaagg
aagcaaaagcccggcaacgggcgccagaggatggcctgtggccgattagctagttggggaggtaaaagcccccaaggcgacgacg
gtagctggtctgagaggatgatcagccacactgggattgagaacggcccagactcctacgggagggcagcagtgaggaaatttcggcaatg
ggcgcaagcctgaccgagtaacgccgcgtgcaggacgaaggtcttcggatcgtaaactgctttctcggggatgaattatgacagtacc
gaggaagaagccccggctgactacgtgccagcagccgcgtaagacgtagggggcaagcgttgcgggagtactgggcgtaaagcgc
ccgcagggcgtcggcgcgtttgaggtaacagctcctggcttacctgggagagcgtctcaaagacgggctgacttgaggaccagagagg
gacacggaattccgggtggagtggtaaatgcgtagatatac

>emp434

agagtttgatcctggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcctgttgcctcggcagcagggagtgccg
gacgggtgagtaacacatgggtcacctacccccctgggtgaggaataaccggcgagaaatcgtcgacaagaccgcatagcttctcgggga
agcaaaagctccgcaagggggcggcagaggatggcctgtggccgattagctggtggggaggtaaaagcccccaaggcgacgacggt
agctggtctgagaggatgatcagccacactgggattgagaacggcccagactcctacgggagggcagcagtgaggaaatttcggcaatg
gcgcaagcctgaccgagcaacgccgcgtgcaggacgaaggtcttcggatcgtaaactgctttctcggggatgaattatgacagtaccg
aggaagaagccccggctgactacgtgccagcagccgcgtaagacgtagggggcaagcgttgcgggagtactgggcgtaaagcgc
cgcagggcgtcggcgcgtttgaggtaacagccccctggcttacctgggagagcgtctcaaagacgggctgacttgaggaccagagagg
acacggaattccgggtggagtggtaaatgcgta

>emp435

agagtttgatcctggctcaggacgaacgctggcggcgtgcctcaaacatgcaagtcgagcgtgggcgcgttgcctcgaagaggtgacg
cgctctagcggcggacgggtgcgtaaacacgtggggaaaccaacctccagacggggacaacctcgggaaaccgggggctaataccggat
acgttcgccgggcccgcacggcctggcagggaaaggggtcgaagatcgtggaggacggccccggcctatcagctagttgtgaggt
aacggcccaccaaggcgcacgggtgagctggcctgagagggtggtcagccacactggaactgagacacggctcagactcctacggga
ggcagcagttgggaatcttgggcaatgggcgcaagcctgaccagcgcgcccgcgtgggggatgaaggcctcgggtcgtaaaccct
gtcaggagggacgaagccgcaaggtgacggtacctccagaggaagccccggtaactacgtgccagcagccgcgtaagacgtaggg
ggcagcgttgcggaattactgggcgtaaagcgcgtgcagggcggcctgtaagtcccgtgtgaaagcctacggctcaactgtaggaag
ccgcgggaaactaccgggct

>emp424

agagtttgatcctggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtctccgcaaggggatgagtaaag
tggcgacgggtgagtaacacgtagctaacctacctcagagtgggggataacctggggaaactcgggctaataatgcataacgtcctggc
ggcgcgagtcgctgggatcaagaccgcaaggggtcgtcgtcaggggggctcggccgattagctagttggtgaggtaacggcccac
caaggcgatgatcggtagccggcctgagagggcgcacggccacactggcactgaaacacgggccagactcctacgggagggcagcag
ggggaatttgcacaatgcccgaaggggtgacgcagcaacgccgcgtgagggatgaaggtcttcggatcgtaaacctcttcgaccgaga
aaaacgggcgggtgaagatccgctctgatgtaccgggtggaagaagccccggccaactacgtgccagcagccgcgtaatacgtagg
gggcgagcgttgcggaattactgggcgtaaagggcgtgtagggcgtcggtaagtcgggcgtgaaatccccgggcttaactcggggc
ggtcgttcgaaactccgggctagagtgcaggagagggaagtgga

>emp438

agagtttgatcctggctcagaacgaacgcccgccatgcctaatacatgcaagtcgaacgcctgtagcaatacaggagtggcggacgg
gtgagtaacgcgtgggaacctgcccttgagtgcgggataaccgagggaaacttcggctaataccgcatgttctcagcaatgagggaaag
atztatcgctcagggagggggcccgcgtcggattagctagttggtggggtaacggcctaccaaggcgcgatccgtagctggtctgagagga
tgatcagccacactgggactgagacacggcccagactcctacgggagggcagcagtagggaatattggacaatgggggcaacctgatcc
agcaatgccgcgtgagtgtgaaggccttaggggtgtaaaactcttcgcgcgcgcgatgatgacggtagcgcgagaagaagccccgg
ctaactcgtgccagcagccgcggttaatacgaagggggcgcagcgtgttcggaactactgggcgtaaaaggcgcgcagggcgcaaca
agtcagggcgtgaaagccccgggctaacctgggaattgcgctgggactgttaggcttgagtacgggagaggagagtggaaatccagtg
agaggtgaaattcgtagatattggggaagaacaccgggtggcga

>emp439

agagttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcgctgagcaatcagcgagtggcggacgg
gtgagtaacacataggtaacctaccccgaagtgaggaataccggcgagaaatcgccgacaagaccgcataggcttcgcgagaagcaaa
gtcgaagacgctttgggatgggctgtggccgattagctggttgggtgggtaaatggcctaccaagacgacgatcggtagctggtctgaga
ggatgatcagccacactgggattgagaacggcccagactcctacgggagggcagcagtgaggaatttcggcaatgggcaaaagcctgac
cgagcaacgccgcgtgcaggacgaaggtcttcggatcgtaaactgctttctcggggaagagaacggacggtaccggaggaagaagcct
cggctaactacgtgccagcagccgcggtaaagcgtagaggcaagtgttccggagtactgggcgtaaaagcggcgagggcgtctg
ttaagttcagggtgacagtcacgcttaactggtgaagtgtctgaaaactgacggacttgaggaccagagagggacctggaatccgggt
ggagcagtgaaatcgtagag

>emp441

agagttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcgctgagcaatcagcgagtggcggacgg
gtgagtaacacataggtaacctaccccgaagtgaggaataccggcgagaaatcgccgacaagaccgcataggcttcgcgagaagcaaa
gtcgaagacgctttgggatgggctgtggccgattagctggttgggtgggtaaatggcctaccaagacgacgatcggtagctggtctgaga
ggatgatcagccacactgggattgagaacggcccagactcctacgggagggcagcagtgaggaatttcggcaatgggcaaaagcctgac
cgagcaacgccgcgtgcaggacgaaggtcttcggatcgtaaactgctttctcggggaagagaacggacggtaccggaggaagaagcct
cggctaactacgtgccagcagccgcggtaaagcgtagaggcaagtgttccggagtactgggcgtaaaagcggcgagggcgtctg
ttaagttcagggtgacagtcacgcttaactggtgaagtgtctgaaaactgacggacttgaggaccagagagggac

>emp442

agagttgatcctggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgcccttcacgcgagtggagggagtggcgg
acgggtgaggaattcatgggtatctacctcgggtgggggataaccggcgagaaatcgccgacaatcccgcatagggtcacggagacgtg
agcaaaagcgcagagggcgggggatgagcctgtggccgattagctggttgggtgtaagggctaccagggcgacgatcggtagct
ggtctgagaggatgatcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaatctccgctatggacgc
aagtctgacggagcaacgccgcgtgcaggaggacgcttttcggagtgtaaactgctttctcgggacgagatggacggtaccctggaa
gaagcctcggctaactacgtgccagctgccgcggttaatacgtaggaggcgcgctgtccggagtactgggcgtaaaagggcgcgcagg
cggcccggcgttgggtggtgaaagttgcgagcttaact

>emp444

agagttgatcatggctcaggataaacgctggcggcgtgcctaacacatgcaagtcgaacgccctggtgcttcggcagcagggagtggcg
gacgggtgagtaacacatgggtcacctacccccgtgtaggaataccggcgagaaatcgtcgacaagaccgcataggcttctcgggga
agcaaaagcgcgaagggggcggcaggggatgggctgtggccgattagctggttggggaggtaaaagcccccaaggcagcagcgg
tagctggtctgagaggatgatcagccacactgggattgagaacggcccagactcctacgggagggcagcagtgaggaatcttcggcaatgg
gcaagcctgaccgagcaacgccgcgtgcaggacgaaggtcttcggatcgtaaactgctttctcgggatgaatttatgacagcaccg
aggaagaagccccggctgactacgtgccagcagccgcggtaaagcgtagggggcaagcgttccggagtactgggcgtaaaagcggc
cgagggcgtcggcgcgtttgaggtaacagctcctggcttacctgggagagcgtctcaaaagcggcgtgacttgaggaccagagaggg
acacgaatccg

>emp210

agagttgatcctggctcagattgaacgctggcggcgtgcctaacacatgcaagtcgaacggcagcagcagtgtagcaatacacggtgatg
gcgagtgggcgacgggtgagtaatgcataggaatctgccagtagcgggggacaaccggagaaatccgggctaataccgcataccg
ctacgggggaaagcgggggaccgaaagcctcgcactattgtagcctatgtcggattagctagttgtagggtaacggcctaccaag

gcgacgatccgtagctggctgagaggacgatcagccacactgggactgagacacggcccggactcctacgggaggcagcagtgggg
aatattggacaatgggggcaaccctgatccagcaatgccgcgtgtgaagaaggcctgcgggtgtaaacacttccggtggggacaaa
aagcgtcgtccaataccacgatgtttgatttaaccatagaagaagcaccggctaactctgtccagcagccgcgtaatacagagggtg
cgagcgttaatcggaaactgggcgtaaacgcacgtagcggcgcgtaagttggatgtgaaatccccgggcttaactgggaattgca
ttcaagactggcgtgctggagtgtgggagagagaagtggaaattccaggagtagcgggtgaaatgcgtagatacctggggaacatcagtg
cgaaggcggcttctggccaacactgacgctgaagtgcggaagcgtgnggatccaacaggattagatacctggtagtcc
>emp7

agagtttgatcctggctcaggataaacgctggcggcgtgcctaacacatgatagtcgaacgctttatccttttctcgaagagcaaaaagg
attgagagtggcggacgggtgatgaacacgcgggttatctgcctctggttgaggaataccggcgagtaategccgactataccgcatagcc
ttgtaacaaggaaagcccggcaacgggtgccgagagatgagcctgcggccgattatgctcgttgggtgggtaaaaggcctaccaaggcc
acgatcggtagctggtctgagaggacgatcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaat
cggcaatgggggcaaccctgaccgagcaacgccgcgtgcaggacgaaggtcttcggatcgtaaactgcttttctgggggacgactccc
acggtaccccagcaatcagccccggctaactacgtgccagcagccgcgtaatacgtagggggcaagcgttgcggatataattggcggt
aaagcgcacgcaggcggctgccaaagtcggagtgacagttccggcttaactgggaaaggtctcgaaaactggtggacttgagactttg
agggggagacggaattccgggtggagcgggtgaaatgcgtagatccggaggaacaccgaaagcgaaggcagctcctggcaaggtc
tgacgctcatgtgcgaaagctatgggagcgaacaggattaaatacctggtagtcc
>emp413

agagtttgatcatggctcaggataaacgctggcggcgtgcctcacacatgcaagtcgaacgcccttccgaagagaggggagtggcggacg
ggtgaggaacacgtgggtacctacctaagatgaggaatactccgagaatcggagacaagaccgatacctctcctcggagaggaa
aaccgcaaggtgcttgagatgggctgcgcctgattagctggttgggtgggtaaaagcctaccaaggccacgatcagtagctggtctgag
aggatgatcagccacactgggattgagaacggcccagactcctacggggggcagcagtgaggaatttctcaatggaggcaactctgaa
cgagcaacgccgcgtgcaggatgacaccttcggggcgtaaacgttttctcggagaagagagtgagcggatctgaggaacaagccc
cggtaactacgtgccagcagccgcggttaatacgtagggggcaagcgttgcggagttactggcgtaaaaggcccgcagggcgggtgg
ctcaagtttgactgacaatcccgcttaacgcgggaggggtgctgcaaaactgggttacttgagcatgagagaggggacaggggaattccgg
gtggagtggggaatgcgtagatatacggaggaaccccgaaggcgaaggcactgtcctggctcattgctgacgctcagggcgaaaagcc
>emp55

agagtttgatcatggctcagaatcaacgctggcggcgtgcctaacacatgcaagtcgaacgagaaagtggagcaatccatgagtaaagt
gcgaccgtggtgagtaacacgtgactaacctacctccagtggggaataactccgggaaaccggggtaataaccgataaacatcgaag
atcaaagcagcaatgcgcttggagagggggtcgcggctgattagctcgttggcgggtaacggcccaccaaggcgaagatcggtatccg
gcctgagagggcgcacggacacactggaactgaaacacggctccagactcctacgggaggcagcagtgggggaatttgcgcaatggggg
aaacctgacgcagcaacgccgcgtggaggatgaagccctttggggcgtaaacctcttcgactcggacgataatgacggtacgagtgga
agaagcaccggtaactctgtgccagcagccgcggttaatacagagggtgcgagcgttgcggaattattgggcgtaaaaggcgcgtagg
cggcgcggttaagtcacctgtgaaatctccggcttaactgggagctcgcagggcaactaccgtgctggagtgtgggagaggtgcgtgga
attccccggtgtagcgggtgaaatgcgtagatatacggaggaacacttggcgaaaagcggcgcactggaccataactgacgctgagggc
gaaagctaggggagcaaacaggattagatacctggtagtcc

Clone sequences recovered from the 3' end-portion-data-set:

>ebs81

aaatgcgtagatatcaggaggaacaccgggtgtagacggctttctggaccgtaactgacgctgaaacacgaaagcgtgggtagcaaac
aggattagataccctgtagtccacgccctaaacgatgcatattggtgtgggtcattcatttgatcccgccggagctaaccggttaaatatg
ccgcctggggagtagcagtcgcaaggctgaaactcaaagggaattgacggggggcccgcacaagcgggtggagcatgtggtcaattcgacgc
agcgcgaagaaccttacctaggtcgaacggcttggacaattctggaaacagggactcccgaaggaccggagtcgaggtgctgc
atggctgtcgtcagctcgtgctgtagatgtgggttaagtcccgaacgagcgaaccctgtcctgtgtgccatacctttaaagtggaa
tcgaggagaccgccaggataacctggaggaaggtggggatgacgtcaagtcctcatggcctttatgcttagggctacacacgtgtac
aatgggcggtacaatgggtgccaaccgcgagggggagccaatcccaaaaaccggtctcagttcgattgcaggctgcaactcgctg
catgaagctggaatgctagtaatgccgtatcagaatgatggcgtgaatacgttcccgggcctgtacacaccgccctg

>ebs78

ccacaactgacgctgatgcgcgaaagctaggggagcaaacaggattagataccctggtagtcctagccctaaacgatgatcgttggtgtg
acaggtagccaatcctgtcgtgccggagctaaccgcttaagcgatccgcctggggagtagcgtcgcaaggctgaaactcaaaggaattga
cggggggcccgcacaagcgggtggagcatgtggtcaattcgacgcaacgcgaagaaccttacctgggctcgaatgtagtgaccgggggt
ggaaacatccctcccgaagggtgctatataggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcggcaacga
gcgcaacccttattgccagttgccatcatttagtgggactctggcgaaactccccggataacggggaggaaggtggggatgacgtcaa
gtcctcatggcctttatgtccagggtacacacgtgtacaatgggtggtacaaaccgtcgaaaaccgcgagggggagctaatacggaaa
aagccatcctcagttcgattggagctgcaactcgactccatgaagctggaatcgctagtaatcgcggtacgcggtgaacac
gttcccgggcctgtacacaccgccctg

>ebs83

gcgaaggcagctcgtacactattttagcgtgagggcacgaaagcgtggggatcaaacaggattagataccctggtagtcacgccctaa
acgatgattactcgcacatcagcgtgactgttgggtgctgagcgaagcattaagtaatccacctgggaagtacgaccgcaagattaaac
tcaaagaaattgacgggggctgcacaagcgggtggaacatgtggttaattcgaggcaacgcgaagaacctaccaggtttgacatgcc
tctgtgggtaccagaatggaaccttcacttcgggtgggagagcacaggtgctgcatggctgtcgtcagctcgtcgtgagatgttggg
ttaagtcggcaacgagcgaacctcgtccctgttctaaccgggtaatgccgagaactctgggaaactccgggtgataagccggagg
aaggtggggatgacgtcaagtcctcatggccttacagcctgggctacacacgtgttacaatggcagtgacaatgctcgtcactccgaa
ggagaggctaatacgaaaaagctgcctcagttcgattgtactctgcaactcgagtacatgaagctggaatcgctagtaatcgcgcatc
gcggcgcggtgaatacgttcccgggtcgtgtacacaccgccctg

