Deanship of Graduate Studies

Al-Quds University

Isolation and Identification of Bacteria from Stored Soil Samples

Sajeda Saleh Yousef Nojoum

M.Sc. Thesis

Jerusalem-Palestine

1442/2020

Isolation and Identification of Bacteria from Stored Soil Samples

Prepared By: Sajeda Saleh Yousef Nojoum

BSc. Life Sciences/ 2013

Al-Quds University-Palestine

Supervisor: Prof. Sameer A. Barghouthi

A thesis submitted in partial fulfillment of requirement for the degree of Master in Medical Laboratory Sciences / Diagnostic microbiology and Immunology Track/ Deanship of Graduate Studies / Al-Quds University.

1442/2020

Al – Quds University

Deanship of Graduate Studies

Medical Laboratory Science

جامعة القدس
AL-QUDS UNIVERSITY

Thesis Approval

Isolation and identification of bacteria from stored soil samples

Prepared By: Sajeda Saleh Yousef Nojoum

Reg. No. 21710034

Supervisor: Prof. Sameer A. Barghouthi

Master thesis submitted and accepted, Date: 6 / 12 / 2020

Names and Signatures of Examining Committee Members are as follow:

1- Head of committee: Prof. Sameer A. Barghouthi Signature: ____________

2- Internal Examiner: Dr. Rania Abu Seir Signature: ____________

3- External Examiner: Dr. Fawzi Al-Razem Signature: ____________

Jerusalem-Palestine

1442/2020

Dedication

I dedicate this work to my Parents,

I dedicate this work to my Brothers my Sisters and my aunt,

I dedicate this work to my uncle Abu Al- Waqqas Dahdol,

I dedicate this work to all the teachers who teach me,

I dedicate this work to My friends.

Sajeda Saleh Yousef Nojoum

Declaration

I certify that this thesis submitted for the degree of master, is the result of my own research, except where otherwise acknowledged, and that this study (or any part of the same) has not been submitted for a higher degree to any other university or institution.

Sajeda Saleh Yousef Nojoum

14-12-2020

Acknowledgments

At first, I would like to thank Allah on what he granted me of useful knowledge, and I thank him for making me patient and able to face the obstacles to continue and get a master degree in Medical Laboratory Science.

I would like to give my gratitude and appreciation to my supervisor**,** Prof. Sameer A. Barghouthi, faculty of Health Professions, Department of Medical Laboratory Sciences, Al-Quds University, Jerusalem, Palestine. I thank him for his contentious directions and suggestions. His guidance helped me through the time of research and writing this thesis. Also, many thanks to Dr. Rania Abu sir for her help, support and continuous encouragement of this research. I would like to thank Al-Quds University and Medical Laboratory staff.

I am so grateful for the two most important people in my life, Dad and Mom whom facilitated my path and encouraged me with their love to continue my educational process, I am highly indebted to them. Also, to my brothers, sisters, aunt, and uncle for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

I am highly indebted to Miss. Suha Butma for her help in this study.

I would like to thank my best friends Hanna'a Abdel Qader, Aman Jaber and Insaf Abu-Odeh for their help and support in this study, they were always there with a word of encouragement and a listening ear.

Finally, I could not have done any of this without you. Thank you all.

Abstract

Soil microbial population is large and diverse, and it is affected by biotic factors such as soil type and climate changes. Interaction of various amounts of water, sunlight, nutrients, temperature and acidity are important determinants of the numbers and types of microorganisms in specific portion of soil. *Bacillus* specie are ubiquitous in soils and are characterized with their ability to form endospores within cells that provides high resistance to desiccation, heat and chemicals.

In this study, thirteen different types of stored soils were used to isolate and identify 16S rDNA of bacteria species that have survived for several years in plastic bags at room temperature using Universal Method based on utilizing the Golden Mixture7 (G7) which used from previous study.

Different types of bacterial genera/species in stored soil samples were identified as; *Lysinibacillus fusiformis* strain WS1-3, *Bacillus sp*. (in: Bacteria) strain CM-CNRG 602, *Peribacillus asahii* strain OM18, *Bacillus cereus* strain BCd16, *Bacillus cereus* strain BHU4, *Bacillus amyloliquefaciens* strain CS13, *Bacillus subtilis* strain UIS0380, *Lysinibacillus fusiformis* strain P-R2A48, *Bacillus cereus* strain NCIM2157, *Bacillus safensis* strain MUGA119 with sequence similarities $\geq$ 98%.

The presence of bacilli in stored soil samples and the rarity of other types of bacteria including Gram-negative bacteria indicated that spore formers bacteria may dominate since they resisted environmental stress and could persist for many years until suitable environment is available.

Depending on results of this study, a numbers of *Bacillus spp.* were identified in different samples of soil which were stored for several years. This indicated that most common type of bacteria that can live in stored soils for long period of time are *Bacillus Spp.* The absence of less tolerant species in stored soils for years under aerobic condition may have resulted in the death of these species which are usually present in fresh soil samples while retaining the endospore-forming species.

عزل وتحديد البكتيريا من عينات التربة المخزنة

إعداد: ساجدة صالح يوسف نجوم

إشراف: برفسوور سمير البرغوثي

ملخص

عدد الميكروبات في التربة كبير ومتنوع ويتأثر بالعوامل الحيوية مثل نوع التربة والتغيرات المناخية. يعتبر التفاعل بين كميات مختلفة من الماء وضوء الشمس والمغذيات ودرجة الحرارة والحموضة محددات مهمة لأعداد وأنواع الكائنات الحية الدقيقة في جزء معين من التربة. توجد العصيات في كل مكان في التربة وتتميز بقدرتها على تكوين الأبواغ الداخلية داخل الخلايا التي توفر مقاومة عالية للجفاف والحرارة والمواد الكيميائية..

في هذه الدراسة ، تم استخدام ثلاثة عشر نوعًا مختلفًا من التربة المخزنة لعزل وتحديد 16s rDNA لأنواع البكتيريا التي نجت لعدة سنوات في أكياس بلاستيكية في درجة حرارة الغرفة باستخدام الطريقة العالمية المعتمدة على استخدام الخليط الذهبي 7 (G7) الذي تم استحدامه من الدراسة السابقة (البرغوثي) 2011.

تم تحديد أنواع مختلفة من الأجناس / الأنواع البكتيرية في عينات التربة المخزنة على النحو التالي:

Lysinibacillus fusiformis strain WS1-3, *Bacillus sp*. (in: Bacteria) strain CM-CNRG 602, *Peribacillus asahii* strain OM18, *Bacillus cereus* strain BCd16, *Bacillus cereus* strain BHU4, *Bacillus amyloliquefaciens* strain CS13, *Bacillus subtilis* strain UIS0380, *Lysinibacillus fusiformis* strain P-R2A48, *Bacillus cereus* strain NCIM2157, *Bacillus safensis* strain MUGA119مع تشابه التسلسل(≤98%) .

يشير وجود العصيات في عينات التربة المخزنة وندرة الأنواع الأخرى من البكتيريا بما في ذلك البكتيريا سالبة الجرام إلى أن مسببات الجراثيم التي تكون أبواغ قد تكون شائعة لأنها قاومت الإجهاد البيئي ويمكن أن تستمر لسنوات عديدة حتى تتوفر البيئة المناسبة. بالإعتماد على نتائج هذه الدراسة ، تم التعرف على أعداد من *Bacillus spp.* في عينات مختلفة من التربة والتي تم تخزينها لعدة سنوات. يشير هذا إلى أن أكثر أنواع البكتيريا شيوعًا التي يمكن أن تعيش في التربة المخزنة لفترة طويلة من الزمن هي *Bacillus Spp.* قد يؤدي عدم وجود أنواع أقل تحملاً في التربة المخزنة لسنوات تحت ظروف هوائية إلى موت هذه الأنواع التي توجد عادةً في عينات التربة الطازجة بينما تبقى الأنواع المكونة للأبواغ.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ARISA	Automated Ribosomal Intergenic Spacer Analysis
BCAs	Biological Control Agents
BLAST	Basic Local Alignment Search Tool
DGGE	Denaturing Gradient Gel Electrophoresis
EDTA	Ethylenediaminetetraacetic acid
G7	Golden Mixture 7
MEGA-X	Multiple Evolutionary Genetics Analysis
NA	Nutrient Agar
NTM	Non-Tuberculosis Mycobacteria
PC R	Polymerase Chain Reaction
PGPR	Plant Growth Promoting Rhizobacteria
QUGP	Al-Quds University General Primer
RISA	Ribosomal Intergenic Spacer Analysis
SAF	Spacecraft Assembly Facility
SDS	Sodium Dodecyl Sulfate

TAE	Tris-Acetate-EDTA
TGGE	Temperature Gradient Gel Electrophoresis
USA	United State of America
UV	Universal Method

Chapter1: Introduction

1.1 Background

Soil related microbes play a significant role in soil properties and affect below-ground ecosystem processes. Microbes can influence soil quality through decomposition of organic matter, recycling of nutrients and biological control of parasites. The limited knowledge of soil-related bacteria and their classification through families and orders hinder our abilities to measure their activities (Stefanis et al, 2012).

Presence of microbes in soils is determined by environmental changes such as; physical and chemical degradation (erosion and compactness), chemical contaminations from industries lead to losses in microbial biodiversity. In natural ecosystem; microorganisms are found in high numbers in spite unknown thousands of microorganisms that have not been described yet. It is estimated that one gram of soil contains counts of bacteria as counted by fluorescent microscopy after staining with a fluorescent dye (Fakruddin and Mannan 2013) .

Bacteria have several uses for human including antibiotic, *Bacillus* species are most common soil bacteria related to their resistant endospore formation and production of vital antibiotics like bacitracin and polymxin, small group of microorganisms belonged to the genera *Penicillium, Streptomyces*, *Cephalosporium* and *Micomonospora* produce antibiotics that have been purified and used in medical practice (Sethi et al, 2013). Also, bacteria have been used as anticancer agent from hundred years back, it has been reported that bacteria can replicate and accumulate within plant tumors through different mechanism (Patyar et al, 2010).

Many molecular and biochemical methods were used in the past decades for isolation and identification of classical bacterial species from soil (Stefanis et al, 2012). Utilizing molecular methods has enhanced discovery of new microorganisms. Many molecular methods have been developed to study microbial diversity including; DNA-DNA and mRNA- DNA hybridization, DNA re-association, DNA Restriction Fragment Polymorphism, DNA cloning, DNA sequencing and other PCR-based methods such as, Temperature Gradient Gel Electrophoresis (TGGE), Denaturing Gradient Gel Electrophoresis (DGGE), Ribosomal Intergenic Spacer Analysis (RISA) and Automated Ribosomal Intergenic Spacer Analysis (ARISA).

In this study, we utilized 13 different soil samples which were collected from different areas of the world and stored for many years at room temperature separately in plastic bags in wooden drawers in the Microbiology Research Laboratory, Department of Medical Laboratory Sciences, Faculty of Health Professions, Al-Quds University, Jerusalem, (Palestine). These different soils were used for isolation and identification of bacteria species utilizing The Universal Method (UM) approach and also aerobic culture methods. This molecular technique depends on using a multiplex mixture known as Golden Mixture 7 (G7) of general primers to amplify the 16S rDNA of isolated bacteria obtained from 13 different soils.

1.2 Aims and Objectives

The aim of this study was to isolate and identify bacteria species that have survived for several years in stored batch soil samples. The specific objectives were:

- Identify most common dominant bacteria that survived storage periods in batch-soil samples.

- Explain the unexpected disappearance of certain bacteria that are usually found in fresh soil samples.

1.3 Hypothesis

Some bacteria can survive in stored soils for a long period of time. If samples of soil from different sources were stored in plastic bags for several years under aerobic condition at room temperature, only “spore-forming” bacteria that can survive under these conditions will remain viable in soil.

Chapter 2: Literature Review

In recent years, it was clear that most natural habitats on earth contain a great diversity of microorganisms, mostly prokaryotes (Archaea and Bacteria). These habitats can vary largely such as; lake water and soil (Madigan 2000). Complex interaction of different amounts of water, sunlight, nutrients, temperature and acidity are important determinants of the numbers and types of microorganisms in specific portion of soil (Baumgardner 2012).

Pathogens may naturally exist among soil microorganisms, or may contaminate soils through dumpsites, sewage, dead animals deposits and manure application. Human contact with such soil microorganisms may result in human disease, which can occur through ingestion of soil contaminated foods or water, ingestion of soil (geophagia) or directly when pathogen enter humans through inoculation into wounds or the respiratory tract through bio-aerosols (Baumgardner 2012).

2.1 Soil-related Infections

Soil-related bacterial and fungal infections of human are well documented. The growth of these microorganism is affected by soil characteristics and may have complex life cycle such as an intermediate host; amoebae or animal hosts. Infections may occur through direct ingestion or inoculation or inhalation. Most types of infection are caused by bacteria and fungi (Baumgardner 2012).

Most important survival factors of bacteria and other microorganisms in soil is their nature and adaption to the different soil conditions and composition, physical and chemical properties of soil and the presence of other competitors.

2.1.1. Soil-related bacterial pathogenesis:

Tetanus and botulism are classical cases of soil-borne bacterial pathogenesis /infections caused by soil bacteria (*Clostridium tetani* and *Clostridium botulinum* respectively) which may cause wound, skin, gastrointestinal and respiratory tract diseases. *Clostridium tetani* and *Clostridium botulinum* are gram positive anaerobic, spore-forming bacteria and produce the toxins that caused tetanus and botulism respectively (Baumgardner 2012). Classic food poisoning caused by *Clostridium perfringens* which is very common in soil, most common cases of gastrointestinal disease are caused by this bacterium that contaminate foods exposed to animal feces (Baumgardner 2012).

Anthrax is a disease caused by gram positive, spore-forming rod *Bacillus anthracis*, this type of bacterium lives in black soils which are rich in organic matter and calcium which in turn enhances spore viability. In the world, the number of human anthrax cases is 2000-20,000 annually (Baumgardner 2012), the disease is associated with sheep-wool processing.

Gastroenteritis caused by *Bacillus cereus* is another soil-related bacterial infection, in the United States of America (USA) it is responsible for 63,000 cases of gastroenteritis annually, *Bacillus cereus* found in nature; in decaying matter, in soil and water and may cause pulmonary infection that may result from inhalation of contaminated dust (Baumgardner 2012).

Listeria monocytogenes is a gram-positive, non-spore-forming bacterium, cause listeriosis and it is responsible for 1591 cases of gastroenteritis in normal human in USA annually. It is widely distributed in soil, sewage, silage, ground water and vegetation. *Listeria monocytogenes* can confront environmental stress such as salinity, pH changes, metal ions and low temperatures (Fenlon 1999; Baumgardner 2012).