>ebs84

aatgcgtagatatcgggaggaacaccagtggcgaaggcgtatcctggggcttacctgacactgaggagcgaagctaggggagcaaaa
cgggattagataccccggtagtcctagccgtaaacgatggactaggtgtaggtggtatcactccatctgtgccgagttaacacaataa
gtgtcccgcctggggagtagcgtcgcaagactaaaactcaaagggaattgacggggggcccgcacaagcgggtgatcatgtggttaattcg
aagcaacgcgaagaaccttacctggattgacatcctctgaagttgtagagatatcaacggctcgaagagacagagagacaggtgctgc
atggctgtcgtcagctcgtgttgaatgttgggttaagtcccgaacgagcgaaccctaccttgggtgccatcattaagttgggcactc
cagagggactgccgtcgtcaagacggagggggtggggatgacgtcaagtcctcatggcctttatccaggggcgacacacgtgatacaa
tgtagatacaaagggcagcgaagtggtgacacggagccaatctcaaaaaatctatctcagttcgattggagctgcaactcgactccatg
aagttggaatcgctagtaatcgcggtacgacacgccggtgaatacgttcccgggcctgtacacaccgccctg

>ebs71

agcggtgaaatgcgtagatatctgcaggaacaccggtggcgaaggcggctggctggaccgcaactgacgctgaggaacgaaagctagg
ggagcaaacaggattagataccctggtagtcctagccctaaacgatcaggacttgggtgcatcccgtcgggggtgctcccgtagctaaccg
cgtaagtctcgcctggggagtagcgtcgaaggctgaaactcaaagggaattgacggggggcccgcacaagcgggtggaacatgtggtt
caattcgacgaacgcgaagaaccttacctggctcgaagtgccggcgaccagcgggtggaacatcgcttctcgaagagacgccggt
agaggtgctgcatggctgtcgtcagctcgtccgtgaggtgttgggttaagtcccgaacgagcgaaccctacacttagttgccaccgc
caagggggaactctaagaggactgctccggataacggagaggaaggtggggatgacgtcaagtcagcacggcctttatgtccagggcta

cacacgtgttacgatgcgcgttacagaccgttgcgaactcgcgagggggagctaatacggtaaaaccgtgctcagttcggattgcagtctgc
aactcgactgcatgaagctggaatcgctagtaatggcgtatcagaacgacgccgtgaatacgttccccggcctgtacacaccgccgt
>ebs72

gtggcgaaggcggctcactggcccgatactgacgctgaggcacgaaagcgtggggagcaaacaggattagataccctggtagtccacg
ccgtaaacgatgaatgccagccgttagtgggttactactagtggcgcagctaacgctttaagcattccgcctggggagtagctgcgaag
attaaaactcaaaggaattgacggggggcccgcacaagcgggtggagcatgtggttaattcgacgcaacgcgcagaacctaccagcccttg
acatcccggctcgggactccagagacggagttcttcagttcggctggaccggagacaggtgctgatggctgtcgtcagctcgtgctga
gatgttgggttaagtcccgaacgagcgcagccccgtccttagttgctaccattcagttgagcactctaaggagactgccggtgataagcc
gcgaggaaggtggggatgacgtcaagtcctcatggcccttacgggctgggctacacacgtgctacaatggcggtgacaatgggatgctaa
ggggcgacccttcgaaatctcaaaaagcgtctcagttcggattgggctcgtcaactcgagccatgaagttggaatcgctagtaatcgtg
gatcagcagccacgggtgaatacgttccccggcctgtacacaccgccgt

>ebs73

tccccgtgtagaggtgaaatcgtagatatacgggaggaacaccagtggcgaaggcggcgacctggacagacattgacgctgagacgcga
aagcgtggggagcaaacaggattagataccctggtagtccacgctgtaaacgatgggtgctagctgtggtgggtgtgacctgccgtgg
cgagctaacgcaataagcaccgccctgggaagtacggccgcaaggctaaaactcaaaggaattgacggggggcccgcgcaagcgggt
ggagcatgtggttaattcgacgcaacgcgcagaaccttacctggtcttgacatctcggaaatcccagagatggggaaagtctctcggg
gagccgagagacaggtgctgcatggctgtcgtcagctcgtgctgagatgttgggttaagtcccgaacgagcgcgacctcgcattag
ttaccatcattcagttgggactctagttgggactgccggtgttaaacgggtggaaggtggggatgacgtcaagtcctcatggcctttatgac
agggctacacacgtgctacaatggccggtacaaaagcgtgccaactcgcgagaggagctaatacgcataaaaccggtctcagttcggatt
ggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgagatcagcacgcttcgggtgaatacgttccccggcctgtacacacc
gccgt

>ebs77

ggtgtagcgggtgaaatcgtagatatacgcaggaacaccgtggcgaagcggcgcgctggaccactactgacgctgaggaacgaaag
ctaggggagcaaacgggattagataccctggtagtcttagccctaaacgatcaggacttgggggtgccaccggtcgggtgtcgtcccgg
gtaaacgcgttaagtctgcacctggggagtacggctcgaagactgaaactcaaaggaattgacggggggcccgcacaagcgggtggaaca
tgtgttcaattcgacgctacgcgaggaaccttacctgggctcgaagcgcacttgacaagctatagaatatggccttcccgaagggaca
ggtgtggaggcgtgcatggctgtcgtcagctcgtgccgtgaggtgttgggttaagtcccgaacgagcgcgcaacctcgcagtagttgcc
actcgaagagggaactctacgtggactgtccggataacggagaggaaggtggggatgacgtcaagtcgcatggcctttatgtccagg
gctacacacgtgttacaatgcagggcacaaccgttccaaccgcgagggggagctaatacggaaaaagccaatcccagtcggattgg
agtctgcaactcgactccatgaagtcggaatcgctagtaatcgagatcagcacgcttcgggtgaatacgttccccggcctgtacacaccg
ccgt

>ebs86

agcgggtgaaatcgtagatatacaggaggaacacctgcgggtgtagacggcttctggaccatcactgacgctgagacacgaaagcgtggg
gagcaaacaggattagataccctggtagtccacgcctaaacgatgcatacttgggtgagccgttcattcgttccggtgagtaacgc
gttaagtatccgcctggggagtacggctcgaagcgtgaaactcaaaggaattgacggggggcccgcacaagcgggtgagcatgtggtta
attcgacgcaacgcgagaaccttacctgggctcgaacggcttcggcaactcgtggaacatgggtcctcccgaagggaccggagtc
gaggtgctgcatggctgtcgtcagctcgtgctgagatgttgggttaagtcccgaacgggcgcaaccttactcgtattgccaccgcga
agggggaactctacgaggactgtccggataacggagaggaaggtggggatgacgtcaagtcgcatggcctttatgtccagggtaca
cacgtgttacaatgcagagtacaaccgtcgaaccgcgagggggagctaatacggaaaaactttgctcagttcggattgcagctcgtca
actcgactgcatgaagccggaatcgctagtaatggcgtatcagatcagccgtgaatacgttccccggcctgtacacaccgccgt

>ebs93

aatgcgtagatatacgggaggaacacctgtggcgaagcggcgcactggaccataactgacgctgaggcgcgaaagctaggggagcaaa
caggattagataccctggtagtcttagccctaaacgatggatgcttgggtgtagcgggtacccaatccgattgtccgaagctaacgcgataa
gcatcccgcctggggagtagctcgggaaggctaaaactcaaaggaattgacggggggcccgcacaagcggcggagcatgtggcttaattc
gaggcaacgcgagaaccttacctgggtttgacatgtagggaaaagccgtagagatacgggtgtccttcggggccttacacaggtggtgat
ggctgtcgtcagctcgtgctgagatgttgggttaagtcccgaacgagcgcgcaaccttacctatgttccagcgcgtaattggcggggac

tcgtaggagactgccggggtcaactcggaggaaggtggggtcgaagtcaagtcacatgcccttacgtccagggtgtacacatgctac
aatggccggtacaacggggtgctattccgcgaggagagccaatccctaaagccgggtcagttcggattggagctgcaactcgactcca
tgaagtcggagtcgctaatacccggatcagcaacgccgggggtgaatacgttcccgggcctgtacacaccgccgt

>ebs95

gaacacctgtggcgaagcgggtcactggaccattactgacgctgatgcacgaaagctaggggagcaaacaggattagataccctggtag
tcctagccctaaactatgatcgcttgggtgacaggtacccaatcctgtcgtccggagctaacgcgtaagcgatccgcctggggagtacg
gtcgcaaggctgaaactcaaggaattgacggggcccgcacaagcgggtggagcatgtggtcaattcgacgcaacgcgaagaacctta
cctgggctcgaagcgtagtggaccgaggtgaaatattcttcccgaaggactgctacgtaggtgctgcatggctgtcgtcagctcgtg
tcgtgagatgttgggttaagtcgcccaacgagcgaaccttattgtcagttgccatcattagttgggactctggcgaaccgcctcgata
acggggagggaaggtggggatgacgtcaagtcctatggccttattgtccagggccacacagtgctacaatggttggtacaaccgtcgc
aaaccgcgagggggagctaactcgaaaaagccaatcccagtcggattggagctgcaactcgactccatgaagctggaatcgctagta
atcgagatcagcagctcggtgaatacgttcccgggcctgtacacaccgccgt

>ebs97

agcgggtgaaatgcgtagatatcaggaggaacacctcgggtgtagacggcttctggaccatagctgacgctgagacagaaagcgtgggt
agcaaacaggattagataccctgtagtccacgccctaaacgatgcatacttgggtcggccattcagttggtcgggtgccggagctaacgcg
ttaagtatccgcctggggagtacggtcgcaaggctgaaactcaaggaattgacggggcccgcacaagcgggtggagcatgtggttaa
ttcgacgcaacgcgaagaaccttacctgggctcgaacggcttctgaatactttgaaacaggggtagtcgccgcaaggacaggagtcga
gggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcgcccaacgagcgaaccttgcctgtgttgaaccccga
gggggactctcaggagaccgccagcgataagttggaggaaaggtggggatgacgtcaagtcacatggccttattgtccagggctacaca
cgtgctacaatggacggtacaagcgcagcgaactcgcgagaggagcaaatcgaaaaaccgttctcagttcggattgcaggctgca
actcgctgcatgaagctggaatcgctagtaatggcatatcagaacgatgccgtgaatacgttcccggccgtgtacacaccgccgt

>ebs99

ctgagaaagctaggggagcaaacaggattagataccctggtagtccagcgtaaacgggtgacgcttctgtaaaaggaatcgaccctt
tgtggcgtagccaacgcgtaaacgtccgcctgggaagtacggtcgcaagattaaaactcaagaaattgacggggcctgcacaagc
ggtggagtatgtggctcaattcgtgcaacgcgaagaaccttacctggccttgacatgcaagtagtagaagggtgaaagcctaacgaggta
gcaataccagcttgacaggtgctgcatggctgtcagctcagctcgtgagatgttgggttaagtcgcccaacgagcgaacccctgtg
cctgttccaccggctcgcgagatcggagcactctggacagactgcctcgttaaacgaggaggaaggtggggatgacgtcaagtcagta
tgcccttataaggtgggtacacacgtgctacaatggcgggtgacaatgggtgctacacggtaacgtgatgtaactctctaaaagccgtc
agttcggattgggctgcaactcgacccatgaagtcggaatcactagtaatcgtggatcagcatgccacgggtgaatacgttcccgggtc
gtacacaccgccgt

>ebs105

ccataactgacgctgaggcacgaaagcgtgggtagcaaacaggattagataccctggtagtccacgccctaaacgatgcacatttgggtg
ggtcattcagttgatccgtgccggagtaacgcgtaaatgtccgcctggggagtacagtcgcaaggctgaaactcaaggaattgacgg
gggcccgcacaagcgggtggagcatgtggttaattcgacgcaacgcgaagaaccttacctgggctcgaatgctgtggaccagccggtga
aagccggcctcccgaaggactgcagtataggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcgcccaacga
gcgcaacccttgcctgtgttgcatactgaaaagtggaactcgcaggagaccgccagggataacctggaggaaggtggggatgacgtca
agtctcatggccttattgtccagggctacacacgtgctacaatggacggtacaaagggtgccaacccgcgagggggagccaatcccaa
aaaaccgttctcagttcggattgcaggctgcaactcgcctgcatgaagctggaatcgctagtaatggcagatcagaacgctgccgtgaatac
gttcccgggccttgtacacaccgccgt

>ebs111

aggaagaccagtggcgaagcgggtctctggacagttcctgacgctgaggcacgaaggtcaggggagcaaacgggattagatacccg
gtagtcctgacagtaaacgggtgacgcttgggtgaggaggtcggcccttctgcgccggagctaacgcgtaaacgtgccgcctggggg
gtacggctcgaagattaaaactcaagaaattgacggggcccgcacaagcgggtggagtagtggctcaattcgatgcaacgcgaagaac
cttaccaggccttgacatgattgtaagttgtaagccagcactggagcaatccagaccctgcacaggtgctgcatggctgtcgtcag
ctcgtgctgtagatgttgggttaagtcgcccaacgagcgaacccctgtgaactgttgcaccgagccgtgaggctcagcactctgtca
gactccccgctcaacggggaggaaggtggggacgacgtcaagtcagtagtggccttacggcctgggctgcacacgtactacaatgcc

cagtacagaacgatccaagaccgaggtggaggaatcactaaaactgggccagttcagattgaagtctgaacccgacttcatgaag
ccggaatcgctagtaatggcgcatcagctacggcgccgtgaatacgtcccgggcctgtacacaccgcccg
>ebs114

ggcgaaaggcagcctcttggtccatatctgacgctgaggcacgaaagcgtggggagcaaacaggattagataccctggtagtccacgctgt
aaacgatgatcactagatgttagttctgtttaacagaattggtatcgagctaacgcgttaagtatccgcctggggactacgatcgcaagatt
aaaactcaaagaattgacgggggcccgcacaagcgggtggtatcatgtggttaattcgactacgcgaggaaccttacctgggctagaca
tgtacttgacagctctgaaacagagcttctaagcgaagccttagacaggtacacaggtgctgcatggctgtcgtcagctcgtcgtgaga
tgtgggttaagtcttaacgagcgaacacctgtcgtcagttgcaactgcttatgcaaggcactctgacgagactgctgggaaaccagg
aggaaggtggggatgacgtcaagtcgcatggcccttatgtccagggtacacacatgatacaatggctagtaaaagagctcgaactc
gcaaggtgagctaatcacaactagctaaagttcagattgaggtctgcaactcgaactcagctcatgaagttgaaatcgctagtaatcgcgcat
cagcaacgacgcggtgaatacgttctcgggcctgtacacaccgcccg

>ebs115

tcccggtgtagcgggtgaaatgcgtagatcgggaggaacaccagtgccgaaggcgatcatctggggcttacctgacactgaggagcga
aagctaggggagcaaacgggattagataccccggtagtctagccgtaaacgatggacactaggtgtaggtggtatcgactccatctgtgc
cgagttaacacaataagtgtcccgcctggggagtagcgtcgaagactaaaactcaaaggaattgacggggcccgcacaagcgggtg
atcatgtggttaattcgaagcaacgcgaagaaccttacctggattgacatcctctgaagttgtagagatatcaacggctcgaagagaca
gagagacaggtgctgcatggctgtcgtcagctcgtgttgaaatgtgggttaagtcggcaacgagcgaacacctaccttggttgccat
cattaagttgggcaactccagagggactgccgtcgtcaagacgggagggtggggatgacgtcaagtcctcatggcctttatatccagggc
gacacacgtgatacaatgtagatacaaaaggcagcgaagtggtgacacggagccaatcaaaaaatctatctagttcggattggagtc
tgaactcgaactcagctgaagttggaatcgctagtaatcgcgatcagcagccgcggtgaatacgtcccgggcctgtacacaccgcccg
t

>ebs117

cgctacactattattgacgctgaggcacgaaagcgtggggatcaaacaggattagataccctggtagtccacgccctaaacgatgattactc
gacatcagcgatacactgttggtgtctgagcgaagcattaagtaaccctgggaagtacgaccgcaaggtgaaactcaaaggaattg
acgggggtccgcacaagcgggtggagcatgtggttaattcgtatgcgacgcgaagaaccttacctgggctcgaacggctatggaccgggtg
tagaaatcgcctttcccgaaggactgtagtcgaggtgctgcatggctgtcgtcagctcgtcgtcgtgagatgtgggttaagtcggca
cgagcgaacacctgtcctgtgtgcaaacggaaaggtgcaactcaggagaccgccagcgataagttggaggaaggtggggatgacgtc
aagtcacatgacctttatgtccagggtacacacgtgctacaatggtcgtacaaagcgtgcaaacccgcgagggggagctaactcga
aaaaccgatctcagttcggattgcaggctgcaactcgcctgcatgaagctggaatcgctagtaatggcgtatcagaacgacgccgtgaata
cgtcccgggcctgtacacaccgcccg

>ebs119

atcgtagatatccggaggaatacaaaaggcgaaggcagtgctctggctggttactgacgctcaggcgcgaaagctaggggagcgaaca
ggattagataccctggtagtctagccgtaaacgatgaccactagatgtacgaggtatcgaccctcgtgtcgtgtagttaacacgctaagt
gtccgcctggggagtagcgtcgaagattaaaactcaaaggaattgacggggaccgcacaagcagcggagcgtgtggttaattcgtatg
cgacgcgaagaaccttaccagggtttgacatgggcttgatccgccagagatggcgtcccgaagggcaagtcacagatgctgcatg
gctgtcgtcagctcgtgccgtgaggtgtgggttaagtcggcaacgagcgaacccctagggtcggttacaggttccggcccactgcc
gcgatcaacgcggaggaaggtggggatgacgtcaagtcagcagcggctcttacctgggcgacacgacgctacaatggctgaaaca
aagagttgcaaacccgcgagggggagctaactcgtcaaaagccagctcagttcggattgcaggctgaaaccagcctgcatgaaggtgga
gttgctagtaaccgtgggtcagcaccacgggtgaatacgtcccgggtctgtacacaccgcccg

>ebs121

gaataccggaggcgaaggcgggtagctgggcagacactgacactatcaccgaaagcgtggggatcaaacaggattagataccctggta
gtccacgccgtaaacgatggatactgtgttgagggtattgaccttcaagtacgaagtaaacgcgttaagatcccgcctggggagtagc
gtcgaagattaaaactcaaaggaattgacggggaccgcacaagcagcggagcgtgtggttaattcgtatgcaacgcgaagaaccttac
caggggttgacatgcaactgaccgccgtaatggggcgtcccttcggggcgggtgcacaggtgctgcatggctgtcgtcagctcgtgctg
gagatgttgggtcagttcccgaacgagcgaacccctgccgtagttgaatgatctagcgggaccgccgcatcaacgcggaggaagg
cggggatgacgtcaagtcagcagcccttacgtcctgggcaacacacgctacaatggctggtacaacggggcccaaccgcgag

ggggagccaatcccatcaaaaccggcctcagttcagattgcaggctgcaacccgcctgcatgaaggaggagttgctagtaaccgtgggtc
agcacactacgggtaatacgttcccgggtctgtacacaccgccgt

>ebs123

gcgggtgcctggaagattcctgactctgaaacacgaaagccaggggagcaaacgggattagataccccggtagtctggccgtaaacggg
gtgcattggcgtagggcggattcagccccgtctgtgctgcagataactcgttaaatgcaccgcctgggaaatacggcgcgaagattaaaact
caaagaaattgacggggggccgcacaagcgggtggagtatgtggcttaattcgatgcaacgcgaagaaccttacctggccttgacatgcatc
tagtagaacttgaaacagggacgaggtagcaataaccaggatgcacaggtgctgcatggctgctcagctcgtgctgagatgtgggtt
aagccccgcaacgagcgaacccctatgattagttgccagcaatggaccctctagtcagactgctctgtgaaacggagaggaaggtggg
gatgacgtcaagtcaggatggcccttacggccagggtgcacacgtactacaatgccagtacagaatgaaccgaaaccgcaaggtgga
ggcaatcctcaaaaactggggccagttcggattgggggctgcaactgccccgtgaagttggaatcgtagtaatggcgcacatcagctacgg
cgccgtgaatacgttcccgggcctgtacacaccgccgt

>ebs126

ccgtaactgacgctgagacacgaaagcgtgggtagcaaacaggattagataccctggtagtcacgcacctaaacgatgcatatttgggtg
ggcattcagttgatccgtgccggagtaacgcgttaaatatgccgctggggagtagcagtcgcaaggctgaaactcaaaggaattgacgg
gggcccgcacaagcgggtggagcatgtggtcaattcagcgaacgcgaagaaccttacctgggctcgaacggctggtaccggcggtgg
aaacatcgtacccccgcaaggggggtccagtcgaggtgctgcatggctgctcagctcgtgctgagatgtgggttaagtcccgaacg
agcgcacacctgtcctgtgttccatcattcagttgggaactcgcaggagaccgccagcgataagctggaggaaggtggggatgacgtc
aagtcctcatggcctttatgtccagggctacacacgtgtacaatggacggtacaaagcgcagcaaacccgcgagggggagccaatcgc
aaaaaacggtctcagttcggattgcaggctgcaactgcctgcatgaagctggaatcgtagtaatggcagatcagcacgctgccgtgaat
acgttcccgggcctgtacacaccgccgt

>ebs171

aatgcgtagatattggaagaacaccgggtggcgtaaagcgtcgtgggacactgacactgaggaacgaaagctaggggagcgaatg
ggattagataccccagtagtcttagccgtaaacgatggatactaggcgtagtggtatcaccgccgtgtgccgagtaacgcgataag
tatcccgcctgagtagtacggccgcaaggtgaaactcaaaggaattgacggggggccgcacaagcgggtggaacatgtggtcaattcga
agcaacgcgaagaacctaccaaggcttgacatctggagaactttcggaaacgagaaggtcccgcaagggagctccaagacaggtgg
tgcattggctgctcagctcgtgctgagatgtgggttaagtcccgaacgagcgaacccccgttgtagttgccatcaggtaatgctgg
gaactctagtgaactgccgtcgaagacgcgaggaaggagggatgatgtcaagtcacatggcctttatcccagggtacacacgtgc
tacaatggggagggacaaagggcagcaacacagtgatgtgaagctaatccccaaaacctctcagttcagattgagctgcaactcgtc
gcatggagctggaatcgtagtaatcgtatatcagcaatgatacgggtaatacgttcccggacctgtacacaccgccgt

>ebs173

acaccgaaagcgaaggcagcacattgggcccactcctgacactgaaacacgaaagcgtgggtagcgaatgggattagataccccagtagt
ccacgcctaaacgatgacaactagctatgcggagtatcaccctctgcgtggcgtagctaacgcgttaagttgtccgcctggggagtagc
gtcgaagattaaaactcaaaggaattgacggggaccgcacaagcgggtggaacatgtggttaattcaggcaacgcaagaaccttac
caggccttgacatgccctctatagaatgcagaaactattcttctcactcgggtgggaggagcacaggtgctgcatggctgctcagctcgt
gtcgtgagatgtgggttaagtcccgaacgagcgaaccctcgtcctctgttctaacaggtaatgctgagcactcaggaaactgccgg
tgataagccggaggaaggtggggatgacgtcaagtcagtagcccttacggcctgggctgcacacgtactacaatgccagtacaaaa
gatcctagaccgcgaggtggaggaatcattaaaactggggccagttcagattgaagctgcaactcgaacttcatgaagccgggaatcgt
gtaatggcgtatcagctacgacccgtgaatacgttcccgggcctgtacacaccgccgt

>ebs174

agcggtgaaatgcgtagatcgggaggaacacctgtggcgaagcggctcactggaccactgacgctgatgcacgaaagctaggg
gagcaaacaggattagataccctggtagtcttagccctaaactatgatcgttggtgacaggtaccaatcctgctgcccggagctaac
gcgttaagcgtaccgcctggggagtagcgtcgaaggctgaaactcaaaggaattgacggggggccgcacaagcgggtggagcatgtgg
ttcaattcagcgaacgcgaagaaccttacctgggctcgaagcgtagtgaccgaggtagaatatcttcccgaagggactgctacgt
aggtgctgcatggctgctcagctcgtgctgagatgtgggttaagtcccgaacgagcgaacaccttattgacgttccatcatttagtt
gggactcctggcgaacccgcctcggataacggggaggaaggtgggatgacgtcaagtcctcatggcctttatgtccagggccacac

gtgctacaatggttggtacaaaccgtcgaaaccgcgagggggagctaateggaaaaagccaatcccagtcggattggagtctgcaac
tcgactccatgaagctggaatcgtagtaatcgagatcagcacgctgcgggtaatacgtfccgggcctgtacacaccgccgt
>ebs70

gacagttcctgacgctgagggcacgaaggtcaggggagcaaacgggattagataccccggtagtcctgacagtaaacgggtgcacgtttggt
gtgggaggttcgacccttctgcgccggagtaacgcgttaacgtgccgcctggggagtacggtcgcaagattaaaactcaaagaaatt
gacggggggcccgcacaagcgggtggagtatgtggctcaatcgtatgcaacgcgaagaaccttaccaggccttgacatgattgctaagttg
gtgaaagccagcactggagcaatccagacctgacaggtgctgcatggctgtcgtcagctcgtgctgagatgttgggtaagtccc
caacgagcgaaccctgtgaactgtgccaccgagccgtgaggtcgtgagcactctgtcagactgccccgcttaacggggaggaaggt
ggggacgacgtcaagtcagtattggcccttacggcctgggctgcacacgtactacaatgccagtcagaaacgatccaagaccgcgaggt
ggaggaaatcactaaaactgggccagttcagattgaagtctgcaactcactcatgaagccggaatcgtagtaattggcgcacagctac
ggcgcctgtaatacgttcccgggcctgtacacaccgccgt

>ebs82

cagagcttgaggacgcgagaggtagagggaaactcatggttaggggtgaaatccgttgatcatggggaacaccaaagcgaaggcac
tttactggcgcgctcctgacgctacacgaaagcgtgggaatcgaacgggattagataccccggtagtcacgcctaaacggtcataa
ctggcttgcggagtatcaccctctgcgagggcaagtaacgcgttaagttatgcgcctgggtagtagcgaagattaaaactcaaag
gaaatagacggggactgcacaagcgggtgatcatcggtcaatcgtatgcaaacgaagaacctcaccagggttgaaccagacgat
cggtttgcgaaagcagacttccgaaagggcgtggcaggtgatgcatggctgtcgtcagttcgtggcttgattgttcccttcagtggggt
aacgaacgaaccgcgttgcctgttttatatgtcagggagactccctctcacgaggagggaaggtgaggatgacgccagatcagcat
gtccctctgatatcctgggctgcacgcatgatacaatgcacgcgacaacgagacgcgacggcgaagcccagcaaatctttagaaacgc
gtgctcagttcggattgaggtctgcaactcactcatgaagccggaatcgtagtaatcggggtcagctataccgcggtgaacacgttct
caagtctgtacacaccgccgt

>ebp416

gtgctcagtggtgggagaggcgcgtggaattgcaggtgtagcgggtgaaatgcgtagatatctgcaggaacaccggtggcgaagcgcacg
cgctggaccactactgacgctgaggaacgaaagctaggggagcaaacaggatcagataccctggtagtcctagccctaaacgatcagga
ctgggggtgccaccggtcgggtgctgcccggagtaacgcgttaagtcctgactggggagtacggtcgcaagactgaaactcaaagg
aattgacggggggcccgcacaagcgggtggaacatgtggtcaatcgcgctacgcgaggaaccttacctgggctcgaatgcttaggacc
agctgtgaaatacggccttcccgaagggactgagtatagggcgtgcatggctgtcgtcagctcgtgccgtgaggtgttgggttaagtcc
cgcaacgagcgaaccctgcacacagttactctgcaagagaggactctgtgtgactgctccggataacggagaggaaggtggggga
tgactcaagtcgcatggcctttatgtccagggtacacacgtgttacaatgcagggcacaaccgttccaaccgcgagggggagct
aatcggaaaaactctgctcagttcggattgaggtctgcaactcactgcatgaagccggaatcgtagtaattggcgtatcagatcgcacgc
gtgaatacgttcccgggcctgtacacaccgccgt

>ebp115

ctgggagtctgcaggcgaactaccgtgctggagtgtgggagaggtgctggaattcccgggtgtagcgggtgaaatgcgtagatatcggga
ggaacacctgtggcganagcggcgcactggaccataactgacgctgagggcgcgaaagctaggggagcaaacaggattagataccctgg
tagtcctagccctaaacgatgatggcttgggtgtagcaggtacccaatcctgctgtgccgcagtaacgcgataagccatccgcctggggagt
acggctgcaaggctgaaaccaaagggaattgacggggggcccgcacaagcgggtggagcatgtggtcaatcgcagcaacgcgaagaac
cttacctgggctcgaatgcaggagatcgggggtgaaacatcccttccccttaaaaggtgctgcataggtgctgcatggctgtcgtcagct
cgtgctgtgagatgttgggttaagtcccgaacgagcgaacccttaccctagttgccatcgggtcaagccggcactctagtgaactgc
ctcggataacgaggaggaaggtgggatgacgtcaagtcctcatggcctttatgtccagggtacacacgtgctacaatggtcggtacaaa
ccgtcgaacaccgtgagggggagtaateggaaaaagccggcctcagttcggattggagtctgcaactcactcctgaagctggaatc
gctagtaatcccggccagggcggccggagtgtaatacgttcccgggcctgtacacaccgccgt

>ebp122

cggcgcactggaccacaactgacgctgatgcgcgaaagctaggggagcaaacaggattagataccctggtagtcctagccctaaacgat
gatcgttgggtgacaggtaccaatcctgtcgtccggagtaacgcgttaagcgtatccgctggggagtacggtcgcaaggctgaaa
ctcaaaggaattgacggggggcccgcacaagcgggtggagcatcggttcaatcgcagcaacgcgaagaaccttacctgggctcgaatg
tagtgaccgggggtgaaacatccctcccgaagggctgctatataggtgctgcatggctgtcgtcagctcgtgctgagatgttgggtta

agtcccgaacgagcgcaacccttattgccagttgccatcattcagttgggactctggcgaaactgccccggataacggggaggaaggt
ggggatgacgtcaagtcctcatggcctttatgtccagggctacacacgtgctacaatgggtggtacaaccgtcgaaaccggcggggg
gagctaatcgaaaaagccatcctcagttcgattggagctgcaactcgactccatgaagctggaatcgctagtaatcgcgatcagcac
gccgcggtgaatacgttccgggcttgtacacaccgccgt

>ebp108

tgtgggagaggtgctggaattccccggtgtagcggtgaaatgcgtagatcgggaggaacacctgtggcganagcggcgcactggac
cacaactgacgtgagggcgcaagctaggggagcaaacaggattagataccctgtagtctagccctaaacgatggatgcttgggtgtg
cgggtaccaatcccagctgccgaagtaacgcgataagcatcccgcctggggagtagctgcaaggctgaaactcaaaggaattga
cggggggcccgcacaagcgggtggagcatgtggtcaattcgacgcaacgcgaagaaccttacctgggctcgaagcgcagtggaccgggg
tagaaatcccttccgcaaggaccgtcgggaggtgctgcatggctgtcgtcagctcgtgctgagatgttgggttaagtcccgaacg
agcgcaacccttatccctagttgccatcgggtcaagccgggactctagtgaaaccgccccggataacggggaggaaggtggggatgac
gtcaagtcctcatggcctttatgtccagggctacacacgtgctacaatggctggtaacaaccgtcgaaaccggcaggggggagctaactg
gaaaaagccgtcctcagttcgattggagctgcaactcgactccatgaagctggaatcgctagtaatcgagatcagcacgctcggtga
atacgttccgggcttgtacacaccgccgt

>ebp116

cgggtgtagcggtgaaatgcgtagagatcgggaggaacaccagagggcgaaggcggcgacctggacagggcactgacactcaactacgaa
agcgtggggagcaaacaggattagataccctggtagtccacgctgtaaacgatgggtactagggttggcgggtcttaccctgctgagcc
ggcgtaaacgcaataagtaccccgcctgggaagtacgctcgaagactaaaactcaaaaggaattgacggggggcccgcacaagcgggtg
atcatgtggttaattcgaagcaacgcgaagaaccttacctgggttgcacacctggaccgacgtagagatacgtcttctagcaatacagc
cgggagacaggtgctgcatggctgtcgtcagctcgtggtgacatgttgggttaagtcccgaacgagcgaacccttgccttagttgcc
agcattcagttgggtctcctaaggagactgccgtggtgaacacggaggaaggtggggatgacgtcaagtcctcatggcctttatatccagg
gcgacacacgtgatacaatggacggtacagagggagcgaaccggcaggtggagccaatctaaaaagccgttctcagttcgattgg
agtctgcaactcgactccatgaagttggaatcgctagtaatcgcgatcagcacgccgcggtgaatacgttccgggcttgtacacacc
cccgt

>ebp112

tcgcaactgacggctgaaggacgaaagttaggggagcaaacaggattagataccctgtagtctagccctaaaccatcaggacttggg
gtttgccctgtacggggcaagtcccggagctaacgcgttaagtctgcgctggggagtagcggcgaagactgaaactcaaaggaattg
acggggggcccgcacaagcgggtggaacgtgtggtcaattcgacgcaacgcgaagaaccttacctgggcttgaaccgcagggacctcc
ctagaaataggactcccgaaggacctgtgtagaggtgctgcatggctgtcgtcagctcgtccgtgaggtgttgggttaagtcccga
acgagcgaacccttactatagttgtacccgcaagggagaactctataggactgctccggataacggagaggaaggtggggatgag
gtcaagtcagcatggcctttatgtccagggctacacacgtgttaccattgcgcgccacaaggtgtcgcgaaccctgaggggggagctaactg
caaaaaagcgcgctcagttcgattgcagctgcaactcgactgcatgaagccggaatcgctagtaatggcgtatcagatcagccgctga
agacgttccgggcttgtacacaccgccgt