Also, soil contains gram-negative enteric pathogens which may enter the soils after contamination by sewage, human or animal waste, the viability of these enteric pathogens is enhanced by soil moisture and adsorption to clay particles. *Salmonella* is one of these pathogens and can survive in soils for prolonged period (Baumgardner 2012; Jacobsen and Bech 2012). *Salmonella* species are facultative, anaerobic, rod-shaped bacteria belonged to family *Enterobacteriaceae,* and cause gastrointestinal disease (Cohen et al, 1987).

Escherichia coli O157: H7is an enteric gram-negative bacterium, belongs to family *Enterobacteriaceae,* it can be isolated from infected manured garden where a child has been reported to have contracted the infection. *E. coli* O157:H7 is capable of replication within protozoan *Acanthamoeba* in part of soil (Baumgardner 2012).

Many types of *Legionella* species cause pneumonia; it is a fastidious, gram-negative bacterium, usually found in biofilms in the environment, it can live in harmony with other microorganism and survive and multiply intracellularly inside free-living amoebae. *Legionella* species can persist in many types of soils such as *L. longbeachae* usually isolated from composts, soil products especially those exposed to heat and moisture (Baumgardner 2012).

Pseudomonas species are aerobic, Gram-negative bacteria and can cause infection to animal and human, *Pseudomonas* is considered one of the most opportunistic microorganisms that is responsible for drug-resistance nosocomial infections. Infections with *Pseudomonas* are usually distinguished by blue pus (Swe et al, 2018). The genus *Pseudomonas* is present in large numbers of all natural environments, it is the most significant bacterium that is associated with plant roots/rhizosphere (Egamberdiyeva 2005).

Some mycobacterium species are widespread in waters and soils ; these microorganisms are known as environmental non-tuberculosis mycobacteria (NTM) (Rahbar et al, 2010). Presently, NTM consist of more than 150 species, and are widespread in both natural and man-made environments. Infection with NTM can be acquired from environment through inhalation, ingestion and dermal contact, which lead to lymphadenitis, pulmonary and disseminated infections, and skin and soft tissue infections (Nishiuchi et al, 2017).

The endospores former bacteria belong to genera; *clostridia* which is strictly anaerobic *Clostridia* are ubiquitous present in soils and can live for long periods of time through maintain very low level of metabolic activity (Heyndrickx 2011). *Bacillus spp.* which are aerobic or facultative anaerobic bacteria. Endospores are created at the end of growth phase when the vegetative mother cells acting as sporangium. These endospores are prevalent in soils, resist heat, they are able to attach to processing equipment through their adhesive characteristics and have ability to germinate and grow under favorable conditions.

2.1.2. Soil-related fungal infections:

The probability of infection by soil fungi are determined by a number of factors including geographic, soil and environmental factors which in turn determine the presence of a particular fungus. Most fungal infections are acquired by inhalation from contaminated soil and near soil environment and cause primary-pulmonary disease (Baumgardner 2012). *Malassezia* belong to fungal order Malasseziales, are widespread and ecologically diverse. Two species of genus *Malassezia* are characterized as human skin associates which *M. globosa* and *M. restricta*. *M. restricta* may be particularly widespread and DNA sequences similar to these species have been detected in habitats such as deep-sea sediments, stony corals, hydrothermal vents and lobster larval guts (Amend 2014).

2.2 Soil microbiology and DNA

In 1980, Torsvik et al were able to extract pure DNA from soil to apply hybridization techniques. In 1990, they reported that zero to ten gram of soil included more than 4000 different genomes of bacteria (Insam 2001).

Traditional culturing techniques used for the identification of soil bacteria are very important and they are still in use today, but using these methods is inadequate and requires more sophisticated techniques which have been developed for the isolation of bacteria from complex microbial habitats (Joseph et al, 2003).

Today, several methods are used to study soil DNA or RNA. Molecular ecologists depend on 16S rDNA analysis as the main feature of prokaryotes. The 16S rDNA consists of a strand of ~1500-bp, it contains both conserved and variable regions, after extraction of DNA for amplification of specific target in the 16S gene, selected primers are used which in turn determine the targeted region to be amplified. After amplification by PCR, analyses are required which give sufficiently detailed insight into specific group (Insam 2001).

2.3 The Universal Method (UM)

In this study, we used the Universal Method (UM) for identification of bacterial genera or species in different soil samples, the principle of this method depends on using pure PCR product of 16S gene, after amplicon sequencing and alignment, the bacterium can be identified. This method does not require any knowledge of the nature of the investigated bacterium, the method utilizes a number of primer multiplexes such as 16S primers (Golden mixtures) against several bacterial isolates. The mixture named Golden mixture 7 (G7) has lend itself very efficiently in detecting tested bacteria. Golden mixtures were designed for practical reasons and to simplify the methods that use single primer pairs. G7 and G10 have shown excellent abilities in detecting bacterial DNA even when diluted 10,000 folds, this important dilution of samples allows getting rid of possible low level contamination with other bacteria (Barghouthi 2011).

Universal Method proved efficient in identification of genus, species, novel species or genera. It allows for detection of variations between species and can be used to detect bacteria in different samples such as cerebrospinal fluid, blood and manufactured samples (Barghouthi 2011; Barghouthi and Al Zughayyar 2012).

In this study, G7 mixture was applied as described previously (Barghouthi 2011) to identify bacterial DNA of isolated bacteria obtained by culture from different stored soils.

Chapter 3: Materials and Methods

3.1 Source of Samples

Thirteen different types of stored soils were used in this study for isolation of different bacterial species, the sources of soil were from different areas in the world which were collected few years back and were stored at room temperature in the research lab of Microbiology in Al-Quds University. The sources of soil were from areas near; Eiffel Tower and another near Louvre Museum, Paris, France (2012). A Basel metro sample, Switzerland (2012); Rhine River side, Switzerland (2012), Austria (2013); Hammamat Maein (2010), Jordan (2009); Sea of Marmara, Avcilar, Turkey (2009). Tiberias Lake, Palestine (2009), Battir village, Palestine (2010); Stanford University, San Francisco, Ca, USA (2009); Oslo, Norway (2010), the Naqab desert, Palestine (2011). Appendix 1 show the sources of soil samples.

3.1.1. Isolation of Soil Bacteria on aerobic nutrient agar (NA) plates:

About 10 mg of each stored soil sample (13 types) were aseptically distributed by sprinkling on the surface of nutrient agar plates and incubated at 28° C for 24hr. Different colony types were then sub-cultured by streaking onto NA plates and MacConkey agar plates and incubated for 24hr.

3.2 Extraction of DNA from Growing Bacteria

The DNA of isolated bacterial cells from each type of soil were extracted based on method used in previous study (Barghouthi 2011; Barghouthi and Al Zughayyar 2012), DNA was extracted as following: a loop full of pure fresh bacterial cells were treated with freshly prepared lysozyme (Sigma Chemical Co.); 5 mg/ml of sterile J-buffer; (100 mM Tris–HCl, 100 mM EDTA, 150 mM NaCl; pH 8; autoclaved), incubated in microfuge tube overnight at 36 C^{o} ±1, cells or incompletely lysed cells were collected by centrifugation and washed three times with J-Buffer, the pellet was lysed with 500 µl of sterile water, after centrifugation, the supernatant was transferred to a new microfuge tube, then 20 µl of 1% SDS were added and boiled for 10 min by steaming, then diluted two-fold with sterile water.

3.3 Application of The Universal Method for identification of isolated soil bacteria

One Golden primer mixture, the Golden 7 (G7) used from previous research (Barghouthi 2011) was utilized in this study to amplify the 16S rDNA of the isolated bacteria by PCR. Table (3.1) shows the primers sequences and their melting temperatures (*Tm*) which were used to prepare Golden mixture 7 (G7) which consisted of seven primers (four forward primers QUGP -Fn3, QUGP-F4, QUGP-Fn5, QUGP-Fn6 and three reverse primers QUGP-Rn1, QUGP-Rn2, QUGP-Rn3).

Table 3.1: Primers used in this study identified as “Al-Quds University 16S general primers” as described previously (Barghouthi 2011).

Primer	PCR primers	Oligomer: length and location	Tm (ºC)
QUGP-Fn3	5′-CAGGATTAGATACCCTGGTAGTCC-3′	24: 744–768	65
QUGP-F4	5′-CCGCCTGGGGAGTACG-3′	16: 840–856	59.5
QUGP-Fn5	5′-ACTCCTACGGGAGGCAGCAG-3′	20: 323–343	65
QUGP-Fn6	5′-CCAGCAGCCGCGGTAATAC-3′	19: 479–497	62
QUGP-Rn1	5′-GGCTACCTTGTTACGACTTC-3′	20: 1471–1468	58
QUGP- Rn2	5′-TGACGGGCGGTGTGTACAAG-3′	20: 1406–1386	63
QUGP- Rn3	5′-GGCGTGGACTACCAGGGTATC-3′	21: 775–752	65

A thermo cycler (Escohealthcare, Swift•Maxpro) heated lid was used to amplify DNA samples; 20 µl reactions were prepared by adding 0.5 µl in a premixed G7 containing 10 pmol of each primer, 0.5 µl of DNA sample, 10 µl Green Master Mix (2X) from Promega, and pure sterile water to 20 µl. All reactions were run using the following amplification protocol; initial denaturation at 95 ºC for 3 min and 33 cycles of denaturation at 94 ºC, annealing at 58 ºC and 60 ºC for 30s each, extension for 72 ºC for 145s and final extension at 72 ºC for 3 min.

3.4 Detection and Documentation

PCR products were run on a 1.2 % agarose gel (0.3g agarose, 25ml 1X TAE and 3µl of Ethidium Bromide at concentration 1mg/ ml). The 1X TAE electrophoresis running buffer that was prepared from a stock of (120.5g Tris base, 28.5ml glacial acetic acid and 50 ml of 0.5M EDTA (pH 8.0)) was used. LKB power supply (Biochrom, Cambridge, England) and UV Transilluminator (Dinco and Rhenium Industrial Ltd.,), 100-bp ladder (Invitrogen, Carlsbad, CA, United States) was used as molecular weight markers for DNA fragment size measurement.

3.5 Identity of amplicon producing QUPG-primers

Based on gel figures of PCR products amplified using G7 mixture and based on Table 3.2 according to Barghouthi study (Barghouthi 2011), the QUPG primers which gave positive amplification of 16s rDNA should be identified and reused in a second PCR run for preparing sufficient pure amplicon for DNA sequence determination (Barghouthi 2011).

Table 3.2: This key-table (from Barghouthi, 2011) allows the interpretation of PCR product based on amplicon size with paired primers (with author's permission).

Table 6 PCR product size with paired primers based on published rDNA gene *H. pylori* (HP) HPAG1 gi|108562424:c1140006-1138506 or *P. fluorescens* (PF) gi|70728250:122811-124349 of Pf-5 NC_004129

Forward primer Reverse primer	QUGP-F1 HP/PF	QUGP-Fn3 HP/PF	QUGP-F4 HP/PF	QUGP-Fn5 HP/PF	QUGP-Fn6 HP/PF
QUGP-Rn1	1463/1503	721/743	*633/639*	1142/1183	995/991 1277[b]
QUGP-Rn2	00/1396	*623/636*	*532/536*	00/1073	00/893 1097[b]
QUGP-Rn3	764/798	NP[a]	NP	465/473	287/298
QUGP-R4	847/881	112	NP	*530/561*	374/379
QUGP-Rn5	334/344	NP	NP	NP	NP
QUGP-Rn6	488/523	NP	NP	171/199	NP

The exact PCR product size for each primer pair may vary from one bacterium to another. Some primer sites are missing and their products are indicated as zero (00 bp)

[a] Primers that did not form PCR pairs are indicated as no pairing (NP). The exact PCR product size depends on bacterial species. Italicized products indicate size similarity

[b] *Homo sapiens* 18S sequence shares homologies with these primers and predicts the possible amplification of 1277 and 1097 bp region

After identification of predicted primers, the PCR products must be reproduced in pure form for sequencing, the selected sequence read from each sample was analyzed for sequence homology with a cutoff value above 98% using BLAST analysis tools in order to identify the bacterial species (Barghouthi 2011).

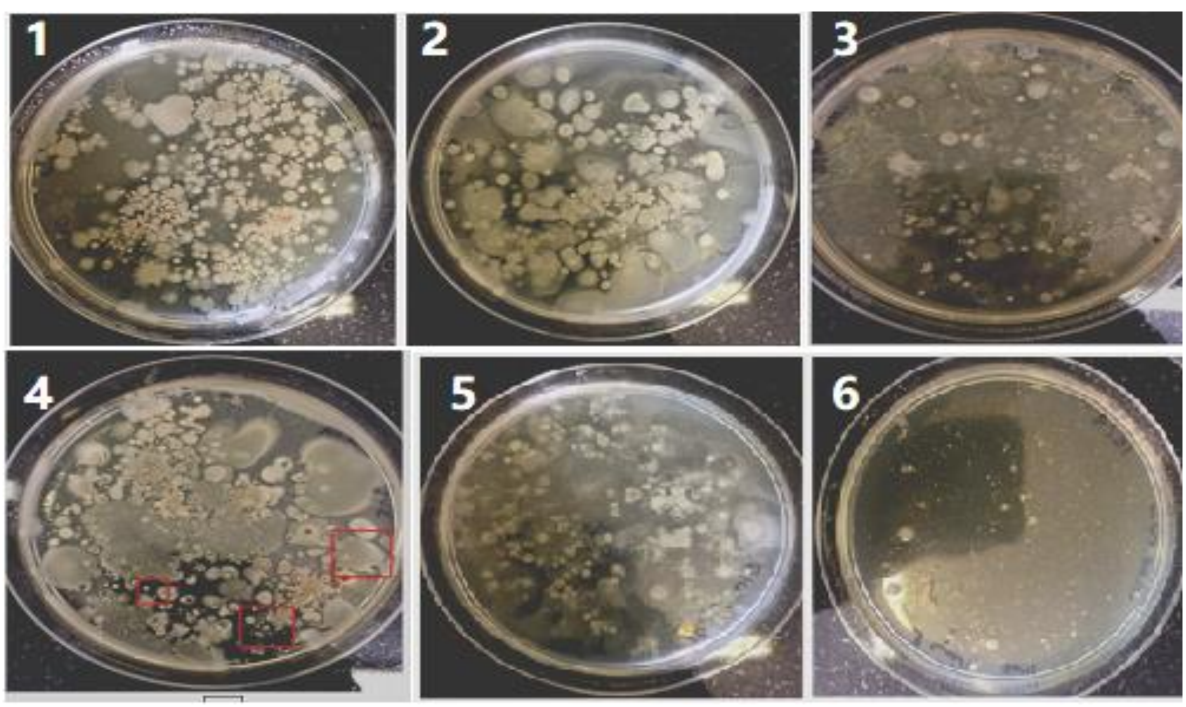
3.6 Bioinformatics analysis

In this study, GeneStudio software were used for analysis of the DNA sequences, then these DNA sequences were analyzed for sequence similarity using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome). Clustal W (https://www.ebi.ac.uk/Tools/msa/clustalo) and Mega X software were used for alignment of multiple sequences and in the construction of the phylogenetic trees.