>ebp157

aattcgtagatattcgcaagaacaccagtgccgaaggccgctcactggcccgaactgacgctgagacgcgaaagcgtggggagcaaa
caggattagataccctggtagtccacgctgtaaacgatggatgctagccgttggggagcttcttctcagtgccgcagccaacgcttaagc
atcccgcctggggagtagatcgcaagattaaaactcaaaggaattgacggggggcccgcacaagcgggtggagcatgtggttaattcga
gcaacgcgcagaaccttaccagctttgacatgtcctggacggattgcagagatcgatcttccctcggggccggggacacaggtgctgc
atggctgtcgtcagctcgtgctgtagatgttgggttaagtcccgaacgagcgaaccctcgccttagttgccatcattcgggtgggact
ccagggggactgccggtgataagccggaggaaggtggggatgacgtcaagtcctcatggccttacaggctgggctacacacgtgctac
aatggcggtgacagcgggagggcgaaggggcaccctgagcaaatctgaaaagccgtctcagttcgattgcgctcgtgcaactcgagcg
catgaagttggaatcgctagtaatcgcgatcagcacgccgcggtgaatacgttccgggcttgtacacaccgccgt

>ebp158

aagcggcttccgtgctcagcaccgacgttccagggcagaaagcgtggggagcaaacaggattagataccctggtagtccacgccataaac
gatgagaactagacgttgagagggtaagcctcccagtgctgtagctaacgcgctaagttctccgctggggagtagcggccgcaaggttga
aactcaaaggaattgacggggaccgcacaagcgggtggagcatgtggttaattcgatgcaacgcgaagaaccttacctggccttgcacatc

ccaggaaccctgaagagatttaggggtgccgaaggagcctggagacaggtgctgcatggctgtcgtcagctcgtcgtgagatgttg
gtaagtcccgaacgagcgaaccctgtccttagtgcagcattcagttgggcactctaaggagaccgccggtgacaaaccggaggaa
ggtgggatgacgtcaagtcacatggcctttatggccagggtacacacgtgctacaatgtagatacagagggaagccaacgcgcgag
cgggagccaatcccagaaaatctatcgtagtccggattggagtctgcaactcgaactcgcattgaagtcggaatcgtagtaatcgcatcagc
attgctcgggtgaatacgttcccgggtctgtacacaccgccgt

>ebp165

tccttactgacgctgagcgcgaaagcgtggggagcaaacaggattagataccctggtagccacgctgtaaacgatgtgtcgtgatgtt
gggacacgtgctcagtgctgtagtaacgcggtaagcacaccgcctaggagtagcggccgcaaggtgaaactcaaaaggaattgg
cggggggccgcacaagcgggtggagcatgtggttaattcgaagcaacgcgcagaacctaccaggcctgacatgggcaggccgcacg
cagagatgctgttcccgaagggtcgcacacaggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcccga
cgagcgaaccctcgcctcagttgcatcggttggctgggcaactgaaaggaactgccggtgacaagccggagggaaggtggggatg
acgtcaagtctcatggccctacgtcctgggtacacacgtgctacaatggcggtgacaacgggaagccaggccgcgagccgagctg
atcccgaagcgtctcagttcagattgcaactcgggtgcatgaaggcggaatcgtagtaatcgccgatcagcatgccgcgg
tgaatacgttcccggcctgtacacaccgccgt

>ebp151

aatgcgtagatatctgcaggaacaccctggcgaaggcggctggctggaccgcaactgacgctgaggaacgaaagctaggggagcaaa
caggattagataccctggtagtccctagccctaaacctacagaacttgggggtggccgcgtaggcgggtcgtcccgtagctaacgcgtaag
ttctgcgctggggagtacggtcgaagactgaaactcaaaggaattgacggggggccgcacaagcgggtggaacatgtggttaattcga
cgcaacgcgaagaaccttacctgggctgaaccgtacggaccaacggtagaaatcgtctcccgaaggaccgtagtagaggtgct
gcatggctgtcgtcagctcgtccgtgaggtgttgggttaagtcccgaacgagcgcgaaccctactcctagttgctaccgcaaggga
actctagtaggactgctccggataacggagaggaaggtggggatgacgtcaagtcagcacggcctttatgtccagggtacacacgtgta
caatgcgcggcacagactgttgcgaaccgcgagggggagctaatggacaaaaccgcgctcagttcggattgagctcgaactcgaactcga
gcatgaagctggaatcgtagtaatggcgtatcagatcgacgccgtgaatacgttcccggcctgtacacaccgccgt

>ebp127

agatatcgggaaggaacaccagtggcgaagcggctcctggctccagcaccgacgttcaggcacgaaagcgtggggagcaaacaggatt
agataccctggtagtccacgccataaacgatgagaactagacgttgaagggtgaagccctgcagtgtcgtagctaacgcgtaagtctcc
gcctggggagtagcggccgaaggtgaaactcaaaggaattgacggggaccgcacaagcgggtggagcatgtggttaattcagatgaa
cgcaagaaccttacctggtcttgacatcctgggaacctgaagagattcgggggtgccgaaggagcccagagacaggtgctgcatgg
ctgtcgtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgcgaaccctgtccttagttgccagcattcagttggcactctaag
gagaccgccggtgacaaccggaggaaggtggggatgacgtcaagtcacatggcctttatgaccagggtacacacgtgtacaatggt
agatacagaggggaagccaacgcgcgagcgggagccaatcccagaaaaatctatcgtagtccggattggagtctgcaactcgaactcga
agtcggaatcgtagtaatcgcatcagcattgctcgggtgaatacgttcccgggtctgtacacaccgccgt

>ebp128

agaacaccagtggcgaagcggctcactggcccgatactgacgctgaggcacgaaagcgtggggagcaaacaggattaggtaccctgg
tagtccacgccgtaaacgatgaatgccagccgttagtgggttactcactagtggcgcagctaacgcttaagcattccgctggggagtag
ggtcgaagattaaaactcaaaggaattgacggggggccgcacaagcgggtggagcatgtggttaattcagcgaacgcgcagaacctta
ccagcccttgacatcccggctcggactccagagatggagtctcagttcggctggaccggagacaggtgctgcatggctgtcgtcagct
cgtgtcgtgagatgttgggttaagtcccgaacgagcgcgaacccccgtccttagttgctaccattcagttgagcacttaaggagactccc
gtgataagccgcgaggaaggtggggatgacgtcaagtcctcatggcccttacgggctgggctacacacgtgctacaacggcggtgaaa
tgggatcgaaggcgcgagccctagcaaatctcaaaaagcgtctcagatcggttgggtctgcaactcagccccatgaagttggaatc
gctagtaatcgtagcagcacgccacgggtgaatacgttcccggcctgtacacaccgccgt

>ebp152

gctggaccactgacgctgaggaacgaaagctaggggagcaaacaggattagataccctggtagtccctagccctaaacgatcaggact
tgggtgccacccttcgggtgtcgtcccggagctaacgcgtaagtctgcacctggggagtaggtcgaagactgaaactcaaagga
attgacggggggccgcacaagcgggtggaacatgtggttaattcagcgtacgcgaggaaccttacctgggctcgaagcgcacttgaaa
gctgtagaaatcggcctcccgaaggacaggtgtggaggcgtgcatggctgtcgtcagctcgtccgtgaggtgttgggttaagtcc

cgcaacgagcgaacccctgcacgcagttgccactcgaagagggaaactctacgtggactgctccggataacggagaggaaggtgggg
atgacgtcaagtccgcatggcctttatgtccagggctacacacgtgttacaatgcagggcacaaaccgttccaaccgcgaggggggagc
taatcggaaaaaactctgctcagttcgattgcagctgcaactcactgcatgaagccggaatcgtagtaatggcgtatcagatcgacgc
cgtgaatacgttcccgggccttgtacacaccgcccgt

>ebp153

caggaacaccgctggcgaaggcggctggctggaccgcaactgacgctgaggaacgaaagctaggggagcaaacaggattagataccc
tgtagtcttagccctaaaccatcagaacttgggggtggcccgcgtaggggggctgctccgtagtaacgcgttaagtctgcgctggggga
gtacggctcgaagactgaaactcaaaaggaattgacggggggcccgcacaagcgggtggaacatgtggttcaattcgacgcaacgcgaagaa
ccttacctgggcttgaaccgtacggaccaacggtagaataatcgtcttcccgaagggaccgtagtagaggtgctgcatggctgtcgtca
gtcgtgccgtgaggtgttgggttaagcccgaacgagcgaaccctactcctagttgctaccggcaagggagaactctagtaggactg
ctccggataacggagaggaaggtggggatgacgtcaagtgcacggcctttatgtccagggctacacacgtgttacaatgcgcggcaca
gactgttgcgaacccgcgaggggagctaatcggacaaaaccgcgctcagttcggattgcagctgcaactcactgcatgaagctggaa
tcgtagtaatggcgtatcagatcgaccgtgaatacgttcccgggccttgtacacaccgcccgt

>ebp154

aggtgtagcgggtgaaatcgtagatatctggaggaacatcagttggcgaaggcggcttctggcccaactgacgctgaggtgcgaaagc
gtgggggagcaaacaggattagataccctgtagtccacgctgtaaacgatgagaactagtcgttggagggctgcttccagtgcgtagc
taacgcgctaagtctccgctggggagtagggccgcaaggtgaaactcaaaaggaattgacggggggcccgcacaagcgggtggagtag
tggttaattcgtgcaacgcgaagaaccttacctggtcttgacatcctcggaaactcgcagagatgctgggagtgcttccgggaactgagag
acaggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagctccgcaacgagcgaaccctcatccctagttgccatcatta
agttgggaactctagggagactgccgggtgacaaaccggaggaaggtggggatgacgtcaagtcacatggcccttatgaccagggctac
acacgtactacaatgctcgtacagagggtagccaagccgcgaggtggagccaatcccagaaaaccgatcgcagtcggattggagctc
gcaactcactccatgaagtcggaatcgtagtaatcgtggatcagcaatgccacgggtgaatacgttcccgggccttgtacacaccgcccgt

>ebp136

ggcttctgtagcgaactgacgctgggggacgaaagctaggggagcaaacaggattagataccctgtagtcttagccctaaaccatca
ggacttgggggttgcctgtacggggcaagtcccggagctaacgcgttaagtctgcgctggggagtagcgtcgaagactgaaactca
aaggaattgacgggggtccgcacaagcgggtggaacatgtggttcaattcgacgcaacgcgaagaaccttacctgggcttgaaccgaca
ggacctccctagaaataggacttcccgaaggacctgtgtagaggtgctgcatggctgtcgtcagctcgtgccgtgaggtgttgggtta
agtcccgaacgagcgaacccttactcatagttgctaccgcaagggagaaactctatgaggactgctccggataacggagaggaaggtg
gggatgaggtcaagtcagcatggcctttatgtccagggctacacacgtgttacaatgcgcgccacaaagtgtcgcgaaccctgaggggg
agctaatcgaaaaaagcgcgctcagttcggattgcagctgcaactcactgcatgaagccggaatcgtagtaatggcgtatcagatcg
acgccgtgaatacgttcccgggccttgtacacaccgcccgt

>ebp155

atattcgaagaacaccagtggcgaaggcggctcactggcccggaaactgacgctgagacgcgaaagcgtggggagcaaacaggattag
ataccctgtagtccacgctgtaaacgatggatgtagccgttggggagcttcttctcagtgccgcagccaacgctttagcatcccgcct
ggggagtagcatcgaagattaaactcaaaaggaattgacggggcccgcacaagcgggtgagcatgtggttaattcgaagcaacgcg
cagaacctaccagctttgacatgtcctggacggattgcagagatgcgatcttcccctcggggccggggacacaggtgctgcatggctgtc
gtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgaaccctcgccttagttgccatattcgggtgggactccaggggg
actgccggtgataagccggaggaaggtggggatgacgtcaagtcctcatggcccttacaggctgggctacacacgtgtacaatggcgg
gacagcgggagggcgaagggggcaccctgagcaaatctcgaagccgtctcagttcggattgctcgtcgaactcagcgcgatgaagtt
ggaatcgtagtaatcgcggatcagcagccggtggaatacgttcccgggccttgtacacaccgcccgt

>ebp164

cgggtgaaatcgtagatatcaggaggaacaccgggtggcgaaggcggctacctggacagaaactgacgctgaggcacgaaagctaggg
gagcaaacaggattagataccctgtagtcttagccctaaacgatggatacttgggtgaccgggattgaatccggctgtgccgaagctaac
gcattaagtatcccgcctggggagtagcgtcgaaggctgaaactcaaaaggaattgacggggggcccgcacaagcgggtggagtagtgggt
taattcgcagcaacgcgaagaaccttacctgggctcgaacggctagtgacaggtggcgaaagtcgctttcccgaagggacactagtcg
aggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagctccgcaacgagcgcgaccctgcctttagttccaccaggtta

tgctgagcactctaaaggactgccggtgaaaaaccggaggaaggtggggatgacgtcaagtctcatggcctttatgtccagggctaca
cacgtactacaatgggcggcacaatgagttgcaaaccgcgagggggagccaatctaaaaaccgctctcagttcggattgcagtctgc
aactcgactgcatgaagctggaatcgtagtaatcgctgatcagcaagcagcgggtgaatacgttcccgggccttgtaacaccgccgt
>ebp141

gacagaaactgacgctgaggcacgaaagctaggggagcaaacaggattagataccctggtagtcctagccctaacgatggatacttgc
tgactgggattgaatccagtcgtggcgaagctaacgcattaagatcccgcctggggagtagcgtcgaaggtgaaactcaaaggaatt
gacggggggcccgcacaagcgggtggagatgtggttaattcgacgcaacgcgaagaaccttacctgggctcgaacggctagtgacaggt
ggcgaagtcgctttcccgaaggacactagtcgaggtgctgcatggctgctgtagctcgtgctgtagatgtgggttaagtcccga
acgagcgcgaaccctgccttagttgccaccaggttaagctgagcactctaaaggactgccggtgaaaaaccggaggaaggtggggat
gacgtcaagtctcatggcctttatgtccagggctacacagctactacaatgggcggcacaatgagttgcaaaccgcgagggggagcca
atctcaaaaaaccgctctcagttcggattgcagtctgcaactcgactgcatgaagctggaatcgtagtaatcgctgatcagcaagcagcg
tgaatacgttcccgggccttgtaacaccgccgt
>ebp167

gagtgtagaggtgaaattcgtagatattcgcaagaaccagtgggcgaaggcggctcactgccccgatactgacgttgaggcacgaaagc
gtggggagcaaacaggattaggtaccctggtagtcacgccgtaaacgatgaatgccagccgttagtgggttactactagtggcgcagc
taacgctttaagcattccgcctggggagtagcgtcgaagattaaactcaaaggaattgacggggggcccgcacaagcgggtggagcatgt
ggttaattcgacgcaacgcgcagaacctaccagccctgacatcccggctcgggactccagagatggagttctcagttcggctggaccg
gagacaggtgctgcatggctgctgtagctcgtgctgtagatgtgggttaagtcccgaacgagcgcgaacccccgccttagttgctacc
atcagttgagcactctaaaggagactcccgggtgataagccgcgaggaaggtggggatgacgtcaagtcctcatggccttacgggctggg
ctacacagtgctacaacggcgggtgacaatgggatgcaagggcgcgagccctagcaaatctaaaaagccgctcagatcggattgggct
ctgcaactcgagccatgaagttggaatcgtagtaatcggtgatcagcacgccacgggtgaatacgttcccgggccttgtaacaccgcc
gt

>ebp170

ctaggggagcaaacaggattagataccctggtagtcctagccctaacgatgattgcttgggtgtagcaggtacccaatcctgctgtgccga
gtaacgcgataagcaatccgcctggggagtagcgtcgaaggctgaaactcaaaggaattgacggggggcccgcacaagcgggtggagc
atgtggttcaattcgacgcaacgcgaagaaccttacctgggctcgaagcgcagtgatcgggggtgaaacatccctcccctaaaaggct
gctcggaggtgctgcatggctgctgtagctcgtgctgtagatgtgggttaagtcccgaacgagcgaacccctatccctagttgccat
cgggtcaagccgggactctagtgaactgcctcggataacgaggaggaaggtggggatgacgtcaagtcctcatggcctttatgtccag
ggctacacagtgctacaatggtcggtaacaaccgtcgaacccgtgagggggagctaatcggaaaaagccggcctcagttcggattg
gagctgcaactcgactccatgaagctggaatcgtagtaatcgcgatcagaacgccgcgggtgaatacgttcccgggccttgtaacacc
gccgt

>ebp413

ggtgtagcgggtgaaatcgtagatcgggaggaacacctgtggcgaagcggcgcactggaccataactgacgctgaggcgcgaaag
ctaggggagcaaacaggattagataccctggtagtcctagccctaacgatgatggcttgggtgtagcaggtacccaatcctgctgtgccga
gtaacgcgataagccatccgcctggggagtagcgtcgaaggctgaaactcaaaggaattgacggggggcccgcacaagcgggtggagc
atgtggttcaattcgacgcaacgcgaagaaccttacctgggctcgaatgcaggagaccgggggtgaaacatccctcccctaaaaggct
gctgcataggtgctgcatggctgctgtagctcgtgctgtagatgtgggttaagtcccgaacgagcgcgaacccctatccctagttgccat
cgggtcaagccgggactctagtgaactgcctcggataacgaggaggaaggtggggatgacgtcaagtcctcatggcctttatgtccag
ggctacacagtgctacaatggtcggtaacaaccgtcgaacccgtgagggggagctaatcggaaaaagccggcctcagttcggattg
gagctgcaactcgactccatgaagctggaatcgtagtaatcgcgatcagaacgccgcgggtgaatacgttcccgggccttgtaacacc
gccgt

>ebp411

ttagcgggtgaaatcgtagatcaggaggaacaccgggtggtgtaaacggcttctggaccataactgacgctgaggcacgaaagcgt
gggtagcaaacaggattagataccctggtagtcacgccctaacgatgcatattggtgtagggccattcatttggctgtgccggagctaac
gcgttaaatatgccgctggggagtagcgtcgaaggctgaaactcaaaggaattgacggggggcccgcacaagcgggtggagcatgtgt
ttaattcgacgcaacgcgaagaaccttacctgggctcgaatgctatggaccagctggtgaaagccggcctcccgaaggaccgtagta

tagtgctcatggctgtcgtcagctcgtgtcgtgagatgttgggtaagtcccgaacgagcgaacccttgcctgtgtgccatcattcag
ttgggaactcgcaggagaccgccagggataacctggaggaaaggtgggggatgacgtcaagtcctcatggcctttatgtccagggtacaca
cgtgctacaatggacggtgtaaaagggctgccaaccgcaagggggagccaatccaaaaaacggtctcagttcggattgcaggctgca
actcgcctgcatgaagctggaatcgttagtaatggtggatcagaacgccaccgtgaatacgttcccgggccttgtacacaccgccctg
>ebp146

agataticgggaggaacacctgtggcgaaagcggcgcactggaccataactgacgctgagggcgcgaaagctaggggagcaaacaggatt
agataacctggtagtcctagccctaaacgatgattgcttgggtgtagcaggtacccaatcctgctgtgccgcagctaacgcgataagcaatcc
gcctggggagtacggtcgaaggctgaaactcaaaggaattgacggggcccgcacaagcgggtggagcatgtggttaattcagcgaa
cgcgaagaaccttacctgggctcgaagcgcagtggtatcgggggtgaaacatccctcccctaaaaggctgctcggagggtgctgcatgg
ctgtcgtcagctcgtgtcgtgagatgttgggtaagtcccgaacgagcgaaccctatccctagttgccatcgggtcaagccgggactct
agtgaactgcctcggataacgaggaggaaggtgggggatgacgtcaagtcctcatggcctttatgtccagggtacacacgtgctacaatg
gtcggtaaaaaccgtcgaaaaccgtgagggggagctaatacggaaaaagccggcctcagttcggattggagctgcaactcgaactccatg
aagctggaatcgttagtaatcgcggatcagaacgccgcgggtaatacgttcccgggccttgtacacaccgccctg
>ebp409

agcggtgaaatcgtagataticgggaggaacacctgtggcgaaagcggctcactggaccacaactgacgctgagggcgcgaaagctagg
ggagcaaacaggattagataacctggtagtcctagccctaaacgatgattgcttgggtgtcgggtaccaatcccacgtgccgcagctaa
cgcgataagcaatccgctggggagtacggtcgaagactgaaactcaaaggaattgacggggggcccgcacaagcgggtggagcatgtg
gttcaattcagcgaacgcgaagaaccttacctgggctcgaatgtagtggaccggggtgaaacatcccctcccgaagggtgctatat
aggtgctgcatggctgtcgtcagctcgtgtcgtgagatgttgggtaagtcccgaacgagcgaacccttattgccagttgccatcattagt
tgggactctggcgaaccgccccgataacggggaggaaggtgggggatgacgtcaagtcctcatggcctttatgtccagggtacacac
gtgctacaatggtcggtaaaaaccgtcgaaaaccgtgagggggagctaatacggaaaaagccggcctcagttcggattgaagtctgcaac
tcactcatgaagctggaatcgttagtaatcgcggatcagaacgccgcgggtaatacgttcccgggccttgtacacaccgccctg
>ebp148

tagataticggaagaacaccagtggcgaaggcgactcactggcccgttactgacgctgagggcgcgaaagcgtggggagcaaacaggatt
agataacctggtagtcacgctgtaaacgatgagtctagtgtcggcatgcatgctcgggtgacgcagctaacgcattaagcactccgc
ctggggagtacggtcgaagattaaaactcaaaggaattgacggggggcccgcacaagcgggtggagcatgtggttaattcgaagcgacg
cgcgaaccttaccacctttgacatgctggaccgttagagagatctaactttcccttcggggactaggacacaggtgctgcatggctgtc
tcagctcgtgtcgtgagatgttgggtaagtcccgaacgagcgaaccctcgcattagttgccatcatttagttgggactctaattgggacc
gccgggtgtaagccggaggaaggtgggggatgacgtcaagtcctcatggccttacggggtgggtacacacgtgctacaatggcgacta
cagaggggtgcaaacccgcgagggggagccaatccctaaaagtcgtctcagttcggattgactctgcaactcgaagtgatgaaagttgga
atcgttagtaatcgcggatcagcatgccgcgggtaatacgttcccgggccttgtacacaccgccctg
>ebp408

ctaggggagcaaacaggattagataacctggtagtcctagccctaaacgatgattgcttgggtgtcgggtaccaatcccacgtgccgca
gctaacgcgataagcaatccgctggggagtacggtcgaagactgaaactcaaaggaattgacggggggcccgcacaagcgggtggagc
atgtggttaattcagcgaacgcgaagaaccttacctgggctcgaatgtagtggaccggggtgaaacatcccctcccgaagggtgctc
tatataggtgctgcatggctgtcgtcagctcgtgtcgtgagatgttgggtaagtcccgaacgagcgaacccttattgccagttgccatcatt
tagttgggactctggcgaaccgccccgataacggggaggaaggtgggggatgacgtcaagtcctcatggcctttatgtccagggtac
acacgtgctacaatggtcggtaaaaaccgtcgaaaaccgtgagggggagctaatacggaaaaagccggcctcagttcggattgaagtctg
caactcgaactcatgaagctggaatcgttagtaatcgcggatcagaacgccgcgggtaatacgttcccgggccttgtacacaccgccctg
>ebp407

ttgagtcgggagaggtgagtggaactcgcagtgtagagtgaaaattcgtagataticggaagaacaccagtggcgaagccggctcactg
gcccgtactgacgctgaggtgcgaaagcgtggggagcaaacaggattagataacctggtagtcacgccgtaaacgatggatgtagc
cgttgtcgggtttaccgtcagttggcgcagctaacgcattaagcatcccgcctggggagtacggtcgaagattaaaactcaaaggaattga
cggggggcccgcacaagcgggtggagcatgtggttaattcgaagcaacgcgaagaaccttaccagcctttgacatcccggctcggggcac
cagagatggagccctcagttcggctggaccggagacaggtgctgcatggctgtcgtcagctcgtgtcgtgagatgttgggtaagtccc
caacgagcgaaccctcgccttagttgccatcatttagttgggactctaaggggactgccggtgataagccgcgaggaaggtggggat

gacgtcaagtcctcatggcccttacaggctgggctacacacgtgctacaatggcggtgacaatgggacgcaatggggaaccctgcgcaa
atctcaaaaagccgtctcagttcggattgtgetctgcaactcagacacatgaagttggaatcgctagtaatcgagatcagcacgctgcggtg
aatacgttcccggcctgtacacaccgccgt

>ebp102

agcgggtgaaatgcgtagatcgggaaggaaaccagtgggcgaagcggcttctgggtccagcaccgacgttcaggcacgaaagcgtgg
ggagcaaacaggattagataccctggtagtccacgccataaacgatgagaactagacgttgagagggtaaacctcccagtgctgtagctaa
cgcgctaagttctccgctggggagtacggccgcaaggtgaaactcaaaggaattgacggggaccgcacaagcgggtggagcatgtgg
ttaaattcagtgcaacgcgaagaaccttacctggccttgacatcccaggaacctgaagagatttaggggtgccgaaggagcctggagac
aggtgctgcatggctgtcgtcagctcgtgctgtagatgttgggttaagtcgccgaacgagcgaaccttgccttagttgccagcattcagt
tgggactctaaggagaccgccgggtgacaaaccggaggaaggtggggatgacgtcaagtcacatggcctttatggccagggtacaca
cgtgctacaatgtagatacagagggaagccaacgcgcgagcgggagccaatcccagaaaatctatcgtagtcggattggagctgcaa
ctcactcctgaagtcggaatcgctagtaatcgagatcagcattgctgcgggtgaatacgttcccgggtctgtacacaccgccgt

>ebp110

agtatgggagaggtgagtggaattcccgggtgtagcgggtgaaatgcgtagatcgggaggaacacctgtggcgaagcggctcactgga
ccattactgacgctgatgcacgaaagctaggggagcaaacaggattagataccctggtagtccctagccctaaacgatgatcgttgggtga
caggtaccaatcctgttggccggagctaaccgcttaagcgtaccgctggggagtacggcgcgaaggctgaaactcaaaggaattgac
gggggcccgcacaagcgggtggagcatgtggtcaattcagcgaacgcgaagaaccttacctgggctcgaagcgtatggatggaggta
gaaatctctttcccgaagggactgctacgtaggtgctgcatggctgtcgtcagctcgtgctgtagatgttgggttaagtcgccgaacga
gcgaaccttattgccagttgccatcattcagttgggactctggcgaaacgccccggataacggggaggaaggtggggatgacgtca
agtctcatggcctttatgtccagggtacacacgtgctacaatggttgggtacaaccgtcgaaacccgcgagggggagctaactggaaa
aagccaatcccagtcggattggagctgcaactcactcctgaagctggaatcgctagtaatcgagatcagcattgctgcgggtgaatac
gttcccggcctgtacacaccgccgt

>ebp161

ctaggggagcaaacaggattagatacccggtagtcctaaacgatcaggacttggggtgccaccgftcgggtgtcgtcccggagctaacg
cgtaagtctgcacctggggagtacggcgaagactgaaactcaaaggaattgacgggggccgcacaagcgggtggaacatgtggttc
aattcagcgtacgcgaggaaccttacctgggctcgaatgcttctgaccagctgtagaaatacggccttcccgaagggacaggagtata
ggcgtgcatggctgtcgtcagctcgtccgtgaggtgttgggttaagtcgccgaacgagcgaaccttgcacgtatgtccactcgaac
gagggaactctactgagctgctccggataacggagaggaaggtggggatgacgtcaagtccgatggcctttatgtccagggtacaca
cgtgttacaatgcagggcacaaccgttccaaccgcgagggggagctaactggaaaaactctgctcagttcggattgacgtctgcaac
tcgactgcatgaagccggaatcgctagtaatggcgtatcagatcagcggcgtgaatacgttcccggcctgtacacaccgccgt

>ems142

cgctcaggtgcgaaagccaggggagcgaacgggattagataccccggtagtcctggccgtaaacgatgaccactagatgtcggggtat
cgaccctcgtgctggagttaacgcgctaagtgtccgctggggagtacggcgcgaagattaaaactcaaaggaattgacggggacc
cgcacaagcagcggagcgtgtggttaattcagtgcaacgcgaagaaccttaccagggcttgacatggcattgaccaccgagagatcgg
gtcccttcggggcgggtgccaccggtgctgcatggctgtcgtcagctcgtgctgtagatgttgggttaagtcgccgaacgagcgaacct
ctatgctggttggagatgtccggcagactgccacgatcaacgtggaggaaggtggggatgacgtcaagtcagcacggccgtacgtc
ctgggcgacacacagctacaatggccggtacaatgggacgccaaccgcaagggggagccaatccagcgaaacgggtctcagttc
gattggaggctgcaaccgctccatgaagctggagttgctagtaaccgtgggtcagcacaccacgggtgaatacgttcccgggtctgtac
acaccgccgt

>ems116

agtgtgaaatgcgtagatcgggaggaacccccgaaggcgaaggcactgtcctggttcattgctgacgctcaggggcgaaagccagg
ggagcgaacgggattagatacccggtagtcctggctgtaaacgatgaccactagacgttcggggatcgcaccctgagtgctgcagcta
acgcgagaagtgtccgctggggagtacggcgaagattaaaactcaaaggaatagacggggaccgcacaagcagcggagcgtgt
ggttaattcagtgcaacgcgaagcacttaccaggggttgacatggcagtgacccccgtgagagcgggcttcttcgggacactgtcaca
ggtgctgcatggctgtcgtcagctcgtgccgtgaggtgttgggttaagtcgccgaacgagcgaacctcaggttagttcttctaccg
agactgccgcgatcaacgtggaggaaggtggggatgacgtcaagtcagcacggcccttacctcggcgacacacagctacaatggc

gaggacaacgagttccaacccgcgaggggagctaactcaaacctcgtcccagttcagattgcaggctgcaacccgcctgcatgaa
gtcggagttgctagtaaccgtgagtcagcacatcacggtgaatccgtcccgggtctgtacacaccgccgt
>ems118

aatgcgtagatatacgggaggaacaccagtgggcgaaggcgatcacctgggggtgtctgacgctgaggagcgaagctaggggagcaaa
cgggattagatacccccgtagcttagccgtaaacgatggatactaggtgtatgtggtatcgactccatatgtccgtagttaacacagtaagt
atcccgcctggggagtagcggccgaaggttaaaactcaaaggaattgacggggcccgcacaagcggaggagcatgtggttaattcgac
gcaacgcgaagaacctaccgggcttgacatcggacgacatccgtagaaatacggtttccccgcaaggggacgtccagacaggtgctg
catggctgtcgtcagctcgtccgtaggtgttgggttaagtcccgaacgagcgaacccctcgtcgtgtgttgctgcaagatctctaca
cgagaccgccggtgagaagccggaggaaggtggggatgacgtcaagtcagcatggccctacgtccgggctacacacgtgctacaat
gggtaggacaacgagttgcgaactcgcgagagtaagccaatctctaaacctatccccagttcggattgaaggctgcaactcgcctcatga
agtcggaatcgtagtaaccgcagatcagctacgctgcggtgaatcggttcccgggctgtacacaccgccgt
>ems120

agatatacgggaggaacacctgtggcgaagcggcgcactggaccataactgacgctgaggcgcgaaagctaggggagcaaacaggatt
agataacctgtagtcttagccctaacgatgattgcttgggtgtagcaggtacccaatcctgctgtccgcagctaacgcgataagcaatcc
gcctggggagtagcgtcgaaggctgaaactcaaaggaattgacggggcccgcacaagcggaggagcatgtggttaattcgacgcaa
cgcgaagaaccttacctgggctcgaatgcaggagaccgggggtggaacatcccttcccctaaaaggctgctgcataggtgctgcatgg
ctgtcgtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgaacccctatccctagttgccatcgggtcaagccgggactct
agtgaactgcctcgggataacgaggaggaaggtggggatgacgtcaagtcctcatggcctttatgtccagggctacacacgtgctacaatg
gtcggtaaaaaccgtcgaaacccgtgagggggagctaatacggaaaaagccggcctcagttcggattggagctgcaactcgcactccatg
aagctggaatcgtagtaatcgcggatcagaacgccgcggtgaatcggtcccgggctgtacacaccgccgt
>ems418

agcgggtgaaatcgtagatatacgggaggaacacctgtggtgcaaacggctctctggaccgtaactgacgctgagacacgaaagcgtggg
gagcaaacaggattagataccctgtagtccacgccctaaacgatgcacatttgggtggtggccattcatttggctcgtccggagctaacc
gttaaatgtgccgctggggagtagcggcgaaggctgaaactcaaaggaattgacggggcccgcacaagcggaggagcatgtggtc
aattcgacgcaacgcgaagaaccttacctgggctcgaacggctggggaccagcgggtgaaacatcgcctcccgaagggaccgcagt
cgaggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgaacccctgtccattgtgccataccgt
aaaaagtggaactcgtgagaccgccagcgaagctggaggaaggtggggatgacgtcaagtcacatggcctttatgtccagggcta
cacacgtgctacaatggcgggtacagtggtcgttaaccgcgagggggagccaatccccaaaagccgtcctcagttcggattgcaggc
tgcaactcgcctgcatgaagctggaatcgtagtaatgctgcatcagaacggtagcgtgaatcggtcccgggctgtacacaccgccgt
>ems403

gacgctgaggcgcgaaagcgtggggagcaaacaggattagataccctgtagtccacgctgtaaacgatgtgcgctagacgttggagga
cgtagtctccggtgtcgcagcaaacgcgataagcgcaccgcctggggagtagcggccgaaggttaaaactcaaaggaattgacggggg
cccgcacaagcggaggagcatgtggttaattcgacgcaacgcgaagaacctaccagccctgacatgtcccttaagggccgagagat
cggcttcgtcgggtcggccggggggaacacaggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcccgaacgag
cgcaacctcacctcagttgccatcgggttaagctgggactctgaaggaactcccgggtgacaagccggagggaaggcggggatgacgt
caagtcctcatggcctttatgggctgggctacacacgtgctacaatggcgggtgacagaggcggcaagggagcgtcctgcgcaatctc
gaaaagccgtcagttcggattgttctctgcaactcagagatgaagctggaatcgtagtaatcgcggatcagcacgccgggtgaata
cgtcccgggctgtacacaccgccgt
>ems419

ggaacacctgtggtgcaaacggctctctggaccgtaactgacgctgagacacgaaagcgtggggagcaaacaggattagataccctgt
agtccacgccctaaacgatgcacatttgggtggtggccattcatttggctcgtccggagctaaccggttaaatgtccgctggggagtagc
gtcgaaggctgaaactcaaaggaattgacggggcccgcacaagcggaggagcatgtggttaattcgacgcaacgcgaagaacctta
cctgggctcgaacggctggggaccagcgggtgaaacatcgcctcccgaagggaccgagtcgaggtgctgcatggctgtcgtcagct
cgtgtcgtgagatgttgggttaagtcccgaacgagcgaacccctgtccattgttgcataaccgtaaaaagtggaactcgtgagaccgc
cagcgacaagctggaggaaggtggggatgacgtcaagtcacatggcctttatgtccagggctacacacgtgctacaatggcgggtacag

tgggtcgctaaccgcgagggggagccaatccccaaaagccgtcctcagttcgattgcaggctgcaactcgctgcatgaagctggaat
cgctagtaatgctgcatcagaacggtagcgtgaatacgttcccgggcctgtacacaccgccgt