Chapter 4: Results

4.1 Culture method

All 13 cultured nutrient agar plates with different soils gave isolated bacterial colonies of different types/morphologies. **Figure 4.1** show cultures of all samples;13 types of soil on NA, some plates showed the growth with different types of bacteria such as Tiberias Lake, Norway and Battir village in which colonies were different in size and morphology. Some plates showed one type of colonies as Sea of Marmara. Sub-cultures of each isolated colony on nutrient agar permitted colony purification as seen in **Figure 4.2 (A-F)**. Cultivation on macConkey agar produced no growth and this indicated that these types of isolated bacteria most likely did not belong to Gram negative bacteria.



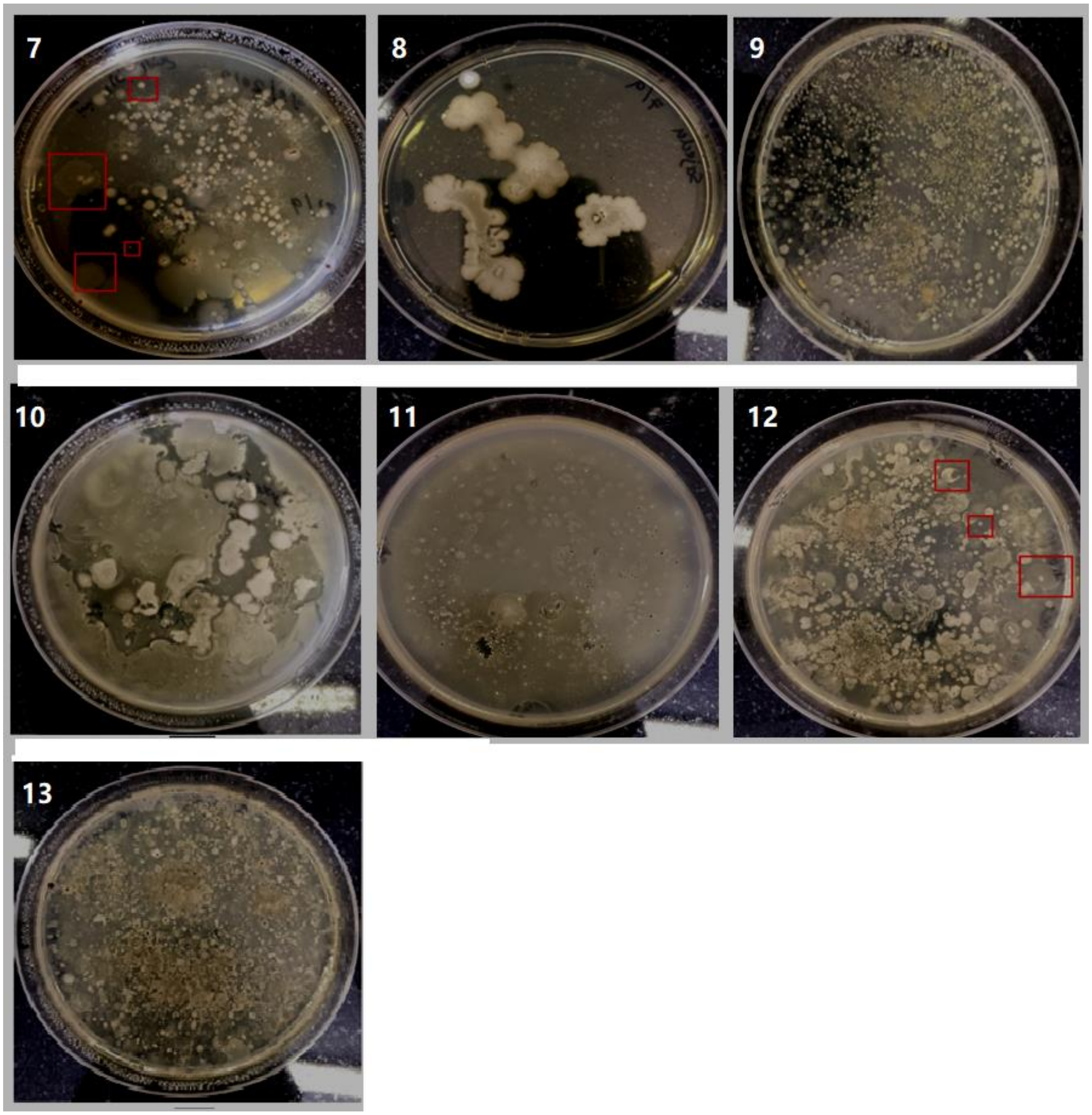


Figure 4.1: nutrient agar plates cultivated with 13 different type of soil; **1:** Tiberias Lake, Palestine (2009), **2:** Stanford University, San Francisco, Ca, USA (2009), **3:** Austria (2013), **4:** Amman, Jordan (2009), **5:** Oslo, Norway (2010), **6:** the Naqab desert, Palestine (2011), **7:** Battir village, Palestine (2010), **8:** Sea of Marmara, Avcilar, Turkey (2009), **9:** Eiffel Tower (2012), **10:** Hammamat Maein (2010), **11:** Switzerland -The Rhine River side (2012), **12**:

Basel metro, Switzerland (2012),**13:** Louvre Museum, Paris, France (2012). The red squares represent the chosen bacteria colony for subcultures.

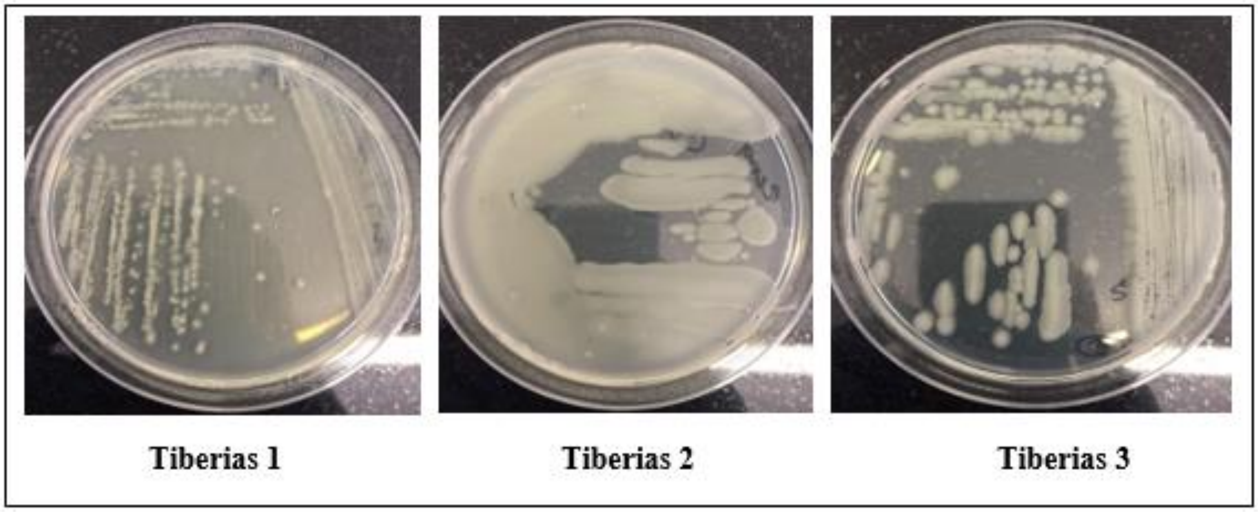


Figure 4.2-A: Subcultures of three types of bacteria isolated from main plate of Tiberias culture.

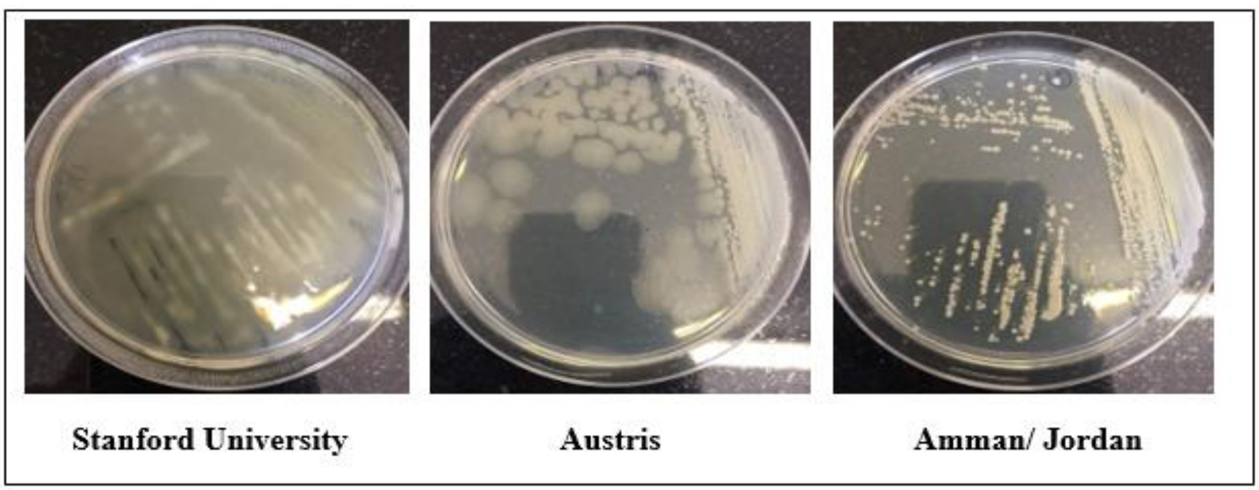


Figure 4.2-B: Sub-cultures of three types of different bacteria isolated from three main plates of Stanford University, Austria and Amman/ Jordan, respectively.

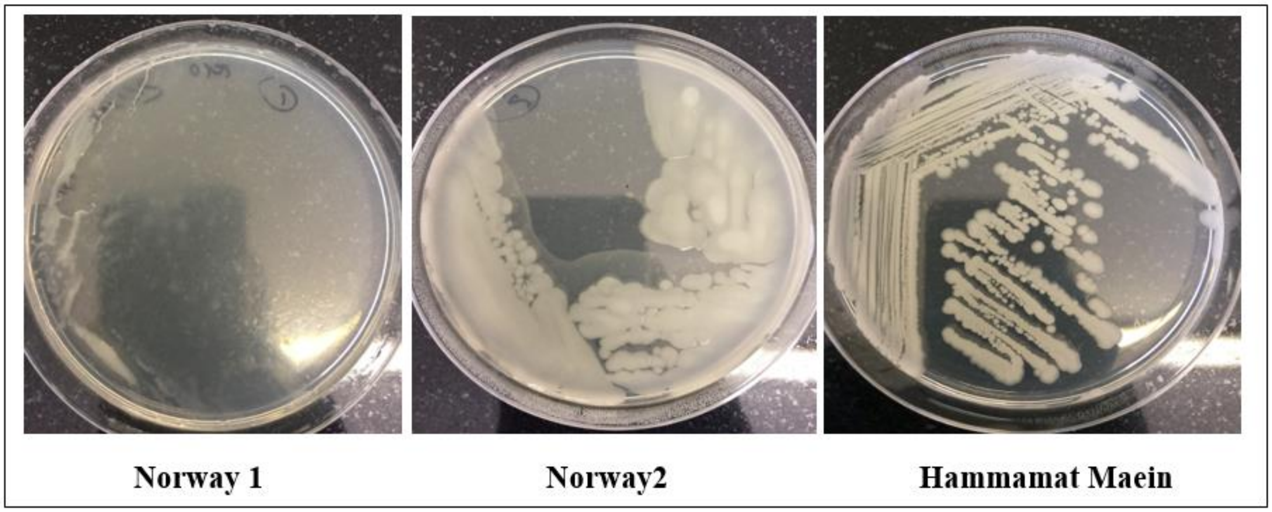


Figure 4.2-C: Sub-cultures of three different types of bacteria isolated from two main plates of Norway and Hammamat Maein, respectively.

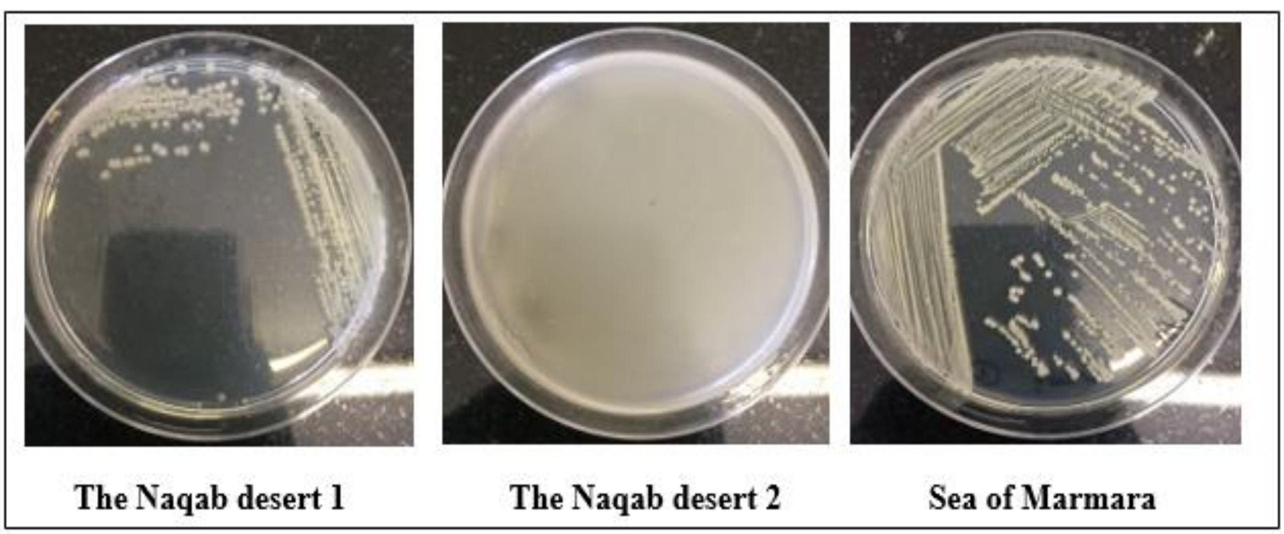


Figure 4.2-D: Sub-cultures of three different types of bacteria isolated from two main plates of the Naqab desert and Sea of Marmara, respectively.

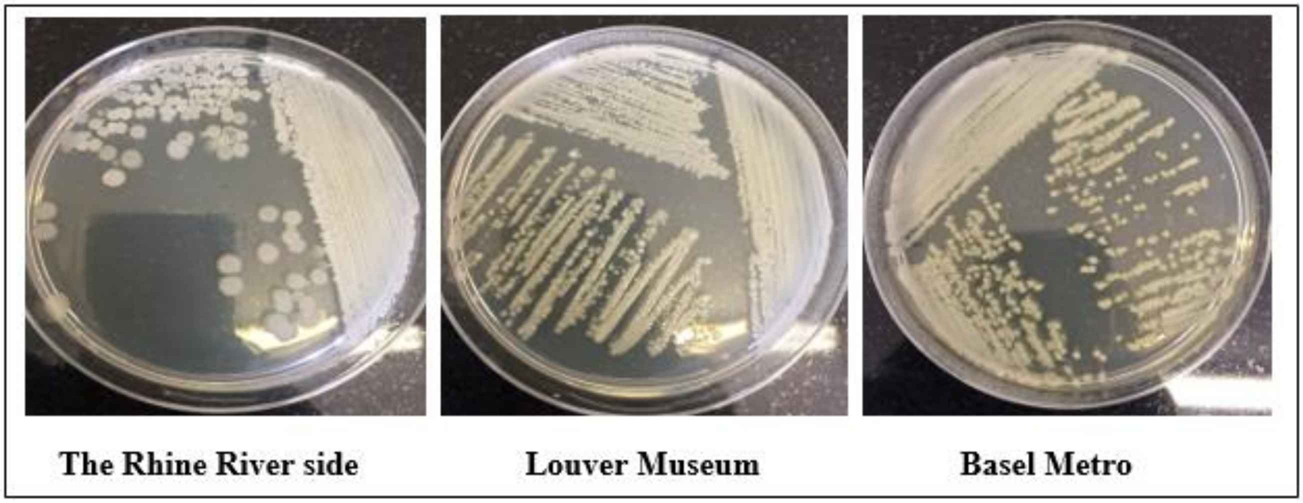


Figure 4.2-E: Sub-cultures of three different types of bacteria isolated from three main plates of Switzerland-the Rhin River, Louvre Musem and Bazel train, respectively.

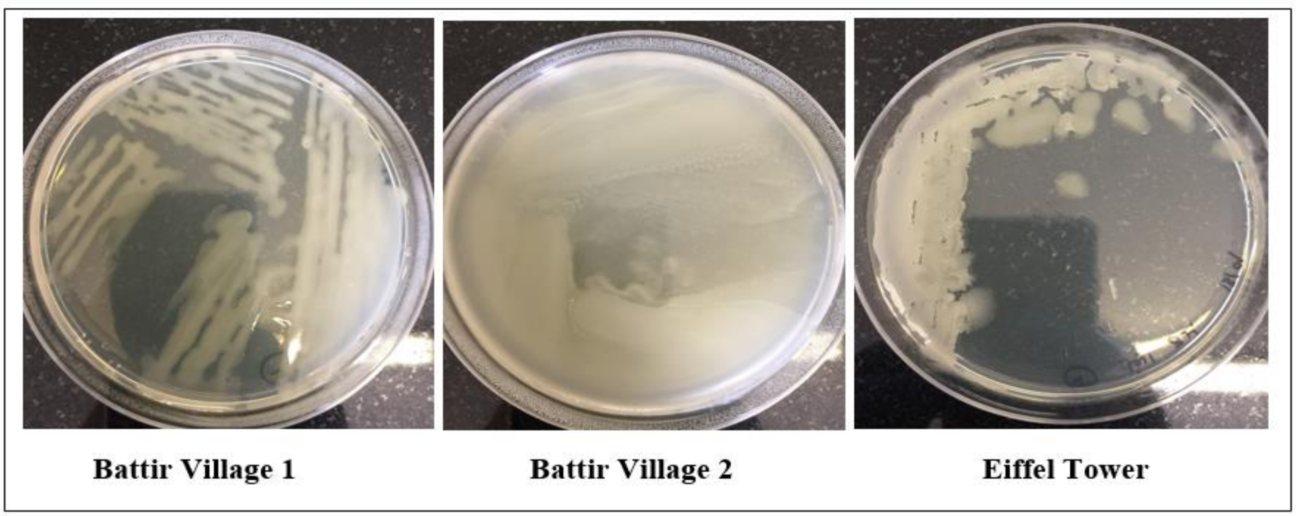


Figure 4.2-F: Sub-cultures of three different types of bacteria isolated from two main plates of Battier Village, Eiffel Twoer, respectively.

4.2 Amplification of the 16s rDNA of 38 sample using Golden Mixture 7 (G7)

DNA sample (38) which were extracted from different soil sources were amplified by G7 mixture and gave positive amplification product of 16s rDNA with different size, as seen in **Figures 4.3;** most samples were gave positive amplification product with size of 500bp and 600bp, some samples exhibit faint band with molecular size at 300bp as in figure 4.3-A, figure 4.3-C.

The primers were identified based on a known PCR products with paired primers which were published in a previous study by Barghouthi in 2011 (Barghouthi 2011), depending on" Table 6", paired primers (QUGP-F4*QUGP-Rn2) and (QUGP-Fn3*QUGP-Rn2) were responsible for amplification of a 500bp and 600bp amplicons, respectively. Figure4.3 A and B Lane number 1 isolate

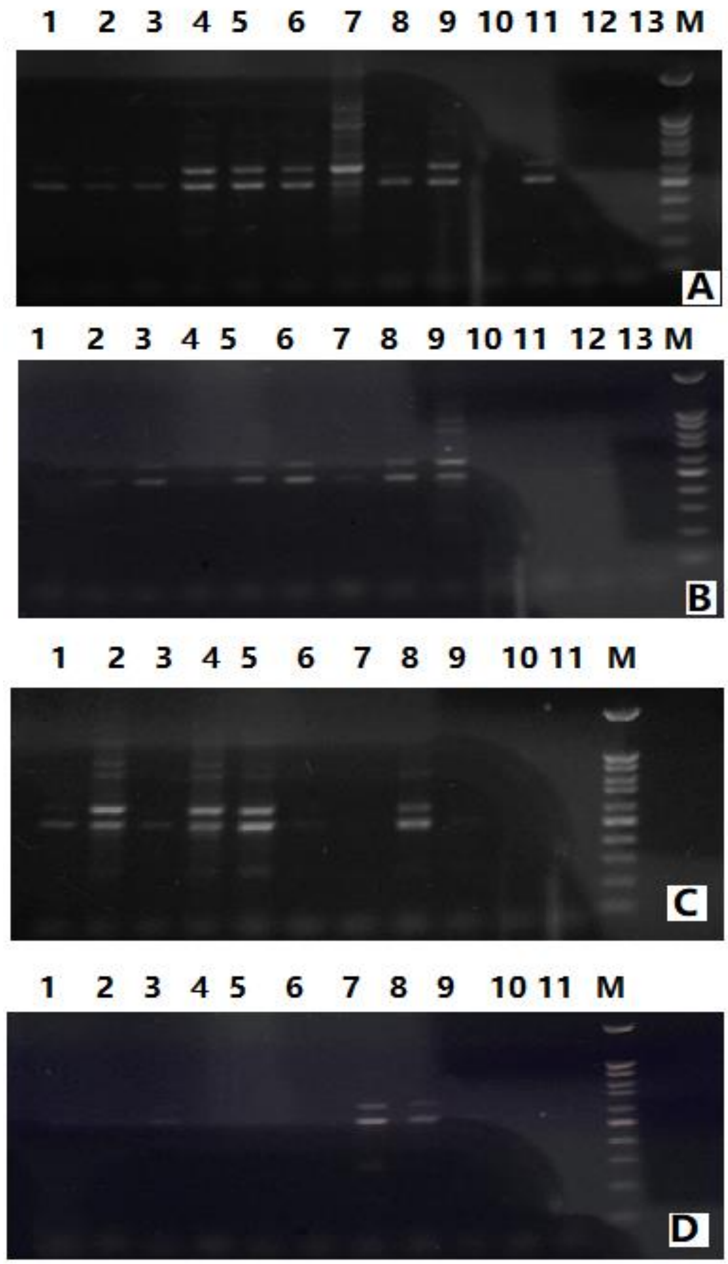


Figure 4.3: 1.2% agarose gel electrophoresis representing results of amplification of extracted DNA obtained from each bacterial isolate by G7 mixture. A and B: lanes 1-11 represent DNA samples, lane 12: *S. pyogenes* and lane 13 is negative control. 3-C D: lanes 1-8 represent DNA samples, lane 9: *S. pyogenes* and lane 11 is negative control.

4.3 DNA sequence alignment and bacterial identification

After amplicons were sequenced and aligned using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome), different bacilli species were identified in different soil sources. *Lysinibacillus. sp* was detected in different soil sources; Stanford University and Tiberias Lake 1 with similarity (>98%), in Eiffel Tower and Sea of Marmara with similarity (<98%). *Lysinibacillus fusiformis* was detected in Louver Museum, Austria, Battier Village (>98%). *Bacillus cereus* with different strain was detected in Tiberias Lake 2 & 3, Basel metro and Norway 2 with similarity (>98%). Other *Bacillus spp* with different strain was detected in Tiberias lake3, Norway 1 and Sea of Marmara with similarity (>98%) and in Jordan and the Naqab desert (<98%). The isolate of Hammamat Maein was *Peribacillus asahii* strain OM18 (99%) based on F4 and Rn2 primers, in the Rhine River side sample, the isolates were *Bacillus amyloliquefaciens* strain CS13 and *Bacillus subtilis* strain UIS0380 (99%) based on used (F4*Rn2) and (Fn3*Rn2) respectively.

In the Naqab desert 2 sample, bacteria were identifying of *Bacillus safensis* with similarity (99%) based on F4*Rn2 primers and *Bacillus paranthracis* strain AA38 (98) based on Fn3* Rn2 primer. While in Eiffel Tower sample, *Bacillus safensis* strain PgKB20 was detected but with sequencing identity (98%) based on FN3*Rn2 primers. In Norway1 and Naqab desert 2, two types of bacteria were detected which *Lelliottia amnigena* strain SatS.3 and *Bacillus paranthracis* strain AA38 with sequencing identity (98%) and (88%), respectively.

In Battier Village 1, *Lysinibacillus sphaericus* strain DH-B01 was identified with sequencing identity (99%) based on F4*Rn2. The control *S. pyogenes* was identified as *Streptococcus* *pyogenes* strain 4063-05 (99%) based on F4 *Rn2 primers. **Table 4.1** show the soil sources and identified bacteria using Blast tool.

Table 4.1 show the soil sources (see Appendix 1) and identified bacteria using Blast tool.

Soil source	Position on Figure 3.3 Gel	Blast identification, DNA sequence (%) identity	PCR and sequencing primers sequenced with (primer)
Stanford University	**A1**	***Lysinibacillus sp.* BAB-5854 (99)**	**(F4), Rn2**
Louvre Museum	**A5**	***Lysinibacillus fusiformis* strain WS1-3 (99)**	**(F4), Rn2**
Louvre Museum	**A5**	***Lysinibacillus fusiformis* strain WS1-3 (99)**	**(Fn3), Rn2**
Tiberias Lake1	**A8**	***Lysinibacillus sp*. strain CL03-7 (99)**	**(F4), Rn2**
Tiberias Lake2	**A9**	***Bacillus cereus* strain af-B25 (99)**	**(F4), Rn2**
Tiberias Lake2	**A9**	***Bacillus cereus* strain UIS0932 (99)**	**(Fn3), Rn2**

Tiberias Lake3	**A11**	***Bacillus sp*. (in: Bacteria) strain CM-CNRG 602 (99)**	**(F4), Rn2**
Tiberias Lake3	**A11**	***Bacillus cereus strain JCM 2152* (98)**	**(Fn3), Rn2**
Austria	**B3**	***Lysinibacillus fusiformis* strain WS1-3 (99)**	**(F4), Rn2**
Austria	**B3**	***Lysinibacillus fusiformis* strain WS1-3 (99)**	**(Fn3), Rn2**
Hammamat Maein	**B5**	***Peribacillus asahii* strain OM18 (99)**	**(F4), Rn2**
Hammamat Maein	**B5**	***Peribacillus asahii* strain OM18 (95)**	**(Fn3), Rn2**
Basel Metro	**B6**	***Bacillus cereus* strain S8 (99)**	**(F4), Rn2**
Basel Metro	**B6**	***Bacillus cereus* strain HFBPR40 (99)**	**(Fn3), Rn2**
The Rhine River	**B9**	***Bacillus amyloliquefaciens* strain CS13 (100)**	**(F4), Rn2**
The Rhine River	**B9**	***Bacillus subtilis* strain UIS0380 (100)**	**(Fn3), Rn2**
Battir 1	**C2**	***Lysinibacillus sphaericus* strain DH-B01 (99)**	**(F4), Rn2**

Battir 1	**C2**	***Lysinibacillus fusiformis* strain P-R2A48 (99)**	**(Fn3), Rn2**
Battir 2	**C3**	***Bacillus cereus strain EdyKolBc*** **(99)**	**(F4), Rn2**
Norway1	**C4**	***Bacillus sp.* H2O-1 (99)**	**(F4), Rn2**
Norway1	**C4**	***Lelliottia amnigena* strain SatS.3 (88)**	**(Fn3), Rn2**
Norway2	**C5**	***Bacillus cereus* strain af-B25 (99)**	**(F4), Rn2**
Naqab Desert 1	**C6**	***Bacillus sp.* strain 207.1 (97)**	**(F4), Rn2**
Naqab Desert 2	**C8**	***Bacillus safensis* strain M46 (99)**	**(F4), Rn2**
Naqab Desert 2	**C8**	***Bacillus paranthracis* strain AA38 (98)**	**(Fn3), Rn2**
Amman/ Jordan	**D3**	***Bacillus sp*. Uz1013 (96)**	**(F4), Rn2**
Eiffel Tower	**D7**	*Lysinibacillus sp.* **BAB-4328 (96)**	**(F4), Rn2**
Eiffel Tower	**D7**	***Bacillus safensis* strain PgKB20 (98)**	**(Fn3), Rn2**
Sea of Marmara	**D8**	***Bacillus sp*. strain DMAJP21 (99)**	**(F4), Rn2**
Sea of Marmara	**D8**	**Lysinibacillus sp. strain NP05 (97)**	**(Fn3), Rn2**

Control **(*Streptococcus pyogenes*)**	**A12, B12,** **C9, D9**	***Streptococcus pyogenes* strain 4063-05** **(99)**	**(F4), Rn2**

4.4 Relatedness of isolates

The multiple sequence alignment of 16s rDNA was done on different soils isolates using ClastalW (https://www.ebi.ac.uk/Tools/msa/clustalo/). Phylogenetic trees shown below, were constructed using MEGA-X software (Molecular Evolutionary Genetics Analysis) to show the degree of similarity among different isolates as represented by the identified matching organism.