>ems406

gagcaaacaggattagataccctggtagtccacgctgtaaacgatgtgcgctagacgttggaggacgtagtctccgggtgcgagcaaac
gcgataagcgcacgcctggggagtagcggccgcaagggttaaaactcaaaggaattgacggggggccgcacaagcgggtggagcatgtg
gttcaattcgacgcaacgcgaagaaccttaccagcccttgacatgtcccttaagggggccgagagatcggttcgtcggttcggccggggg
gaacacaggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgaacacctcagttcagttgccat
cgggttaagctgggcactctgaaggaactgccggtgacaagccggaggaaaggcgggggatgacgtcaagtcctcatggcctttatgggct
gggctacacacgtgctacaatggcgggtgacagagggcggcaaggagcgtcctgcgccaatctgaaaagccgtctcagttcggattgt
tctctgcaactcgagagcatgaagctggaatcgctagtaatcgcgatcagcacgccgcgggtgaatacgttcccgggcctgtacacaccg
cccgt

>ems407

gagcaaacaggattagataccctggtagtccaaacgatgcatactgggtgtgagccgttcattcggttcgtccggagctaacgcattaagt
atccgcctggggagtagcggcgaaggctgaaactcaaaggaattgacggggggccgcacaagcgggtggagcatgtggttaattcga
cgcaacgcgaagaaccttacctgggctcgaacggcttcgaaagatcgtgaaacacggtttctcccgaagggatcgagtcgaggtg
ctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgaacacctgtcctgtttggcatgtgtcacgcaagt
ggcgcgcaactatcaggagaccgccagcgataagttggaggaagggtggggatgacgtcaagtcacatggcctttatgtccagggtaca
cacgtgctacaatggcggtacaaaagcgtcgaacccgcgagggggaagtaatcgaaaaaaccgtcctcagttcggatcgaggtg
caattcgctgcgtgaagctggaatcgctagtaatggcatatcagaacgatgccgtgaatacgttcccgggcctgtacacaccgccgt

>ems412

agcgggtgaaatcgctagatcagagggaacacttgccgtgtagacggctttctgacctcactgacgctgagacacgaaagcgtgggg
agcaaacaggattagataccctggtagtccacgccctaaacgatgcatactgggtgtgagccgttcattcggttcgtccggagctaacgcat
taagtatccgcctggggagtagcggcgaaggctgaaactcaaaggaattgacggggggccgcacaagcgggtggagcatgtggttaat
tcgacgcaacgcgaagaaccttacctgggctcgaacggcttcgaaagatcgtgaaacacggtttctcccgaagggatcgagtcga
gggtcgtcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgaacacctgtcctgtttggcatgtgtcacgc
aagtggcgcgcaactatcaggagaccgccagcgataagttggaggaagggtggggatgacgtcaagtcacatggcctttatgtccagggc
tacacacgtgctacaatggcggtacaaaagcgtcgaacccgcgagggggaagtaatcgaaaaaaccgtcctcagttcggatcgag
gctgcaattcgctgcgtgaagctggaatcgctagtaatggcatatcagaacgatgccgtgaatacgttcccgggcctgtacacaccgcc
gt

>ems413

cctggcacagggccttgcagactgtgagacttgaggcctcagagggacatggaactccaggtgtagtgggaaatcgctagatctgg
aagaacaccaaagcgaaggcagtgctcctgggaggtacctgacgctcaggtgcgaaagccaggggagcgaacgggattagatacccc
ggtagtcctggccgtaaacgatgaccactagatgtcgggggtatcgacccctgcgtgctggagttaacgcgtaagtggcgcctggg
gagtagcgtcgaagattaaaactcaaaggaattgacggggaccgcacaagcagcggagcgtgtggttaattcagtgcaacgcgaag
aaccttaccagggcttgacatggcattgcaccaccgagagatcggtgtccctcggggcgggtgccaccgggtgctgcatggctgtcgtcagc
tcgtgtcgtgagatgttgggttaagtcccgaacgagcgaacccctatcgtcggttggagatgtccggcgagactgccacgatcaacgtg
gaggaagggtggggatgacgtcaagtcagcacggcccgtacgtcctgggcgacacacacgctacaatggccggtacaatgggacgcaa
cccgaagggggagccaatccagcgaaaaccggctcagttcggattggaggctgcaacccgctccatgaagctggagttgctagtaac
cgtgggtcagcacaccacgggtgaatacgttcccgggtctgtacacaccgccgt

>ems147

agatatctggaagaacaccaaagcgaaggcagtgctcctgggaggtacctgacgctcaggtgcgaaagccaggggagcgaacgggatt
cgataccccggtagtctggccgtaaacgatgaccactagatgtcgggggtatcgacccctgcgtgctggagttaacgcgtaagtggtc
cgcttggggagtagcggcgaagattaaaactcaaaggaattgacggggaccgcacaagcagcggagcgtgtggttaattcagtgcaac
cgcaagaaccttaccagggcttgacatggcattgcaccaccgagagatcggtgtccctcggggcgggtgccaccgggtgctgcatggctg
tcgtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgaacccctatcgtcggttggagatgtccggcgagactgccacgat
caactgggaggaagggtggggatgacgtcaagtcagcacggcccgtacgtcctgggcgacacacacgctacaatggccggtacaatggg

acgccaaaccgcaagggggagccaatccagcgaaaaccggtctcagttcggattggaggctgcaaccgcctccatgaagctggagttg
ctagtaaccgtgggtcagcacaccacggtgaatacgttcccgggtcgtgtacacaccgccgt

>ems408

ctgtgagacttgagggcctcagagggacatggaactccaggtgtagtggtgaaatgcgtagatatctggaagaacaccaaaggcgaagg
cagtgctctgggaggtacctgacgctcaggtgcgaaagccaggggagcgaacgggattagatccccggtagtctggccgtaaacgat
gaccactagatgtgcggggtatcgaacctcgtgctggagttaacgcgctaagtggctccgcctggggagtacggcgcgaagattaaaa
ctcaaaggaattgacggggaccgcacaagcagcggagcgtgtggttaattcagtgcaacgcgaagaaccttaccagggttgacatgg
cattgaccaccgagagatcgggtgcccctcggggcgggtgccaccgggtgctgcatggctgctgacgctcgtgctgtgagatgtgggtaa
gtcccgaacgagcgaaccctatcgtcgggtggagatgtccggcgagactgccacgatcaacgtggaggaaggtggggatgacgtca
agtgcacggcccgtacgtcctgggagcacacacgctacaatggccggtacaatgggacgccaaccgcaagggggagccaatcc
agcgaaaaccggtctcagttcggattggaggctgcaaccgcctccatgaagctggagttgctagtaaccgtgggtcagcacaccacggt
gaatacgttcccgggtctgtacacaccgccgt

>ems450

gactgccaggctagaggacggaagaggagcgcggagttcccgggtgtagcggtaagtgcgtagagatcgggaggaaggccgggtggc
gaaggcggcgtctggaacgttctgacgctgaggcacgaaagcgtggggagcaaacaggatcagataccctgtagtccacgcctaa
acgatggacactaagtgtcggcgggtaccgtcgggtccgcagctaaccgagtaagtgtcccgcctggggagtacggccgcaaggttga
aactcaaaggaattgacggggcccgcacaagcgggtggagcatgtggttctatcagcgaacgcgaagagccttaccagctctgacat
gcctcgtatggggccgagaaaatcggttccttcagttcgggtggcggggacacaggtgctgcatggctgctgacgctcgtgctgtgagatgt
tgggttaagtcccgaacgagcgaaccctcctccttagttgccatcagggttaagctggcactctgaggaaactgccggtgacaagccgg
aggaaggcggggatgacgtcaagtcctcagttcggccttatgggctgggctacacacgtgctacaatggcgggtgacagggggcggcggg
cagcgtatccaagccaatctcaaaaagccgtctcagttcggattgctctctgcaactcagagcatgaagtcggaatcgtagtaatcggg
atcagcatgcccggtgaatacgttcccggcctgtacacaccgccgt

>ems451

agagagggacctggaattccgggtggagcagtgaaatgcgtagagatccggaggaacacaaaaggcgaaggcagggtcctggctggc
acctgacgctcagggcgaaagctaggggagcgaacgggattagatccccggtagtcttagccgtaaacgatgaccactagatgcgcg
gggttcgacctcgtgcccagctaacgcaataagtggtccgcctggggagtacggcgcgaagactaaaactcaaaggaattgacg
gggaccgcacaagcagcggagcgtgtggttaattcagcgcacagaagaaccttaccagggttgacatgacgtgcaccgtccag
agatggacgtgccacaaggacggctgcacaggtgctgcatggctgctgacgctcgtccgtgaggtgtgggttaagtcccgaacg
agcgaaccctgtcgtcagttaccatgtcgtcagagactgccaaaactaggtggaggaaggcggggatgacgtcaagtcagcagggc
ccttacatcctgggtacacacacgctacaatggccgagacaaagtagcgaaccgcaaggcgcgagcgaatctgagcaaacccggct
tcagttcggattcagggctgcaaccgcctcgtcatgaagtcggagttgctagtaaccgtgagtcagcatatcacggtaatacgttcccgggt
ctgtacacaccgccgt

>ems409

ccgtgtaggggtcgggagggggaagtggaattcccgggtgtagcggtaaatgcgtagatcgggaggaacacctcggcgaaggcg
gcttctggaccgtaaccgacgctgagggcgcgaaagctaggggagcaaacaggattagatccccggtagtcttagccctaaacgttgag
cactgggtgtagccctcagggggtgtgccgaagctaaccgattaagtctccgcctggggagtacggcgcgaaggctgaaactcaaag
gaattgacggggcccgcacaagcgggtggagcatgtggttaattcagcgaacgcgaagaaccttacctggactcgaatctagttgac
cgtctcagagatgagacttcccgaaggacggctggataggtgctgcatggctgctgacgctcgtgctgtgagatgtgggttaagtcc
cgcaacgagcgaaccctcacccttggctcgtcccattcgtcgggaaagcaccgaaagggaaccgctcggataacggggaggaa
gggtggggatgacgtcaagtcctcagttcggccttatgccagggtacacacgtgctacaatggccggtacagagcgtcgaaacccgcgag
ggggagccaatcgcaaaaaccggtctcagttcggattgcaactcactgcatgaagctggaatcgtagtaatcgcgtatcag
aacgatgcgggtgaatacgttcccggcctgtacacaccgccgt

>ems411

gtagagatccggaagaacaccgggtggcgaaggcggcagctgggggtgtgtgtgacgctgatgcgcgacagcgtggggagcaaacagg
attagatccccggtagtccacgccgtaaacgatggcactaggtgccgggggagcgaccctcgggtgccgtcgtacacgattaagt

gccccgcctggggagctacggccgcaaggctgaaactcaaaggaattgacggggggccgcacaagcggaggagcatgtggtttaattcga
cgcaacgcgaagaacctacctaaggcttgacatgcacgtgattctctccagagatggaggggtccctcggggcactgacacaggtgctgc
atggctgtcgtcagctcgtgctgagatgttgggtaagtcccgaacgagcgaaccttgccccctgttgccagcggcggcgtaagcc
ggccggggactctggggggactgccggtgccaaaccggaggaaggtggggacgacgtcaagtcacatggtccttacgtctagggctac
acacgtgctacaatggacagtagagggctgcgaggtgcaatgccaagccaatcccaaaagctgtcctcagtgccgattgtcgcctgc
aaccggcggcatgaagctggaatcgtagtaatcgtggatcagcatgccacggtgaatacgttccccggccttgtaacacaccgccgt
>ems136

agcggcgcactggaccacaactgacgctgaggcgcgaaagctaggggagcaaacaggattagataccctggtagtcctagccctaaac
gatgattgcttgggtgtagcaggtaccaatcctgccgtgccgaagtaaacgcgataagcaatccgcctggggagtaggtcgcaaggctg
aaactcaaaggaattgacggggggccgcacaagcgggtggagcatgtggtcaattcagcgaacgcgaagaacctacctgggctcgaa
gcgagtgaggaccgggggtgaaacatcccttttcgcaagaactgctgcggaggtgctgcatggctgtcgtcagctcgtgctgtagatgtg
ggtaagtcccgaacgagcgaacccttatccctagttgccatcgggtcaagcctggcactctagtgaaccgccccggataacgggga
ggaaggtggggatgactcaagtcctcatggcctttatgtccagggtacacacgtgctacaatggtcgggtacaaccgctcgaacccgt
gagggggagctaategaaaaagccggcctcagttcggttggagctgcaacccgactccatgaagctggaatcgtagtaatecgcgga
tcagaacggcgggtgaatacgttccccggccttgtaacacaccgccgt
>ems137

gcgactggaccataactgacgctgaggcgcgaaagctaggggagcaaacaggattagataccctggtagtcctagccctaaacgatgat
tcttgggtgtagcaggtagccaatcctgctgtgccgcagctaacgcgataagcaatccgcctggggagtaggtcgcaaggctgaaactca
aaggaattgacggggggccgcacaagcgggtggagcatgtggtcaattcagcgaacgcgaagaacctacctgggctcgaagcgcagt
ggaccggggtagaaatatcccttcccctaaaaggctgctgcggaggtgctgcatggctgtcgtcagctcgtgctgtagatgttgggta
gtcccgaacgagcgaaccttatccctagttgccatcgggtcaagccgggactctagtgaactgcctcggataacgaggaggaagg
tggggatgactcaagtcctcatggcctttatgtccagggtacacacgtgctacaatggtcgggtacaaccgacgcaaacccgtgagggg
gagctaategaaaaagccggcctcagttcggttggagctgcaactcactccatgaggctggaatcgtagtaatecgcgatcagaac
gccgctggaatacgttccccggccttgtaacacaccgccgt
>ems138

gaaagctaggggagcaaacaggattagataccctggtagtcctaaacgatgtaacagtaagtccgcctggggagtagggaagactaaaact
gcagtaacgcagtaagtgtcccgcctggggagtaggccgcaaggtgaaactcaaaggaattgacggggggccgcacaagcgggtgg
agcatgtggttaattcagcaacgcgaagaacctacctaactcttgacatcctgcgaacttggtgagaggtgaggtgccgaaaggaac
gcagagacaggtgccgcatggctgtcgtcagctcgtgtgtagatgttgggtaagtccgtaacgagcgaaccttgctcttagttgcca
tcatttagttgggactctaaggagactgccgggtgaggaaccggaggaaggggggacgacgtcaagtcacatggtcctttatgagtagg
gctacacacgtgctacaatggggcgtacagagggtagccaaccgcgagggggagccaatctctaaagcgtctcgtagtcggattgga
gtctgcaactcactccatgaagtcggaatcgtagtaatecgcaatcagcattgtcgcggtgaatacgttccccggccttgtaacacaccgc
cgt
>ems1

gaaagctaggggagcaaacaggattagataccctggtagtcctaaacgatgtaacagtaagtccgcctggggagtagggaagactaaaact
caaaggaattgacggggggccgcacaagcgggtgcatgtggttaattcgaagcaacgcgaagaacctacctgggttgacatccctcg
accgacgtagagatacgtttttgtagcaataaacggggagacaggtgctgcatggctgtcgtcagctcgtgtggtgacatgttgggtaag
tcccgaacgagcgaaccttgctcttagttgccagcatttagttggctcttaaggagactgccgtggttaaacaggaggaaggtgggg
atgacgtcaagtcctcatggcctttatccaggggcgaacacacgtgatacaatggagggtagagaggaagccaaagcgaagctggagc
caatcctaaaaagcccttctcagttcggttggagctgcaactcactccatgaagttggaatcgtagtaatecgcgatcagcacggcgc
ggtgaatacgttccccggccttgtaacacaccgccgt
>ems8

gcgaaggc gatcacctgggggtgtctgacgctgatgagcgaagctaggggagcaaacgggattagatacccccggtagtcctagccgtt
aacgatggatactaggtgatgtggtatcactccatgtgccgtatgtgacacaataagtatcccgcctggggagtagggccgaaggtta
aaactcaaaggaattgacggggggccgcacaagcgggtggagcatgtggttaattcagcgaacgcgaagaacctacctgggcttgaca
tgtcggaaatcctatgaaagtaggctccggggcaaccggccgaacacaggtgctgcatggctgtcgtcagctcgtgccgtgaggtgt

gggtaagtcccgaacgagcgaaccctgtcgtgttgggtccacaagatctctcacagagaccgccggtgagaagccggagggaag
gtggggatgacgtcaagtcagcatggcccttacgtccagggctacacacgtgctacaatgcgcgcgacaacgagctgcgaaccgcgag
ggcaagccaatctcataaacgcgcgctcagttcggattgaaggctgcaactcgcctcatgaagtcggaatcgctagtaaacgcagatcag
ctacgctgcgtgaatacgttcccggcctgtacacaccgccgt