4.4.1. Maximum Likelihood relationship among different species of Bacillus in different soil samples:

The multiple alignment file in **Appendix 2** was used to create phylogram using MEGA-X software based on bootstrap analysis of 1000 replicates to evaluate the confidence of trees topologies. The *Pseudomonas aeruginosa* strain DSM 50071 was used as an outgroup in construct the phylogeny tree of 500bp DNA sequences.

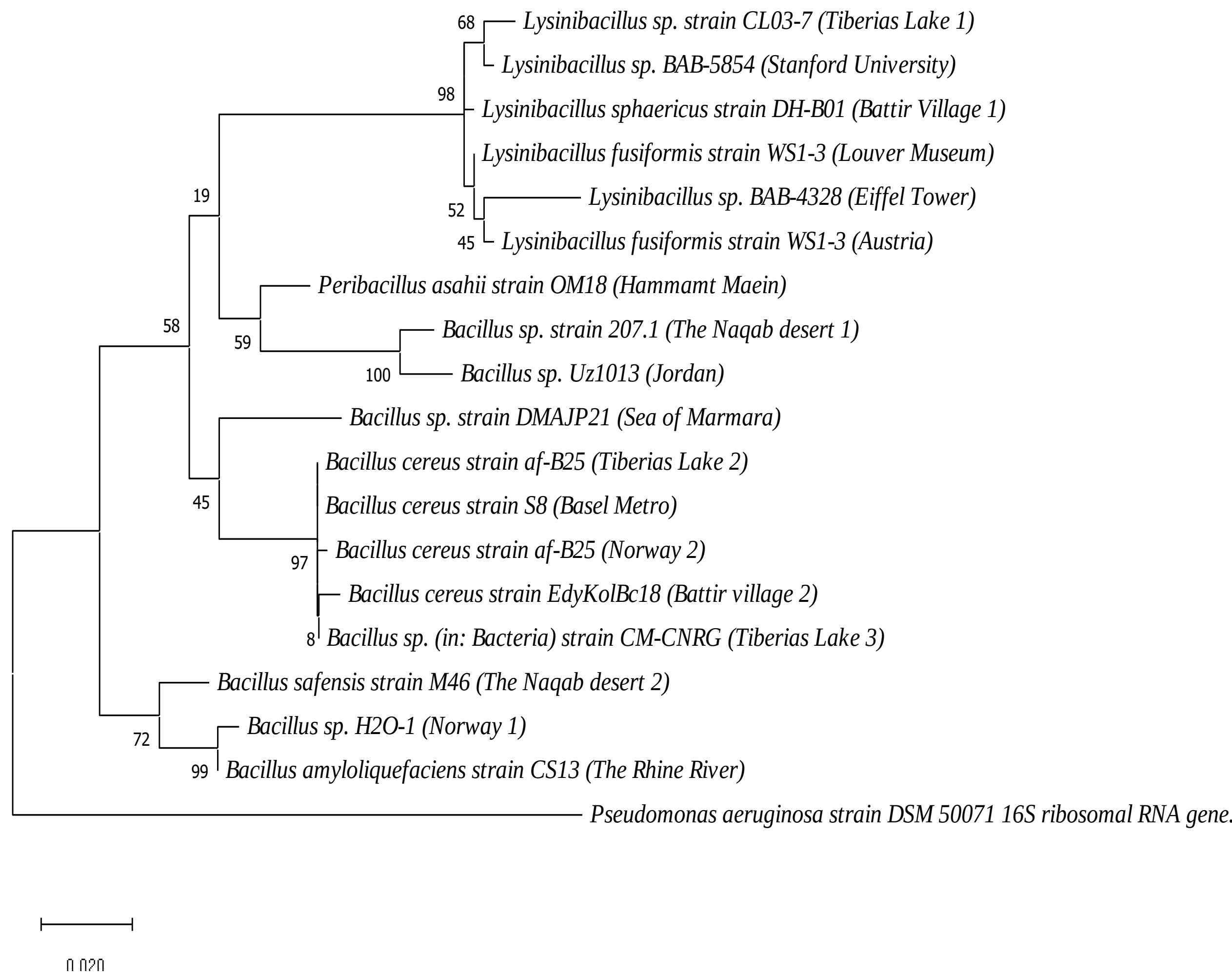


Figure 4.4.1: Maximum Likelihood relationship among different species of Bacillus in different soil samples, The *Pseudomonas aeruginosa* strain DSM 50071 was used as an outgroup.

4.4.2 Maximum Likelihood relationship among different species of Bacillus in different soil samples:

The multiple alignment file in **Appendix 2** was used to create phylogram using MEGA-X software based on bootstrap analysis of 1000 replicates to evaluate the confidence of trees topologies. The *Pseudomonas kribbensis* strain CHA-19 was used as an outgroup in construct the phylogeny tree of 600bp DNA sequences.

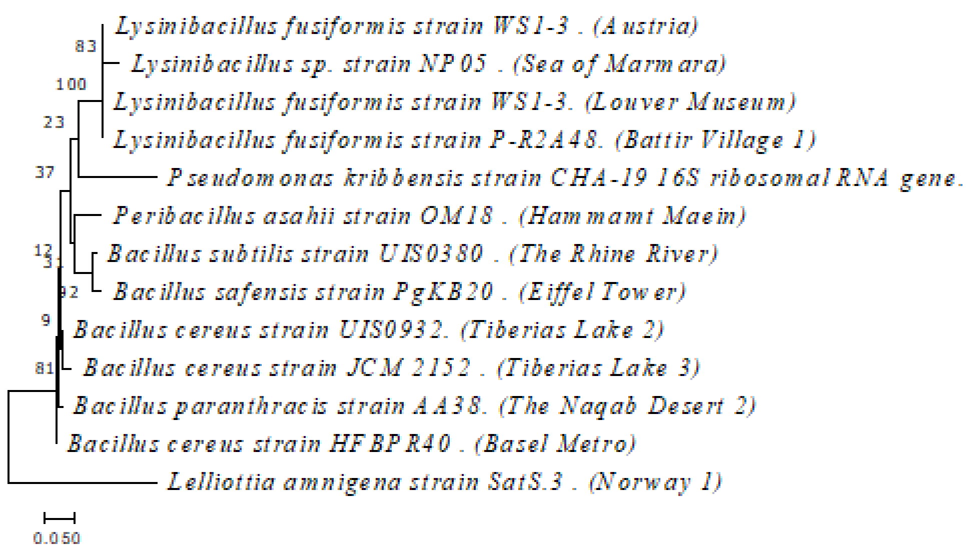


Figure 4.4.2: Maximum Likelihood relationship among different species of Bacillus in different soil samples, The *Pseudomonas kribbensis* strain CHA-19 was used as an outgroup .

Chapter 5: Discussion

5.1 Discussion

Population of bacteria in soil can boom or decline in parts of soil in few days depending on changes in soil moisture, temperature, and carbon substrate (Bhattarai et al, 2015).

In this study, most isolates from different soil sources belonged to endospore forming genera including *Bacillus, Lysinibacillus,* and *Peribacillus.* Clearly, their ability to form endospores enabled them to survive harsh environments; An indicative of their abilities to resist environmental stress. Most aerobic spore formers are ubiquitous and can be readily isolated from different sources, also soil is considered the primary habitat of endospore-forming bacteria (Todar 2004).

Lysinibacillus fusiformis is rod-shaped, gram-positive bacterium that belongs to family *Bacillaceae* and genus *Lysinibacillus*, it can cause topical skin ulcers, respiratory illnesses and severe sepsis. It can be isolated from different environmental sources including farming soil and factory waste water. Its endospores can resists high temperatures, ultraviolet, and desiccation; therefore can survive for long-periods of time (Sulaiman et al, 2018). *Lysinibacillus* species differ from *Bacillus* species in cell wall composition; the cell walls of *Lysinibacillus* consist of two diagnostic types of amino acids; lysine and aspartate, whereas cell walls of genus *Bacillus* contain meso-diaminopimelic acid (Sulaiman et al, 2018). *Lysinibacillus fusiformis* was isolated from cosmetic samples by using 16s rDNA sequencing methods (Sulaiman et al, 2018).

Bacillus subtilis which was isolated from the Rhine River side soil belongs to family *Bacillaceae.* It is found naturally in soils, in 2019; *Bacillus subtilis* was isolated from different soil samples by Madika et al for screening its ability to produce amylase (Madika et al, 2019). In another study; *Bacillus subtilis* MH-4 isolates from soils showed good antimicrobial activity, the antibiotic was proven to be bacitracin (Jamil et al, 2007).

Bacillus cereus isolates were isolated from different soil sources in this work including; Norway, Basel Metro and Tiberias Lake. In 2014, Shukla et al showed that most isolates of *Bacillus spp* from different samples of soil have a high level of resistance or decreased susceptibility to Ampicillin and Ceftriaxone (Singh et al, 2018).

Bacillus amyloliquefaciens isolate was from Rhine River soil with 100% similarity, in 2015, Xiong et al isolated the *Bacillus amyloliquefaciens* JK6 from the rhizosphere soil of healthy tomato plants and to evaluate the antimicrobial compound produced by JK6 against *Ralstonia solanacearum* (Xiong et al, 2015); a Gram negative bacterium that causes wilt diseases in plants. *B. cereus* and *B. amyloliquefaciens* are considered the most types of *Bacillus* species with abilities of producing Biological Control Agents (BCAs) which in turn control soil-borne plant pathogens (Xiong et al, 2015).

Bacillus paranthracis with similarity of 98% was isolated from the Naqab desert soil, Osman and Yin identified *B. paranthracis* as plant growth-promoting rhizobacteria (PGPR), they showed how it can enhance the growth of pea plant (Osman and Yin 2018). *Bacillius safensis* was isolated from Naqab soil and Eiffel Tower samples with more than 98% similarity, common habitats of *B. safensis* are industrial effluents, oil polluted sites, saline desert, insect guts, rhizosphere, human and animal excreta (Lateef et al, 2015). *B. safensis* was first isolated as contaminant in a spacecraft assembly facility (SAF) at the Jet Propulsion Laboratory (Satomi et al, 2006).

Peribacillus asahii is a mesophilic bacterium that was isolated from soil also belongs to family *Bacillaceae*, in 2004; Yumoto et al discovered a novel bacterium species which was isolated from soil with its ability to deodorize the bad smell generated from short fatty acid oxidation, this bacterium is classified as *Bacillus asahii* (Yumoto et al, 2004).

Only one isolate was a Gram negative bacterium that belonged to family *Enterobacteriaceae*, it was identified as *Lelliottia amnigena*, according to taxonomy browser NCBI and has other homotypic names as *L. amnigena* and *Enterobacetr amunigenus.* It is a new *Enterobacter sp.* that causes soft rot disease on harvested onion (Liu and Tang 2016)

Of the identified bacteria in this study, the *Bacillus spp.* were dominant. *Bacillus spp.* are considered most ubiquitous among other bacterial genera of the soil, it has multiple ecological functions in soil ecosystems, the cause of abundance of *Bacillus spp*. (Saxena et al, 2020). Many species of *Bacillus* have a wide range applications in biodefense, biofuel production, biofertilizer, biocontrol, and bioenzyme production (Liu et al, 2019). *Bacillus* species are responsible of several infections in human including self-limited food poisoning, localized infections related to trauma such as ocular infections, and anthrax (Baumgardner 2012).

Bacillus spp show antagonistic activity through exerting extracellular metabolites such as antibiotics and cell wall hydrolases. Also, *Bacillus spp*. can improve plant response to pathogen attack by triggering systemic resistance. *Bacillus spp* enhance plant growth through nitrogen fixation, phytohormone production and phosphate solubilization.

All soil samples in this study produced positive results, they were stored under aerobic conditions for many years (more than 9 years) in the laboratory at room temperature and this could be the cause for presence of mostly endospore formers and the rarity of other types of bacteria including most Gram-negative bacteria. This was confirmed when soil samples were cultured directly on MacConkey agar, mostly plates remain clean with no growth, except for one appeared to be of the pseudomonads; probably *Lelliottia* which showed similarity to *Pseudomonas krebbensis* (600-phylogenetic tree, Section 4.4.2). To show the differences between isolates; phylogenetic tree was constructed using ClustalW (https://www.ebi.ac.uk/Tools/msa/clustalo/) and Mega X software for sequence lengths of 500bp and another for 600bp sequences. Distance between isolates showed little variation among similar species but wider variations could be observed as seen with *Lelliottia amnigena* and *Pseudomonas* serving as outgroup sequences in the phylogenetic trees.

Depending on this study, the presence of *Bacillus spp.* in different soil samples which were stored for several years made it the most dominant Genus found in soils. Ability of this type of bacteria to survive under stress environment through formation of endospores help to protect it from these harsh conditions. The absence of other types of bacteria which are usually found in fresh soils may be related to storage conditions which were not suitable for maintaining them viable (inside the plastic bags in drawers (in dark)) leading to their death. Additionally; some types pf bacteria which are usually found in fresh soils are non-spores former and were most likely unable to survive these storage conditions.

5.2 Conclusions

Depending on results of this study, numbers of *Bacillus spp.* were identified in different samples of soil which were stored for several years. This indicated that most common type of bacteria that can live in stored soils for long period of time were endospore formers belonging to Family *Bacillaceae* mostly of the genera *Bacillus*, *Lysinibacillus,* and *Peribacillus.* Therefore, keeping clean environment is a necessity for maintaining healthy microbial population and healthy balanced living environment for higher organisms.

The absence of less tolerant species in stored soils for years under aerobic condition may have resulted in the death of these species which are usually present in fresh soil samples while retaining the endospore-forming species.

Based on this study, some sources of samples may contain specific species such as observed with Hammamt Maein groundwater, Battir village, Eiffel Tower and Austria. It also can be concluded that *Bacillus spp.* were the most dominant bacteria that found in stored soils for years rather than other types of bacteria.

5.3 Recommendations

This study could carry important recommendations to given the great economic and ecological importance of *Bacillus spp*. that found in soils. *Bacillus spp*. have many applications in biodefense, biocontrol, biofuel production and biofertilizers. Due to this importance of *Bacillus* and due to their ability to survive for several years in soil, it is necessary to find advanced methods for rapid and comprehensive research to study the important applications of these *Bacillus spp.* which have low cost of maintenance, preservation and propagation leading to low-cost production of their byproducts on cheap growth media.