>ems3

gggccttgacagactgtgagacttgagggcctcagaggacatggaactccaggtgtagtgggaaatgcgtagatatctggaagaacacc
aaaggcgaaggcagtgctctgggaggtacgtgacgctcaggtgcgaaagccaggggagcgaacgggattagataccccggtagtctg
gccgtaaacgatgaccactagatgtcgggggtatcgaaccctgcgtgctggagtaacgcgctaagtgtccgctggggagtacggtc
gcaagataaaaactcaaaggaattgacggggaccgcacaagcagcggagcgtgtggttaattcagtgcaacgcgaagaacctacca
gggcttgacatggcattgcaccaccgagagatcgggtgctcctcggggcgggtgccaccgggtgctgcatggctgctcagctcgtgctg
agatgttgggtaagtcccgaacgagcgaacccctatcgtcggttggagatgtccggcgagactgccacgatcaacgtggagggaaggt
ggggatgacgtcaagtcagcacgcccgtacgtcctggggcagacacacgctacaatggccggtacaatgggacgccagcccgaag
ggggagccaatccagcgaaaaccggtcagttcggattggaggtgcaaccgcctcatgaagctggagttgctaagtaaccgtgggtc
agcacaccaggtgaatacgttcccgggtctgtacacaccgccgt

>ems5

gaaagcaggggagcaacgggattagataccccgtagtcccctaacgatgaacagtaagtggccgcctggggagtacaagtaaaactc
aaaggaattgacggggcccgcacaagcgggtggaacatgttgttaattcagcgaacgcgaagaaccttacctgggctcgaagtcca
gtgaccagcgggtgaaacatcgcctcccgaaggacactgtagaggtgctgcatggctgctcagctcgtgccgtgaggtgttgggt
taagtcccgaacgagcgaacccctactcttagtcgccaccgcgaagggggaactctaagaggactgctccggataacggagaggaag
gtgggatgacgtcaagtcagcacggcctttatgtccagggctacacacgtgtacaatgcgcggtacaaccgctcgaacccgtgagg
gggagctaatacggaaaaaactgcgctcagttcggattgcagctgcaactcactgcatgaagccggaatcgctagtaatggcgtatcaga
acgacccgtgaatacgttcccggcctgtacacaccgccgt

>ems65

tgacgtcagaggcgacagctaggggagcgaacgggattagataccccgtagtccctagccgtaaacgatgaccacgagatgtccggg
ggatcgacccttgggtgctgaagcaaacgcagtcagtggtccgctggggagtacggctcgaagattaaaactcaaaggaattgacgg
ggaccgcacaagcagcggagcgtgtggttaattcagtgcaacgcgaagcacttaccagggttgacatgctcttgaccgcttgaga
caagcgtcctctcggaggcgggagcacaggtgctgcatggctgctcagctcgtgccgtgaggtgtgggtaagtcccgaacgagc
gcaaccctcattgctagttgaacatctctagcagactgccacgatcaacgtggaggaaggtggggatgacgtcaagtcagcacggcctt
acatcctgggcgacacacagctacaatggtcagacaacgagttgccaaccgcgagggggagctaatacacaactcgaatcag
tcagattgacggctgcaaccgcctgcatgatggtggagttgctaaatcagtgaaatcgttcccggcctgtacacaccgccgt

>ems91

aaggcgtacacctgggggttctgacgctgatgagcggaaagctaggggagcaaacgggattagataccccgtagtccctagccgtaaac
gatggatactaggtgatgtggtatcgactccatgtgccgtagttaacacaataagtatccgcctggggagtacggccacaaggttaaaa
ctcaaaggaattgacggggggcccgcacaagcgggtggagcatgtggttaattcagcgaacgcgaagaaccttacctgggcttgacatgt
gcggaatcctatgaaagtaggctccggggctacccggccgcaacacaggtgctgcatggctgctcagctcgtccgtgaggtgttgg
gttaagtcccgaacgagcgaacccctgtcgtgttgggtccacaagatctctcacagagaccgccggtgagaagccggaggaaggtg
gggatgacgtcaagtcagcatggccttacgtccagggctacacacgtgctacaatgcgcgcgacaacgagctcgaacccgcgaggg
caagccaatctcataaacgcgcgctcagttcggattgaaggctgcaactcgcctcatgaagtcggaatcgctataaactcttcaagta
cgttcccggcctgtacacaccgccgt

>ems131

tggaattcccgggtgtagcgggtgaaatgcgtagatcgggaggaacacctgtggcgaagcggctcactggaccacaactgacgctgag
gcgcgaaaagctaggggagcaaacaggattagataccctgtagtccctagccctaacgatggatgcttgggtgtagcaggtaccaatcctg
ctgtgccggagctaacgcatccgcctggggagtacggctcgaaggtgaaactcaaaggaattgacggggggcccgcacaag
cggtggagcatgtggttaatacagcgaacgcgaagaaccttaccaggcttgacatccctgaaagccctagaaatagggtcttcatcgc
aagatgacagggtgacaggtgctgcatggctatcgtcagctcgtgccgtgaggtgttgggtaagtcccgaacgagcgaacccctcgtt
cgtgttggctcgaagatctctcaggggaccgccggtgagaagccggagggaaggtggggatgacgtcaagtcagcatggccttac

gtctggggctacacacgtgctacaatgggtgcgacaacgagttccaagctcgcgagggtaaggcaatctcttaaacgcacacctcagttcgg
attgaaggctgcaactcgccttcataagtcggaatcgctatcaactcagtgaaacgttcccgggccttgacacaccgccgt
>emp62

cggatactgacgctgagacgcgaaagcgtgggtagcaaacaggattagatacctggtagtccacgccgtaaacgatgagaactaggtgt
cgtgggtgtgacccccgcggtgccgtagtaaacgcattaagtctccgcctgggaagtacggccgcaaggctaaaactcaaaggaattga
cggggggcccgcacaagcggaggcatgtggttaattcgacgcaacgcgcagaaccttacctggtcttgacatcctcggaaatctgcaga
gatgacggagtgctcgcaagagaaccgagagacaggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggttaagtcccgcaa
cgagcgaaccctcgcctttagtccatgtaaaagtggatctctagagggactgccggtgtaaacggaggaaggtggggatgacgtca
agtctcatggcctttatgaccagggtacacacgtgctacaatggccggtacaaagcgtgccaactcgcgagagggagccaatcgcata
aaaaccggctcagttcagattggagctgcaactcgaaccatgaagcgggaatcgtagtaatcgtagatcagcagcgtacgggtgaatac
gttcccgggccttgacacaccgccgt

>emp203

cgaagctggggatcaaacaggattagatacctggtagtccacgctgtaaacgatgagaactagacgttgggaagatctgcctttcagttg
cgtagtaacgcgctaagtctccgcctggggagtaggccgcaaggtgaaactcaaaggaattgacggggggcccgcacaagcgggtgg
agtatgtggttcaattcgatgcaacgcgaagaaccttacctggtcttgacatcctcggaaacctgcagaaatcggggggtgccttcgggaac
cgagagacaggtgctgcatggctgtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgaaccctcgtcctagtggcc
agcattcagttgggactctagggagactgccggtgacaaaccggaggaaggtggggatgacgtcaagtcacatggcccttatgaccag
ggctacacagtactacaatggtcgggtacaaaggggtgccaagccgcgagggcgagccaatcccagaaaaccggctgtagtccggattg
gagctgcaactcactccatgaagtcggaatcgtagtaatcgcgatcagcaatgccggtgaatacgttcccgggccttgacacac
cgccccgt

>emp204

cgaagcggggatcaaacaggattagatacctggtagtccacgctgtaaacgatgagaactagacgttgggaagatctgcctttcagttg
ttagtaacgcgctaagtctccgcctggggagtaggccgcaaggtgaaactcaaaggaattgacggggggcccgcacaagcgggtgga
gtatgtggttcaattcgatgcaacgcgaagaaccttacctggtcttgacatcctcggaaacctgcagaaatcggggggtgccttcgggaacc
gagagacaggtgctgcatggctgtcagctcgtcgtgagatgttgggttaagtcccgaacgagcgaaccctcgtcctagtggcca
gcattcagttgggactctagggagactgccggtgacaaaccggaggaaggtggggatgacgtcaagtcacatggcccttatgaccagg
gctacacagtactacaatggtcgggtacaaaggggtgccaagccgcgagggcgagccaatcccagaaaaccggctgtagtccggattgg
agtctgcaactcactccatgaagtcggaatcgtagtaatcgcgatcagcaatgccggtgaatacgttcccgggccttgacacacc
gccccgt

>emp402

gacactgaggcacgaaagctaggggagcgaaacagattagatacctgtgtagtccatagccgtaaacgatgctcgttagtcttgccgca
aggtcgggggtgtaagctaacgcgtaagcagccgcctggggagtagccgcaaggtgaaactcaaaggaattgacgggggacccg
cacaacctgtggagcgtgtggttcaattcgagacgaaacgaaggacctcaccaggggttgacatgctaccgtacattcttcagaaatgaga
ggaaatccgcaaggatgtagcacaggtgctgcatggctgtcagctcgtcgtgaggtgttcccttaagtggggaacgagcgcgaa
ccccatccaatgttacatattcattgaaacagtcctgttttacaaggaaggaagcgggtagacgtcaagtcagcacggcccttatacc
ctgggccacacactacaatggcgaaaaacaatgagctgcaagctagcaatagtaagctaactctgaaattcgcctcagttcagattgg
ggtctgcaaccgacccatgaagccggaataggtagtaatcgcgatcagcaagccgcggtgaatacgttctcgggtcgtgtacacacc
gccccgt

>emp404

cgaagctaggggagcaaacaggattagatacctggtagtccatagccgtaaacgatgagcacttgggtggcccttacggggccgtgc
cggagctaacgcgtaagtgtccgcctggggagtagcgtcgaaggtgaaactcaaaggaattgacggggggcccgcacaagcgggtg
gagcatgtggttcaattcgacgcaacgcgaagaaccttacctggactcgaatgcagttgaccgtctcagagatgagatttcccgaaggg
acggctgtataggtgctgcatggctgtcagctcgtcgtgagatgttgggttaagtcccgaacgagtgcaacctcactcttagttgca
cggccgcaagggtgacccctaaggaaaccgcccgcataagggggaggaaggtggggatgacgtcaagtcctcatggcctttatgcc
agggctacacacgtgctacaatggccggtacagagcgtcgcgaactcgcgagagggagccaatcgtaaaaaccggctcagttcggatt

gcagtctgcaactcgactgcatgaagctggaatcgctagtaatcgctgtcagaacgacgcggtgaatacgttctcgggccttgtaacacacc
gcccgt

>emp405

ggttcctgacgctgaggtgcgaaagctaggggagcgaacgggattagataccccggtagtcttagccgtaaacgatgaccactagggcgt
gtggggtatcgaccccctactgcccactctaacgaagtaagtgtccgcctggggagtagcggtcgcaagattaaaactcaaaggaattga
cggggaccgcacaagcagcggagcgtgtggttaattcgaatgcaacgcgcagaacctaccagggttgacatggctctgatcgggcga
gaaatcgcttccttcgggacagagtcacaggtgctgcatggtgtcgtcagctcgtccgtgaggtgtgggtaagtcccgaacgagc
gcaaccctgtcgtcgggttacgggtgtccgacgagactgccgcgatacaactggaggaaggtggggatgacgtcaagtgcacggccc
ttacatcctgggcgacacacacgctacaatggctgggacaagagcggccaacctgcgaaggggagctaattcacaaccggctca
gttcagattagaggctgcaaccgctctatgaagtgggagttgctagtaaccgtaggtcagcacactacgggtaatacgttcccgggtctg
tacacaccgcccgt

>emp409

aggcagggtcctggctggtacctgacgctcgtggcgaaagctaggggagcgaacgggattagataccccggtagtcttagccgtaaac
gatgaccactaggtgtgcacgtatcgacccgttgcgtgccggaggcaccaccattaagtgtccgcctggggagtagcggtcgcaagatta
aaactcaaaggaattgacggggaccgcacaagcagcggagcgtgtggttaattcgaatgcaacgcgaagcacctaccagggttgaca
tgcgcgctgagcggcagagatggggggggccgcaagggaggcgcgcacaggtgctgcatggctgtcgtcagctcgtccgtgaggt
gttgggtaagtcccgaacgagcgaaccctgtcgggttgcgtatgatccggcgagactgccacgtgaaacgtggaggaaggtggggga
tgactcaagtcagcacggcctgatgtcctgggcgacacacacgctacaatggccgagacaacgggtcggcaaccgcgagggggga
gccaatcccagcaaagtcggcctcagttcagattgcaggtgcaaccgctcatgaaggaggagttgctagtaaccgtgggtcagcat
accacggtgaatacgttcccgggtctgtacacaccgcccgt

>emp412

aggcgaaggcagcctcctggaaagcatctgacgctcagggacgaaagcagggggagcgaacgggattagataccccggtagtctcgc
cgtaaacgatgaccactagatgtcgggggagtcgaccccctgcgtgtcgcagcttacgcagtaagtgtccgcctggggagtagcggtcgc
aagattaaaactcaaaggaattgacggggaccgcacaagcagcggagcgtgtggttaattcgaagcaacgcgaagcacctaccagg
gcttgacatgctcctgaccacggcgaaacgtcgtcctccttcggggcaggagcacaggtgctgcatggctgtcgtcagctcgtcgtga
gatgttgggtaagtcccgaacgagcgaaccctatggtcgggtactggtgtccggccagactgccgcgacaacgcggaggaaggt
ggggatgatgcaagtcagcacggcccttacgtcctgggcgacacacacgctacaatggccgggacaacgggacgccaaccgcgag
ggggagccaatcccaacaaccgggtcccagttcagattgcaggtgcaaccgctcatgaagtcggagttgctagtaaccgtgagtc
agcacatcacggtgaatgcgttcccgggtctgtacacaccgcccgt

>emp416

cgaaagcaggggagcgaacgggattagataccccggtagtctcgcgtaaacgatgaccactagatgtcgggggagtcgaccccctgc
gtgtcgcagcttacgcagtaagtgtccgcctggggagtagcggtcgcaagattaaaactcaaaggaattgacggggaccgcacaagca
gcggagcgtgtggttaattcgaagcaacgcgaagcacctaccagggttgacatgctcctgaccacggcgaaacgtcgtcctcctc
gggcaggagcacaggtgctgcatggtgtcgtcagctcgtcgtgagatgttgggtaagtcccgaacgagcgaaccctatggtc
gttactggtgtccggccagactgccgcgacaacgcggaggaaggtggggatgatgtcaagtcagcacggcccttacgtcctgggcgac
acacacgctacaatggccgggacaacgggacgccaaccgcgagggggagccaatcccaacaaccgggtcccagttcagattgcag
gctgcaaccgctgcatgaagtcggagttgctagtaaccgtgagtcagcacatcacggtgaatgcgttcccgggtcgtgtacacaccgccc
cgt