References

- Amend, A. (2014). "From dandruff to deep-sea vents: Malassezia-like fungi are ecologically hyper-diverse." PLoS pathogens **10**(8): e1004277-e1004277.

- Barghouthi, S. A. (2011). "A universal method for the identification of bacteria based on general PCR primers." Indian journal of microbiology **51**(4): 430-444.

- Barghouthi, S. A. and D. K. Al Zughayyar (2012). "Detection of Neisseria meningitidis and unknown Gammaproteobacteria in cerebrospinal fluid using the two-step universal method." African Journal of Microbiology Research **6**(14): 3415-3424.

- Baumgardner, D. J. (2012). "Soil-Related Bacterial and Fungal Infections." The Journal of the American Board of Family Medicine **25**(5): 734.

- Bhattarai, A., B. Bhattarai and S. Pandey (2015). "Variation of soil microbial population in different soil horizons." J Microbiol Exp **2**(2): 00044.

- Cohen, J. I., J. A. Bartlett and G. R. Corey (1987). "Extra-intestinal manifestations of salmonella infections." Medicine **66**(5): 349-388.

- Egamberdiyeva, D. (2005). "Characterization of Pseudomonas species isolated from the rhizosphere of plants grown in Serozem soil, semi-arid region of Uzbekistan." TheScientificWorldJournal **5**.

- Fakruddin, M. and K. Mannan (2013). "Methods for analyzing diversity of microbial communities in natural environments." Ceylon Journal of Science (Biological Sciences) **42**(1).
- Fenlon, D. R. (1999). "Listeria monocytogenes in the natural environment." FOOD SCIENCE AND TECHNOLOGY-NEW YORK-MARCEL DEKKER-: 21-38.
- Heyndrickx, M. (2011). "The Importance of Endospore-Forming Bacteria Originating from Soil for Contamination of Industrial Food Processing." Applied and Environmental Soil Science **2011**: 561975.
- Insam, H. (2001). "Developments in soil microbiology since the mid 1960s." Geoderma **100**(3): 389-402.
- Jacobsen, C. S. and T. B. Bech (2012). "Soil survival of Salmonella and transfer to freshwater and fresh produce." Food Research International **45**(2): 557-566.
- Jamil, B., F. Hasan, A. Hameed and S. Ahmed (2007). "Isolation of Bacillus subtilis MH-4 from soil and its potential of polypeptidic antibiotic production." Pakistan journal of pharmaceutical sciences **20**: 26-31.
- Joseph, S. J., P. Hugenholtz, P. Sangwan, C. A. Osborne and P. H. Janssen (2003). "Laboratory Cultivation of Widespread and Previously Uncultured Soil Bacteria." Applied and Environmental Microbiology **69**(12): 7210.
- Lateef, A., I. Adelere and E. Gueguim-Kana (2015). "The biology and potential biotechnological applications of Bacillus safensis." Biologia **70**: 411-419.

- Liu, J., X. Cui, Z. Liu, Z. Guo, Z. Yu, Q. Yao, Y. Sui, J. Jin, X. Liu and G. Wang (2019). "The Diversity and Geographic Distribution of Cultivable Bacillus-Like Bacteria Across Black Soils of Northeast China." Frontiers in microbiology **10**: 1424-1424.
- Liu, S. and Y. Tang (2016). "Identification and Characterization of a New Enterobacter Onion Bulb Decay Caused by Lelliottia amnigena in China." Applied Microbiology: open access **2**.
- Madigan, M. T. (2000). Bacterial Habitats in Extreme Environments. Journey to Diverse Microbial Worlds: Adaptation to Exotic Environments. J. Seckbach. Dordrecht, Springer Netherlands**:** 61-72.
- Madika, A., J. Ameh and D. Machido (2019). "Isolation and Screening of Bacillus subtilis from Soil for Amylase Production." **2**: 82-86.
- Nishiuchi, Y., T. Iwamoto and F. Maruyama (2017). "Infection Sources of a Common Non-tuberculous Mycobacterial Pathogen, Mycobacterium avium Complex." Frontiers in Medicine **4**(27).
- Osman, N. I. and S. Yin (2018). "Isolation and characterization of pea plant (Pisum sativum L.) growth-promoting Rhizobacteria." African Journal of Microbiology Research **12**(34): 820-828.
- Patyar, S., R. Joshi, D. S. P. Byrav, A. Prakash, B. Medhi and B. K. Das (2010). "Bacteria in cancer therapy: a novel experimental strategy." Journal of Biomedical Science **17**(1): 21.

- Rahbar, M., A. Lamei, H. Babazadeh and S. A. Yavari (2010). "Isolation of rapid growing mycobacteria from soil and water in Iran." African Journal of Biotechnology **9**(24): 3618-3621.
- Satomi, M., M. Duc and K. Venkateswaran (2006). "Bacillus safensis sp nov., isolated from spacecraft and assembly-facility surfaces." International journal of systematic and evolutionary microbiology **56**: 1735-1740.
- Saxena, A. K., M. Kumar, H. Chakdar, N. Anuroopa and D. Bagyaraj (2020). "Bacillus species in soil as a natural resource for plant health and nutrition." Journal of Applied Microbiology.
- Sethi, S., R. Kumar and S. Gupta (2013). "Antibiotic production by microbes isolated from soil." International Journal of Pharmaceutical Sciences and Research **4**(8): 2967.
- Singh, P., R. Sharma, A. K. Shukla and R. Singh (2018). "Isolation of Bacillus spp. from Soil for Antimicrobial Production and Antibiotic Resistance." Adv Biotech. & Micro **8**(4): 1-5.
- Stefanis, C., A. Alexopoulos, C. Voidarou, S. Vavias and E. Bezirtzoglou (2012). "Principal methods for isolation and identification of soil microbial communities." Folia microbiologica **58**.
- Sulaiman, I. M., Y.-H. Hsieh, E. Jacobs, N. Miranda, S. Simpson and K. Kerdahi (2018). "Identification of Lysinibacillus fusiformis Isolated from Cosmetic Samples Using MALDI-TOF MS and 16S rRNA Sequencing Methods." Journal of AOAC INTERNATIONAL **101**(6): 1757-1762.

- Swe, S., K. Lae and H. Ngwe (2018). "Isolation and Identification of Pseudomonas aeruginosa from the Clinical Soil."
- Todar, K. (2004). Todar's online textbook of bacteriology, University of Wisconsin Madison^ eWisconsin Wisconsin.
- Xiong, H., Y. Li, Y. Cai, Y. Cao and Y. Wang (2015). "Isolation of Bacillus amyloliquefaciens JK6 and identification of its lipopeptides surfactin for suppressing tomato bacterial wilt." RSC Adv. **5**.
- Yumoto, I., K. Hirota, S. Yamaga, Y. Nodasaka, T. Kawasaki, H. Matsuyama and K. Nakajima (2004). "Bacillus asahii sp nov., a novel bacterium isolated from soil with the ability to deodorize the bad smell generated from short-chain fatty acids." International journal of systematic and evolutionary microbiology **54**: 1997-2001.

Appendix 1

Table A1.1. The sources of sample soil, the type of isolate in each soil and who gave the soil samples.

*The yellow star of figures below refers to the specific place of soil samples which were collected from a depth of 5-10cm by Prof. Sameer A. Barghouthi.

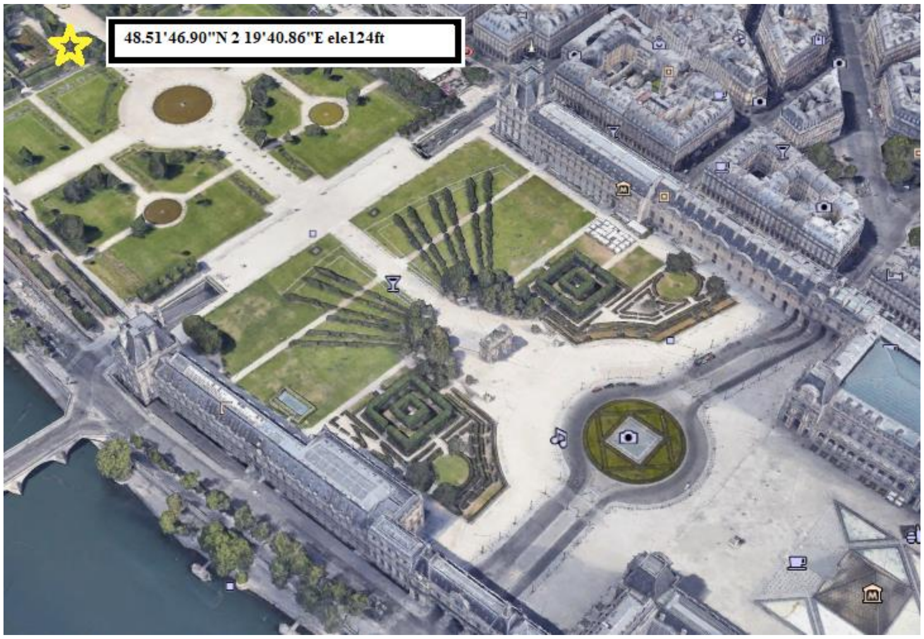
Sample location	**Position on Figure 3.3 Gel**	**Blast identification, DNA sequence (%) identity**	**Gift from**
Stanford University	**A1**	***Lysinibacillus sp.*** **BAB-5854 (99)**	***Dr. Samira Barghouthi***
Austria	**B3**	***Lysinibacillus fusiformis*** **strain WS1-3 (99)**	***Dr. M Qadi***
Austria	**B3**	***Lysinibacillus fusiformis*** **strain WS1-3 (99)**	***Dr. M. Qadi***
Norway1	**C4**	***Bacillus sp.*** **H2O-1 (99)**	***Traveler***
Norway1	**C4**	**Bacillus sp. (in: Bacteria) strain NRC5 (89)**	***Traveler***
Norway2	**C5**	***Bacillus cereus*** **strain af-B25 (99)**	***Traveler***
Naqab Desert 1	**C6**	***Bacillus sp.*** **strain 207.1 (97)**	***Gift from***

			Ms. Dua Abo Sway
Naqab Desert 2	**C8**	***Bacillus safensis* strain M46 (99)**	***Gift from*** ***Ms. Abo Sway***
Naqab Desert 2	**C8**	***Bacillus paranthracis* strain AA38** **(98)**	***Gift from*** ***Ms. Abo Sway***
Amman/ Jordan	**D3**	***Bacillus sp*. Uz1013** **(96)**	***Gift from*** ***Mrs.Samia*** ***Barghouthi***
Control **(*Streptococcus pyogenes*)**	**A12, B12,** **C9, D9**	***Streptococcus pyogenes* strain 4063-05** **(99)**	***Clinical isolate,*** ***Faculty of*** ***Medicine***

Eiffel Tower	**D7**	***Lysinibacillus sp.* BAB-4328 (96)**
Eiffel Tower	**D7**	***Bacillus safensis* strain PgKB20 (98)**

Hammamat Maein	**B5**	***Peribacillus asahii* strain OM18**	**(99)**
Hammamat Maein	**B5**	***Peribacillus asahii* strain OM18**	**(95)**

Basel Metro	**B6**	***Bacillus cereus* strain S8**	**(99)**
Basel Metro	**B6**	***Bacillus cereus* strain HFBPR40**	**(99)**



Louvre Museum	**A5**	***Lysinibacillus fusiformis* strain WS1-3 (99)**
Louvre Museum	**A5**	***Lysinibacillus fusiformis* strain WS1-3 (99)**

The Rhine River	**B9**	***Bacillus amyloliquefaciens* strain CS13 (100)**
The Rhine River	**B9**	***Bacillus subtilis* strain UIS0380 (100)**

Tiberias Lake1	**A8**	***Lysinibacillus sp*. strain CL03-7 (99)**
Tiberias Lake2	**A9**	***Bacillus cereus* strain af-B25 (99)**
Tiberias Lake2	**A9**	***Bacillus cereus* strain UIS0932 (99)**
Tiberias Lake3	**A11**	***Bacillus sp*. (in: Bacteria) strain CM-CNRG 602 (99)**
Tiberias Lake3	**A11**	***Bacillus cereus strain JCM 2152* (98)**

Sea of Marmara	**D8**	***Bacillus sp*. strain DMAJP21 (99)**
Sea of Marmara	**D8**	**Lysinibacillus sp. strain NP05 (97)**

Battir 1	**C2**	***Lysinibacillus sphaericus* strain DH-B01 (99)**
Battir 1	**C2**	***Lysinibacillus fusiformis* strain P-R2A48 (99)**
Battir 2	**C3**	*Bacillus cereus strain EdyKolBc* **(99)**

Appendix 2

CLUSTAL O (1.2.4) multiple sequence alignment

https://www.ebi.ac.uk/Tools/msa/clustalo

A2.1: The 500bp sequences Isolate number as indicated in Figure 3.1 except No 14 isolate referred to *Pseudomonas aeruginosa* strain DSM 50071 as an outgroup (not related reference bacterium).

CLUSTAL O (1.2.4) multiple sequence alignment

```
No14>14-500bp-F4          CCGCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAATGAATTGACGGGGGCCCGCACAA 60
No9>17-500BP-F4_A03_001   ----------------------------------------------------------AA 2
No7>10-500BP-F4_B02_002   -----------------------------------------------------------G 1
No>3-500BP-F4_C01_003     ------------------------------------------TGTAGGGGGCCCGCACAA 18
No2>1-500BP-F4_A01_001    -------------------------------------------TGAGGGGGCCCGCACAA 17
No3>6-500BP-F4_F01_002    ----------------------------------------------------------AA 2
No13>2-500BP-F4_B01_002   --------------------------------------GGAATTGAGGGGGCCCGCACAA 22
No8>18-500BP-F4_B03_002   -----------------------------------------------------GCAaCAA 7
No4>16-500BP-F4_H02_004   -------------------------------------------------------CACAA 5
No6>14-500BP-F4_F02_002   ------------------------------------------------------GCACAA 6
No5>12-500BP-F4_D02_004   --------------------------------------------ACGGGGGCCCGCACAA 16
No6>15-500BP-F4_G02_003   ------------------------------------------------------------ 0
No11>9-500BP-F4_A02_001   ------------------------------------------------GGGCCCGCACAA 12
No10>7-500BP-F4_G01_003   -------------------------------------------------------CACAA 5
No5>13-500BP-F4_E02_001   ----------------------------------------------------------AA 2
No12>8-500BP-F4_H01_004   ----------------------------------------------------------AA 2
No1>4-500BP-F4_D01_004    ------------------------------------------------------------ 0
No1>5-500BP-F4_E01_001    -------------------------------------------------GGCCCGCACAA 11
No7>11-500BP-F4_C02_003   --------------------------------------------------------CACA 4

No14>14-500bp-F4          GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCTGGCCTTGACATG 120
No9>17-500BP-F4_A03_001   gCGGtGGAGCaTGTGGTTTAATTCGAAGCAACGC-GA-gACCTTACCAGGTCTTGACATC 60
No7>10-500BP-F4_B02_002   CCGGTGGAGCATGTGGTTTAATTCGAAGCAACGC-gAaGAACCTTACCAGGTCTTGAATC 60
No>3-500BP-F4_C01_003     gCGGTGGAGCATGTGGTTTAATTCGAAGCAACGC-GAAAACCTTACCAGGTCTTGACATC 77
No2>1-500BP-F4_A01_001    tCGGTGGAGCATGTGGTTTAATTCGAAGCAACGC-GAgAACCTTACCAGGTCTTGACATC 76
No3>6-500BP-F4_F01_002    GCGGaGGAGCATGTGGTTTAATTCG-AAGCACGC-GAgAACCTTACCAGGTCTTGACATC 60
No13>2-500BP-F4_B01_002   GCGGTGGAGCATGTGGTTTAATTCG-AgCAACGC-GAgAACCTTACCAGGTCTTGACATC 80
No8>18-500BP-F4_B03_002   GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATC 67
No4>16-500BP-F4_H02_004   TCGGT-GGAAATGTGGTTTAATT-CG-AGCGCGCGAACAACCTTACCGGTCT---TGAaT 59
No6>14-500BP-F4_F02_002   tCGGA-GGAAATGTGGTTTAtGTCGG-AGCGCGCGAACAACCTTACCAGGTCTTGACATC 64
No5>12-500BP-F4_D02_004   GCGGTGGAGCATGTGGTTTAATTCGAAGCAACG-CGAgAACCTTACCAGGTCTTGACATC 75
No6>15-500BP-F4_G02_003   -CGGTGGAGCATGTGGTTTAATTCGAAGCAACG-CGAgAACCTTACCAGGTCTTGACATC 58
No11>9-500BP-F4_A02_001   GCGGTGGAGCATGTGGTTTAATTCGAAGCAACG-CGAgAACCTTACCAGGTCTTGACATC 71
No10>7-500BP-F4_G01_003   tCGGTGGAGCATGTGGTTTAATTCGAAGCACG--CGAgAACCTTACCAGGTCTTGACATC 63
No5>13-500BP-F4_E02_001   GCGGTGGAGCATGTGGTTtAATTCGAAgCAACGCGAAGAACCTTACCAGGTCTTGACATC 62
No12>8-500BP-F4_H01_004   GCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATC 62
No1>4-500BP-F4_D01_004    --GGTGGAGCATGTGGTTTAATTCGAAGCAACGC-GAgAACCTTACCAGGTCTTGACATC 57
No1>5-500BP-F4_E01_001    tCGGTGGAGCATGTGGTTTAATTCGAAGCAACGC-GAgAACCTTACCAGGTCTTGACATC 70
No7>11-500BP-F4_C02_003   AtCGGTGGACATGTGGTTTAATTCGAAGCAACGC-GAAAACCTTACCAGGTCTTGACATC 63
                               *  *  ********** *        * * * * *       *

No14>14-500bp-F4          CTGAGAACTTTCCAGAGAT-GGATTGGTG-CCTTCGGGAACTCAGACACAGGTGCTGCAT 178
No9>17-500BP-F4_A03_001   CCGTTGACCACTGTAGAGATATAGTTTCCCCTTCGGGGCAACGG-TGACAGGTGGTGCAT119
No7>10-500BP-F4_B02_002   CCGTTGACCACTGTAGAGATATGGTTTCCCCTTCGGGGCAACGG-TGACAGGTGGTGCAT 119
```

```
No>3-500BP-F4_C01_003       CCGTTGACCACTGTAaAaATATGGtTTTCCCTTCGGGGACAACGGTGACAGGTGGTGCAT 137
No2>1-500BP-F4_A01_001      CCGTTGACCACTGTAGAGATATGGTTTTCCCTTCGGGGACAACGGTGACAGGTGGTGCAT 136
No3>6-500BP-F4_F01_002      CCGTTGACCACTGTAGAGATATAGTTTCCCCTTCGGGGGCAACGGTGACAGGTGGTGCAT 120
No13>2-500BP-F4_B01_002     CCGTTGACCACTGTAGAGATATAGTTTCCCCTTCGGGGGCAACGGTGACAGGTGGTGCAT 140
No8>18-500BP-F4_B03_002     CTCTGACAACTCTAGAGATAGAGCGTTCCCCTTCGGGGGACAGAGTGACAGGTGGTGCAT 127
No4>16-500BP-F4_H02_004     CCTCTGAGaCC-CTATgATAGGGCTTTCCCCTTCGGGGGACAGAGTGACAGGgGGTGCAT 118
No6>14-500BP-F4_F02_002     CTCTGACAACC-CTATaATAGGGCTTTCCCCTTCGGGGGACAGAGTGACAGGgGGTGCAT 123
No5>12-500BP-F4_D02_004     CTCTGACAATCCTAGAGATAGGACGTCC—CCTTCGGGGGCAGAGTGACAGGTGGTGCAT 133
No6>15-500BP-F4_G02_003     CTCTGACAACCCTAGAGATAGGGCTTTC—CCTTCGGGGACAGAGTGACAGGTGGTGCAT 116
No11>9-500BP-F4_A02_001     CTCTGACAATCCTAGAGATAGGACGTCC—CCTTCGGGGGCAGAGTGACAGGTGGTGCAT 129
No10>7-500BP-F4_G01_003     CTCTGCCAACCCTAGAGATAGGGCGTTCCCCTTCGGGGGACAGAGTGACAGGTGGTGCAT 123
No5>13-500BP-F4_E02_001     CTCTGAAAACCCTAGAGATAGGGCTTCT—CCTTCGGGAGCAGAGTGACAGGTGGTGCAT 120
No12>8-500BP-F4_H01_004     CTCTGAAAACCCTAGAGATAGGGCTTCT—CCTTCGGGAGCAGAGTGACAGGTGGTGCAT 120
No1>4-500BP-F4_D01_004      CTCTGAAAACCCTAGAGATAGGGCTTCT—CCTTCGGGAGCAGAGTGACAGGTGGTGCAT 115
No1>5-500BP-F4_E01_001      CTCTGAAAACCCTAGAGATAGGGCTTCT—CCTTCGGGAGCAGAGTGACAGGTGGTGCAT 128
No7>11-500BP-F4_C02_003     CTCTGAAAACCCTAGAGATAGGGCTTCT—CCTTCGGGAGCAGAGTGACAGGTGGTGCAT 121
                *                   * *  ***       ***** * *****

No14>14-500bp-F4            GGC-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGTAACGAGCGCAACCCT 237
No9>17-500BP-F4_A03_001     GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAcgGAGcGCaCCGCT 178
No7>10-500BP-F4_B02_002     GGTTGTCGTCAAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAaCGaGCGCAaCCCT 179
No>3-500BP-F4_C01_003       GGT-TGTCGtCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAaCGAGCGCAACCCT 196
No2>1-500BP-F4_A01_001      GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAaCGAGCGCAACCCT 195
No3>6-500BP-F4_F01_002      GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCT 179
No13>2-500BP-F4_B01_002     GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCT 199
No8>18-500BP-F4_B03_002     GGT-TGTCGTCAGCTCGTGTCGTGAgATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCT 186
No4>16-500BP-F4_H02_004     GGT-TGTCgtCAGCTCGTGTCGtGAGATGTTGGgTTAAGTCCCGCaaCGAgCGCAaCCCT 177
No6>14-500BP-F4_F02_002     GGT-TGTCgtCAGCTCGTGTCGtGAGATGTTGGgTTAAGTCCCGCAaCgAgCGCAaCCCT 182
No5>12-500BP-F4_D02_004     GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCT 192
No6>15-500BP-F4_G02_003     GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCT 175
No11>9-500BP-F4_A02_001     GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCT 188
No10>7-500BP-F4_G01_003     GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAaCGAGCGCAACCCT 182
No5>13-500BP-F4_E02_001     GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCT 179
No12>8-500BP-F4_H01_004     GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAaCCCT 179
No1>4-500BP-F4_D01_004      GGT-TGTCGTCAGCTCGTGTCGtGAGATGTTGGGTTAAGTCCCGCAaCGAGCGCAaCCCT 174
No1>5-500BP-F4_E01_001      GGT-TGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCT 187
No7>11-500BP-F4_C02_003     GGT-TGTCGTCAgCTCGTGTCGTGAGATGTTGGGTTAAgTCCCGCAaCGAGCGCAaCCCT 180
                **      ********************************* *  ******* * **

No14>14-500bp-F4            TGTCCTTAGTTACCAGCACCTCGGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGG 297
No9>17-500BP-F4_A03_001     TGATCTTAGTTGCCATCATTT-AGTTGgGCACTCTAAGGtGACTGCCGGTGACAAACCGt 237
No7>10-500BP-F4_B02_002     TGATCTTAGTTGCCATCATTT-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 238
No>3-500BP-F4_C01_003       TGATCTTAGTTGCCATCATTT-AgTTGGgCACTCTAAGGTGACTGCCGGTGACAAACCGG 255
No2>1-500BP-F4_A01_001      TGATCTTAGTTGCCATCATTT-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 254
No3>6-500BP-F4_F01_002      TGATCTTAGTTGCCATCATTT-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 238
No13>2-500BP-F4_B01_002     TGATCTTAGTTGCCATCATTT-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 258
No8>18-500BP-F4_B03_002     TGATCTTAGTTGCCAGCATTT-AgTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 245
No4>16-500BP-F4_H02_004     TGATCTTAGTTGCCcGCATTC-AaTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 236
No6>14-500BP-F4_F02_002     TGATCTTAGTTGCCagCATTC-AaTTGGgCACTCTAAGGtGACTGCCGGTGACaAACCGG 241
No5>12-500BP-F4_D02_004     TGATCTTAGTTGCCAGCATTC-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 251
No6>15-500BP-F4_G02_003     TGATCTTAGTTGCCAGCATTC-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 234
No11>9-500BP-F4_A02_001     TGATCTTAGTTGCCAGCATTC-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 247
No10>7-500BP-F4_G01_003     TGATCTTAGTTGCCAGCATTC-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 241
No5>13-500BP-F4_E02_001     TGATCTTAGTTGCCATCATTA-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 238
No12>8-500BP-F4_H01_004     TGATCTTAGTTGCCATCATTA-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 238
No1>4-500BP-F4_D01_004      TGATCTTAGTTGCCATCATTA-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 233
No1>5-500BP-F4_E01_001      TGATCTTAGTTGCCATCATTA-AGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG246
No7>11-500BP-F4_C02_003     TGATCTTAGTTGCCATCATTA-aGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGG 239
                **  ******* **  **      ************** *******************

No14>14-500bp-F4            AGGAAGGTG-GGGATGACGTCAAGTCATCATGGCCCTTACGGCCAGGGCTACACACGTGC 356
No9>17-500BP-F4_A03_001     AGGAAGGTGTGCGATGACGTcAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 297
No7>10-500BP-F4_B02_002     AGGAAGGtG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 297
No>3-500BP-F4_C01_003       AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 314
No2>1-500BP-F4_A01_001      aGGAAGGTG-GGGATGACGTCAAATCATCATGCCCcTTATGACCTGGGCTACACACGTGC 313
No3>6-500BP-F4_F01_002      AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 297
No13>2-500BP-F4_B01_002     AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 317
No8>18-500BP-F4_B03_002     AGGAAGGtG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 304
No4>16-500BP-F4_H02_004     AGGAAGGtG-GGGATGACgTCAAATCATCATGCCCCTTATGACCTGGGCTACACaCGTGC 295
No6>14-500BP-F4_F02_002     AGGAAgGtG-GGGATGACgTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC|300
```

```
No5>12-500BP-F4_D02_004     AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 310
No6>15-500BP-F4_G02_003     AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 293
No11>9-500BP-F4_A02_001     AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 306
No10>7-500BP-F4_G01_003     AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 300
No5>13-500BP-F4_E02_001     AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 297
No12>8-500BP-F4_H01_004     AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 297
No1>4-500BP-F4_D01_004      AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 292
No1>5-500BP-F4_E01_001      AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 305
No7>11-500BP-F4_C02_003     AGGAAGGTG-GGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGC 298
                  ********* * *********** ******** ****** * ** ***************

No14>14-500bp-F4         TACAATGGTCGGTACAAAGGGTTGC-CAAGCCGCGAGGT---GGAGCTAATCCCATAAAA 412
No9>17-500BP-F4_A03_001     TACAATGGACGATACAaACGGTTGCCAACTCGcGAGAG----ggAGCTAATCCGATAAAG 353
No7>10-500BP-F4_B02_002     TACAATGGACGATACAAACGGTTGCcAACTCGCGAGAG----GGAGCTAATCCGATAAAG 353
No>3-500BP-F4_C01_003     TACAATGGACGATACAAACGGTTGCCAAcTCGCGAGAG----GGAGCTAATCCgATAAAG 370
No2>1-500BP-F4_A01_001     TACAATGGACGATACAAACGGTTGCCAACTCGCGAGAG----GGAGCTAATCCGATaAAG 369
No3>6-500BP-F4_F01_002     TACAATGGACGATACAAACGGTTGCCAACTCGCGAGAG----GGAGCTAATCCGATAAAG 353
No13>2-500BP-F4_B01_002     TACAATGGACGATACAAACGGTTGCCAACTCGCGAGAG----GGAGCTAATCCgATAAAG 373
No8>18-500BP-F4_B03_002     TACAATGGATGGTACAAAGGGGCTGCaAGACCGcGGaGGtCAAGCCaAATCCCATAAAAA 364
No4>16-500BP-F4_H02_004     TACAaTGGATGGtACAaAGGGCTGCAAACCTGCGA-AGGT---AAgCGAATCCCATAAAG 351
No6>14-500BP-F4_F02_002     TACAaTGGATGGTACAaAGGGCTGCAAACCTGCGA-AGGT---AAgCGAATCCCATAAAG 356
No5>12-500BP-F4_D02_004     TACAATGGACAGAACAAAGGGCAGCGAAACCGCGA-GGTT---AAGCCAATCCCACAAAT 366
No6>15-500BP-F4_G02_003     TACAATGGACAGAACAAAGGGCTGCgAGACCGCaa-GGTT---TAGCCAATCCCATAAAT 349
No11>9-500BP-F4_A02_001     TACAATGGACAGAACAAAGGGCAGCGAAACCGCGA-GGTT---AAGCCAATCCCACAAAT 362
No10>7-500BP-F4_G01_003     TACAATGGATGGTACAAAGAGCTGCGAACCCGCGA-GGGT---AAGCGAATCTCATAAAG 356
No5>13-500BP-F4_E02_001     TACAATGGACGGTACAAAGAGCTGCAAgACCGCGA-GGTG---GAGCTAATCTCATAAAA 353
No12>8-500BP-F4_H01_004     TACAATGGACGGTACAAAGAGCTGCAAGACCGCGA-GGTG---GAGCTAATCTCATAAAA 353
No1>4-500BP-F4_D01_004     TACAATGGACGGTACAAAGAGCTGCAAGACCGCgA-GGTGG—gAGCTAATCTCATAAAA|349
No1>5-500BP-F4_E01_001     TACAATGGACGGTACAAAGAGCTGCAAGACCGCGA-GGTG---GAGCTAATCTCATAAAA 361
No7>11-500BP-F4_C02_003     TACAATGGACGGTACAAAGAGCTGCgAGACCGCGA-GGTG---GAGCTAATCTCATAAAA 354
                  ********    ***** *           *       * *   ***

No14>14-500bp-F4         CCGATCGTAGTCCGGATCGCAGTCTGCAACTCGACTGCGTGAAGTCGGAATCGCTAGTAA 472
No9>17-500BP-F4_A03_001     TCGTTCTCAGTTCGAATTGTaGGCTGCAaCTCGCCTACtTGACGCCGGA-TCGCTAGTAa 412
No7>10-500BP-F4_B02_002     TCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCCGGGATCGCTAGTAA 413
No>3-500BP-F4_C01_003     TCgTTCTCAgTTCGgATTGTaGGCTGCAaCTCGCCTACaTGAAgCCGGAATCGCTAGTAa 430
No2>1-500BP-F4_A01_001     TCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAcGCCGGAATCGCTAGTAA 429
No3>6-500BP-F4_F01_002     TCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAcGCCGGAATCGCTAGTAA 413
No13>2-500BP-F4_B01_002     TCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAaGCCGGAATCGCTAGTAa 433
No8>18-500BP-F4_B03_002     CCATTCTCAGTTCGGATTGtAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAA 424
No4>16-500BP-F4_H02_004     CCATTCTCAGTTCGGATTGCAcGCTGCaAcTCGCCTGCaTGAAGCCGGAcTCGCTAGTAa 411
No6>14-500BP-F4_F02_002     CCATTCTCAgTTCGGATTGCAGGCTGCAACTCgCCTGCATGAAgCCGGAcTCGCTAGTAA 416
No5>12-500BP-F4_D02_004     CTGTTCTCAGTTCGGATCGCAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGCTAGTAA 426
No6>15-500BP-F4_G02_003     CTGTTCTCAGTTCGGATCGCAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGCTAGTAA 409
No11>9-500BP-F4_A02_001     CTGTTCTCAGTTCGGATCGCAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGCTAGTAA 422
No10>7-500BP-F4_G01_003     CCATTCTCAGTTCGGATTGTAGGCTGCAaCTCGCCTACATGAAGCCGGAATCGCTAGTAa 416
No5>13-500BP-F4_E02_001     CCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAA 413
No12>8-500BP-F4_H01_004     CCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAA 413
No1>4-500BP-F4_D01_004     CCGTTCTCAGTTCGGATTGTAGGCtGCAACTCGCCTACaTGAAGCTGGAATCGCTAGTAA 409
No1>5-500BP-F4_E01_001     CCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGCTGGAATCGCTAGTAA 421
No7>11-500BP-F4_C02_003     CCGTTCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACaTGAAGCTGGAATCGCTAGTAA 414
                     **  *** ** ** * *  ********** ** * *** *  **  **********

No14>14-500bp-F4         TCGTGAATCAGAATGTCACGGTGAAT-ACGTTCCC-GGGCC-TT-GTA-CACACCG-CCC      526
No9>17-500BP-F4_A03_001     TCGCGGAtCAaCATGCCGCGTgA--ATACGTcCCCGGCCTT-------------------      451
No7>10-500BP-F4_B02_002     TCGCGGATCAgCatGCCGc-GGTGAATACGTTCCCGGcCCTTGtA-CACA----------      461
No>3-500BP-F4_C01_003     TCGCGGATCAgCATGCCG-CGGTGAATACgTTCCCGGGCCTTGTA-CACACCGCCCGTC- 487
No2>1-500BP-F4_A01_001     TCGCGGATCAgCATGCCCGCGGTGAATACgTTCCCGGGCCTTGTA-CACACCGCC----- 483
No3>6-500BP-F4_F01_002     TCGCGGATCAgCATGCCGc-GGTGAATACgTTCCCGGGCCTTGTA-CACACCGC------  465
No13>2-500BP-F4_B01_002     TCGCGGATCAgCATGCCCGCGGTGAATACgTTCCCGGGCCTTGTA-CACACCGCCCGTCA 492
No8>18-500BP-F4_B03_002     TCGCGGATCCcCCaTGCCGcGGTGAATACgTTCCCGGGcCCTTGTACACACCGCCCgGcC 484
No4>16-500BP-F4_H02_004     TCGCGGATCAaCaTG-CCGCGGTGAATA------------------------------          438
No6>14-500BP-F4_F02_002     TCGCGGATCAgCaTG-CCGCGGTGAATACGTTCCCGGGCCTTGTACACACCCGCCCCGTC 475
No5>12-500BP-F4_D02_004     TCGCGGATCACCctG-CCGCGGTGAATACtTTCCCGGGGCCTTGTACAC-----------      474
No6>15-500BP-F4_G02_003     TCGCGGATCAGCAtG-CCGCGGtGAATACGTTCCCGGGCCTTGTACACACCGC-------  461
No11>9-500BP-F4_A02_001     TCGCGGATCAgCAtG|-CCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCG----477
No10>7-500BP-F4_G01_003     TCGCGGATCAGCATg-CCGCGGTGAATACGTTCCCGGG-CCTTGTACAC-----------      463
No5>13-500BP-F4_E02_001     TCGCGGATCAcCatG-CCCGCGGTGAATACGTTCCCGGGCCtTTGTACACACCC------    466
No12>8-500BP-F4_H01_004     TCGCGGATCAgCATG-CCCGCGGTGAATACgTTCCCGGGCCTtGTAacACACCGCC----- 468
No1>4-500BP-F4_D01_004     TCGCGGATCAgCATG-CCGCGGTGAAT-ACgTTCCCGGGCCTTGTACACACCCGCCCG--465
```

```
No1>5-500BP-F4_E01_001      TCGCGGATCAgCATG-CCGCGGTGAAT-ACGTTCCCGGGCCTTGTACACACCGCCCGTCA 479
No7>11-500BP-F4_C02_003      TCGCGGATCAgCATG-CCGCGGTGAAT-ACgTTCCCGGtCCTTGTACACACCGCCcGTC| 471
                   *** * ***      *

No14>14-500bp-F4          GTC         529
No9>17-500BP-F4_A03_001    ---         451
No7>10-500BP-F4_B02_002    ---         461
No>3-500BP-F4_C01_003     ---         487
No2>1-500BP-F4_A01_001    ---         483
No3>6-500BP-F4_F01_002    ---         465
No13>2-500BP-F4_B01_002    A--         493
No8>18-500BP-F4_B03_002    ---         484
No4>16-500BP-F4_H02_004    ---         438
No6>14-500BP-F4_F02_002    A--         476
No5>12-500BP-F4_D02_004    ---         474
No6>15-500BP-F4_G02_003    ---         461
No11>9-500BP-F4_A02_001    ---         477
No10>7-500BP-F4_G01_003    ---         463
No5>13-500BP-F4_E02_001    ---         466
No12>8-500BP-F4_H01_004    ---         468
No1>4-500BP-F4_D01_004    ---         465
No1>5-500BP-F4_E01_001    A--         480
No7>11-500BP-F4_C02_003    ---         471
```