>emp417

cgaaagctaggggagcaaacaggattagataccctggtagtcctagccctaaacgatgagcactgggtgtggccccgtacggggccgtg
ccgaagctaacgcattaagtgtccgcctggggagtagcggtcgcaaggctgaaactcaaaggaattgacggggggccgcacaagcgg
ggagcatgtggtcaattcagcgaacgcgaagaaccttacctggactcgaatctagtcaccgtctcagagatgaggttcccgaag
ggacggctggataggtgctgcatggctgtcgtcagctcgtcgtgagatgttgggtaagtcccgaacgagcgaaccctaccgttag
tgtcgcggcgaagggagacccctaacggaaccgctcagataatggggaggaaggtggggatgacgtcaagtcctcatggcctttatg
cccagggctacacacgtgctacaatggccggtaacaagcgtcggcaaccgtaagggggagccaatcgtaaaaaccgggtcctcagttc

gattgcaggctgcaactgcctgcatgaagctggaatcgctagtaatcgctatcagaacgacgcggtgaatacgttcccggcctgtaca
caccgcccgt

>emp57

gcgaaagccaggggagcgaacgggattagataccccgtagtctggcctaacgatgggtactagatagtagactttcatgggttact
gtcgaagcaaaagtggtaagtaccccctggggagtatgctcgaaggctgaaactcaaaggaattgacgggggctcacacaagcggg
ggagcatgtggctcaattcaggctacgcgaagaaccttacctgggcttgacatgtgcgaaagcgttggggagtagcgggtggaaacac
tttccaacggatccagtcggaaccaatacaggtgctgcatggctgtcgtcagctcgtcgtgagatgtcgggttaagtcataacga
gcgaaaccttgttccagtggcagcgggtcatgccggggactctggaaagactgccgggtttaaaccggaggaaggtggggatgacgt
caagtctcatggcctttatgccagggtgcacacgtgctacaatgccgtgaacaaaggacgcaaacctgcgagggggagcaaatcc
caaaaacacggcccagttcagatcgaaggctgcaactcgccttcgtgaagttggaatcgctagtaatcgaggtcagcaaacactgcgggt
aatgttctctgagcctgtacacaccgcccgt

>emp420

cgaagcgtggggagcaaacaggattagataccctgtagtccacgccgtaaacgatgtgcgctagacgttgggggacctagtctcccgg
tgtcgagctaacgcgtaagcaccgctggggagtacggccgaagggttaaaactcaaaggaattgacggggcccgcacaagcgg
gtggagcatgtggtcaattcagcgaacgcgaagaaccttaccagccctgacatgtcccgtatggggccgagagatcgcttcgtagtt
cggctggcgggaacacaggtgctgcatggcgtcgtcacctcgtgctgtagatgttgggttaagtcggcaacgagcgaaccctacc
ttcagttccatcgggttaagccgggactctgaaggaactgccgggtgacaagccggaggaaggcggggatgacgtcaagtctcatgg
cctttatgggctgggctacacacgtgcttcaatggcggcgacagagggaggaaggcgggggacccctgcgccaatctcgaaaagccgtct
cagttcggattgttctctgcaactcgagagcatgaagctggaatcgctagtaatcgcgatcagcacgccgcggtgaatacgttcccgggc
ctagtacacaccgcccgt

>emp437

aaatgcgtagatatcaggaggaacaccggcggcgaagcggccctctggatcgttctgacgctgagacgcgaaagcgtggggagcaa
acaggattagataccctgtagtccacgctgtaaacctatgggtactaggttccgggtattgaccccctgagtccgcagctaacgcattaa
gtaccccgcctggggactacggccgaagggttaaaactcaaaggaattgacggggcccgcacaagcgggtggagcatgctgtcaattc
gacgcaacgcgaagaaccttacctgggctagacaacggcggacagcctgagaaatcaggtcttccctcggggactgccggttcaggtgt
tgcattggctgtcgtcagctcgtcgtgagatgttgggttaagtcggcaacgagcgaaccctgcccttagttgccaccggataatgccg
agcactctatggggactgcaggggtcaaccgtgagggggggggacgacgtcaagtcacatggcccttatgtccagggtacaagcgt
gctacaatggacggtacaaaggggtcgaactcgcgagagccagccaatcccagaaaaccgcccctcagctcggatcggagctgcaac
cgcatgaagtggaatcgctagtaatcgtgaatacgttcccgggctgtacacaccgcccgt

>emp445

ctcagggggcgaagctaggggagcgaacggggttagataccccgtagtcttagctgtaaacgatgaccactaggtgtcggggatcgc
accccctgcgtgccgcagtaaacgcaataagtgtccgcctggggagtacggctcgaagattaaaactcaaaggaattgacggggacc
gcacaagcagcggagcgtgtggttaattcagcgcgacgcgagaaccttaccaggggttgacatgccggtgcaccggcccgagaaggt
cgtcccgcaggacgccggcacaggtgctgcatggctgtcgtcagctcgtgccgtgaggtgttgggttaagtcggcaacgagcgcga
accccctcgtcgttcatgtatccgacgagactgcctc gatcaacggggaggaaggtggggatgacgtcaagtcagcacggccttac
atcctgggcgacacacacgctacgatggtcgcggacaatggggagccaaccgcgagggggagccaatcccagcaaaagcggcctca
gttcggattgcaggctgcaaccgctcgtatgaagtcggagttgctagtaaccgtgggtcagcacaccaggtgaatacgttcccgggtctt
gtacacaccgcccgt

>emp446

cgaagctaggggagcaacgggttagataccccgtagtcttagctgtaaacgatgaccactaggtgtcggggatcgcaccccctgcgtg
ccgcagctaacgcaataagtgtccgcctggggagtacggctcgaagattaaaactcaaaggaattgacggggaccgcacaagcagc
ggagcgtgtggttaattcagcgcgacgcgagaaccttaccaggggttgacatgccggtgcaccggcccgagaaggtcgtcccgcaa
gggacgccggcacaggtgctgcatggctgtcgtcagctcgtgccgtgaggtgttgggttaagtcggcaacgagcgaacccccatcgtc
ggttcatgtatccgacgagactgcctc gatcaacggggaggaaggtggggatgacgtcaagtcagcacggccttcatcctggcgac
acacacgctacaatggtcgcggacaatggggagccaaccgcgagggggagccaatcccagcaaaaggcggcctcagttcggattgca

ggctgcaacccgcctgcatgaagtcggagttgctagtaaccgtgggtcagcacaccacggtgaatacgttcccgggtctgtacacaccg
ccgt

>emp447

cgaagctaggggagcgaacgggattagataccccggtagtctagccgtaaacgatgaccactaggtgtcggggatcgacccccg
cgtgccgtagctaacgcaataagtggccgctggggagtagcggtcgcaagattaaaactcaaaggaattgacggggacccgcacaagc
agcggagcgtgtggttaattcgacgcgacgcgcagaacctaccagggttgacatgctgggtgaccagccagagatggtgtaccgca
agggacgccggcacaggtgctgcatggtgtcgtcagctcgtgccgtgaggtgttgggtaagtcccgaacgagcgaacccctgtcgt
cggttggagtctccgacgagactgcctgatcaacggggaggaaggcggggatgacgtcaagtcagcacggcccttacctggtggc
acacacacgctacaatgggcgcgacaacgggaagccaaccgcgagggggagccgatccagcaaggcgtcctcagttcggtggt
aggctgcaacccgcctgctgaagtcggagttgctagtaaccgtgggtcagcactacggtgaatacgttcccgggtctgtacacaccg
ccgt

>emp448

cgaagctaggggagcaaggaattagataccccggtagtctagccgtaaacgatgaccaccaggtgtgggaggtatcgacccttccgc
gccccggcgttaacacactaagtggctcgcctggggagtagcggtcgcaagattaaaactcaaaggaattgacggggacccgcacaagcag
cggagcgtgtggttaattcgatcgacgcggaagaacctaccagggcttgacatgacttgacgccctctgagaggagggtcccttggg
gccccgtcacaggtgctgcatggctgtcgtcagctcgtgccgtgaggtgttgggtaagtcccgaacgagcgaacccctgttgcggct
ggatgtgtccgacgagactgccacgatcaacgtggaggaaggtggggatgacgtcaagtcagcacggccctgacgtcctgggcgacac
acacgctacaatggtcgggacaacgggtcgccaaccgcgagggggagccaatcccagcaaacccgatctcagttcggtggcaggt
gcaacccgcctgctgaagccggagttgctagtaaccgtgggtcagcacaccacggtgaatacgttcccgggtctgtacacaccgccc
>emp443

gcagctaacgccaataagtggccgctggggagtagcggtcgcaagattaaaactcaaaggaattgacggggacccgcacaagcagcg
gagcgtgtggttaattcgacgcgacgcgcagaacctaccagggttgacatgccggtgaccggcccgagaaggtcgtgccgcaag
ggacgccggcacaggtgctgcatggctgtcgtcagctcgtgccgtgaggtgttgggtaagtcccgaacgagcgaacccccatcgtc
gttacatgtatccgacgagactgcctgatcaacggggaggaaggtggggatgacgtcaagtcagcacggcccttacctggtggcgaca
cacacgctacaatggtcgcggacaatggggagccaaccgcgagggggagccaatcccagcaaacggcggcctcagttcggtgag
gctgcaacccgcctgcatgaagtcggagttgctagtaaccgtgggtcagcacaccacggtgaaggcgttcccgggtctgtacacaccg
ccgt

>emp430

ctggctggcgactgacgctgagggggcgaagccaggggagcgaacgggattagataccccggtagtctggctgtaaacgatgaccac
caggtgtcggcgatcgaccgctgctgcccggcgttaacacaatcagttgctcgcctggggagtagcggtcgcaagattaaaactcaa
ggaattgacggggacccgcacaagcagcggagcgtgtggttaattcgatcgacgcggaagaacctaccagggcttgacatgcctctgc
accgcccagagagggggcgtgcccgcaaggagcggagggcacaggtgctgcatggctgtcgtcagctcgtgccgtgaggtgttgggta
gtcccgaacgagcgaacccctgtggccggtgaaactcctccggcagactgccacgatcaacgtggaggaaggtggggatgacgtc
aagtcagcacggcccttactcctgggcgacacacacgctacaatggccgagacaaagggggcgaacccgcgagggggagccaat
ccggcaaacctggtctcagttcagattgcaggctgcaacccgcctgcatgaagtcggagttgctagtaaccgtgggtcagcacaccacggt
gaatacgttcccgggtctgtacacaccgcccgt

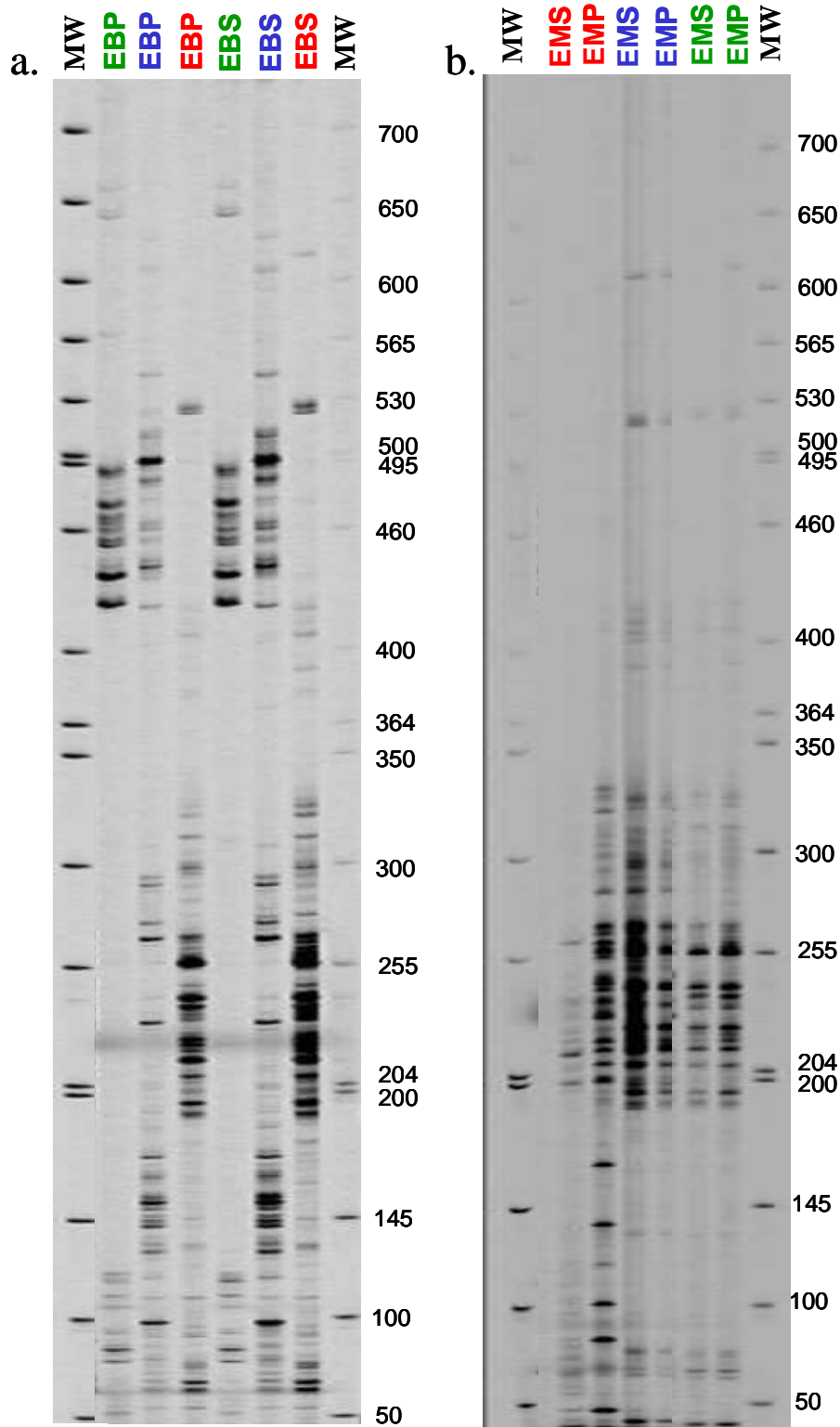
>emp206

ggggatcaaacaggattagataccctggtagtcacgctgtaaacgatgagaactagacgttggaggtatctgcctttcagttcgtgtagta
acgcgctaagttctccgctggggagtagcggccgcaaggtgaaactcaaaggaattgacgggggcccgcacaagcgtggagtagtg
gttcaattcgatgcaacgcggaagaaccttacctggtcttgacatcctcggaaacctgcagaaatcggggggtgcttcgggaaccgagaga
caggtgctgcatggctgtcgtcagctcgtgtcgtgagatgttgggtaagtcccgaacgagcgaacccctgctcctagttgccagattca
gttgggactctaggagactgccggtgacaaaccggaggaaggtggggatgacgtcaagtcacatggtcccttaccaggggtaca
cacgtactacaatggtcggtaaaaggttccaagccgcgagggcggagccaatcccagaaaaccggtcgtagtcggattggagctg
caactcgactccatgaagtcggaatcgctagtaaccggtgacgcaatcccggtgaatacgttcccggcctgtacacaccgcccgt
>emp414

gtcctagccctaaacgatcaggacttgggggtgccacccgttcgggtgtcgtcccggagctaacgcgttaagtctgcacctggggagtag
gtcgcgaagactgaaactcaaaggaattgacggggggcccgcacaagcgggtggaacatgtggtcaattcgacgctacgcgaggaaccttac
ctgggctcgaatgcttctgaccagctgtagaaatacggccttcccgaagggacaggagtataggcgtgcatggctgtcgtcagctcgt
gccgtgaggtgtgggtaagtcccgaacgagcgaacccctgcacgtagttgccactcgaagagggaactctacgtggactgctccg
gataacggagaggaaggtggggatgacgtcaagtccgcatggcctttagtccagggtacacacgtgttacaatgcagggcacaaccg
ttgccaaccgcgagggggagctaateggaaaaactctgctcagttcggattgcagtctgcaactcgactgcatgaagccggaatcgta
gtaatggcgtatcagatcgacgccgtgaatacgttcccggccttgtacacaccgccgt
>emp440

ggggctcgtgcccgtagctaacgcgataagtctgcgcctggggagtaggtcacaaggctgaaactcaaaggaattgacggggggccc
cacaagcgggtggaacatgtggtcaattcgacgcaacgcgaagaaccttacctgggcttgaagtgtactggaccattcctggaaacaggga
cttcccgaaggaccggtgtagaggtgctgcatggctgtcgtcagctcgtgccgtgaggtgtgggtaagtcccgaacgagcgaac
ccttactttagttgectcccgaaggggggaactctacaaggactgctccggataacggagaggaaggtggggatgaggtcaagtcagca
tggcctttagtccagggccacacacgtttacaatgcgcgtacaaagcgtcgcgaaccgcgagggggagctaategcaaaaaagcg
cgctcagttcggattgcagtctgcaactcgactgcatgaagccggaatcgctagtaatggcgcacagaacggcgccgtgaatacgtcac
ggccttgtacacaccgccgt

Appendix VI



T-RFLP fingerprints of bacterial communities from Bosque del Pueblo's soil samples. Samples in red color identified samples digested by *HaeIII*, samples in blue color identified samples digested by *MspI*, and samples in green color were digested by *RsaI*. Panel A shows T-RFLP profiles for EBS and EBP soil samples; and Panel B shows T-RFLP profiles for EMS and EMP soil samples.