CLUSTAL O (1.2.4) multiple sequence alignment

https://www.ebi.ac.uk/Tools/msa/clustalo

A2.2: The 600bp sequences isolate numbered as in Figure 3.1 except No 14 isolate referred to *Pseudomonas kribbensis* strain 46-2 as an outgroup (not related reference bacterium).

CLUSTAL O (1.2.4) multiple sequence alignment

```
No10>5-600BP-FN3_E04_001    --------GGTTTCCGCCCTTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGA      52
No14>14-600BP-Fn3           -------------------TTAGTGGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGA     40
No8>13-600BP-FN3_D05_004    -----GGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAaGCACTCCGCCTGGGGA      55
No7>8-600BP-FN3_H04_004     -GTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGA      59
No13>1-600BP-FN3_A04_001    --TTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGA      58
No3>4-600BP-FN3_D04_004     TGTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGA      60
No9>12-600BP-FN3_C05_003    ------GGGGGTTTCCGCCCTTAGTGCTGCAGCtAACGCATTAAGCACTCCGCCTGGGGA      54
No6>11-600BP-FN3_B05_002    -GTTAGAGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCGCCTGGGGA      59
No11>7-600BP-FN3_G04_003    -GTTAGGGGGTTTCCGCCCCTTAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGA      59
No12>6-600BP-FN3_F04_002    --------GGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCGCCTGGGGA      52
No1>2-600BP-FN3_B04_002     -GTTAGcGGGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCGCCTGGGGA      59
No1>3-600BP-FN3_C04_003     --------GGTTTCCGCCCTTTAGTGCTGAAGTTAACGCATTAAGCACTCCGCCTGGGGA      52
                       ****** * ** ************    ***********

No10>5-600BP-FN3_E04_001    GTACGGCCGCcG—GCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAgCGGTGGAGC 110
No14>14-600BP-Fn3           GTACGGCCGCAAG-GTTAAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGC       99
No8>13-600BP-FN3_D05_004    GTACGGtCGCAAGTACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAaGCGGTGGAGC 115
No7>8-600BP-FN3_H04_004     GTACGGTCGCAAG-ACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGC118
No13>1-600BP-FN3_A04_001    GTACGGTCGCAAG-ACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGC117
No3>4-600BP-FN3_D04_004     GTACGGTCGCAAG-ACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGC119
No9>12-600BP-FN3_C05_003    GTACGGTCGCAAG-ACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGC113
No6>11-600BP-FN3_B05_002    GTACGGCCGCAAG-GCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGC118
No11>7-600BP-FN3_G04_003    GTACGGTCGCAAG-ACTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGC118
No12>6-600BP-FN3_F04_002    GTACGGCCGCAAG-GCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGC111
No1>2-600BP-FN3_B04_002     GTACGGCCGCAAG-GCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGtGGAGC 118
No1>3-600BP-FN3_C04_003     GTACGGCCGCAAG-GCTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGC|111
                 ****** ***     * ********* ********************************

No10>5-600BP-FN3_E04_001    ATGTGGTTTAATTCGAAGCAaCgCGAAcAACCTTACCAGGTcTTGACATCCTCTGCCAAC 170
No14>14-600BP-Fn3           ATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGCCTTGACATCCAATGAACTT 159
No8>13-600BP-FN3_D05_004    ATGTGGTTTAATTCGAAGCAaCGCGAAgAACCTTACCAGGTCTTGACATCCCGTTGACCA 175
No7>8-600BP-FN3_H04_004     ATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCCGTTGACCA 178
No13>1-600BP-FN3_A04_001    ATGTGGTTTAATTCGAAGCAACgCGAAgAACCTTACCAGGTCTTGACATCCCGTTGACCA 177
No3>4-600BP-FN3_D04_004     ATGTGGTTTAATTCGAAGCAaCgCGAAcAACCTTACCAGGTCTTGACATCCCGTTGACCA 179
No9>12-600BP-FN3_C05_003    ATGTGGTTTAATTCGAAGCAACGcGAAGAACCTTACCAGGTCTTGACATCCTCTGAC-AA 172
No6>11-600BP-FN3_B05_002    ATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCTCTGAA-AA 177
No11>7-600BP-FN3_G04_003    ATGTGGTTTAATTCGAAGCAACGCGAAgAACCTTACCAGGTCTTGACATCCTCTGAC-AA 177
No12>6-600BP-FN3_F04_002    ATGTGGTTTAATTCGAAGCAaCGCGAAGAACCTTACCAGGTCTTGACATCCTCTGAA-AA 170
No1>2-600BP-FN3_B04_002     ATGTGGTTTAATTCGAAGCAaCgCGAAGAACCTTACCAGGTCTTGACATCCTCTGAA-AA 177
No1>3-600BP-FN3_C04_003     ATGTGGTTTAATTCGAAGCAaCgCGAAGAACCTTACCAGGTCTTGACATCCTCTGAA-AA 170
                 ************************* ************ **********  *

No10>5-600BP-FN3_E04_001    CCTAGAGATAGGGCGTTCCCcTTCGGGGGACAaAGTGACAGGtGGgGCATGGKTGWMSTM230
No14>14-600BP-Fn3           TCCAGAG-AT-GGATTGGTGCCTTCGGGAACATTGAGACAGGTGCTGCATGGCTGTCGTC 217
No8>13-600BP-FN3_D05_004    CTGTaGAGATATAGTTTCCCCTTCGGGGgCAACGGTGACAGGTGtaGCATGgtTGTCGTC235
No7>8-600BP-FN3_H04_004     CTGTAGAGATATGGTTTCCCCTTCGGGGGCAACGGTGACAGGTGGTGCATGGTTGTCGTC 238
No13>1-600BP-FN3_A04_001    CTGTAGAGATATAGTTTCCCCTTCGGGGGCAACGGTGACAGGTGGTGCATGGTTGTCGTC 237
No3>4-600BP-FN3_D04_004     CTGTAGAGATATAGTTTCCCCTTCGGGGGCAACGGtGACAGGTGGtGCATGGtTGTCGTC 239
No9>12-600BP-FN3_C05_003    CCCTAGAGATAGGgCTTTTCCCTTCGGGGACAgAGTGACAGGTGGTGCATGGTTGTCGTC 232
No6>11-600BP-FN3_B05_002    CCCTAGAGATAGGGCTTCT-CCTTCGGGAGCAGAGTGACAGGTGGTGCATGGTTGTCGTC 236
No11>7-600BP-FN3_G04_003    TCCTAGAGATAGGACGTCC-CCTTCGGGGGCAGAGTGACAGGTGGTGCATGGTTGTCGTC 236
No12>6-600BP-FN3_F04_002    CCCTAGAGATAGGGCTTCT-CCTTCGGGAGCAGAGTGACAGGTGGTGCATGGTTGTCGTC 229
No1>2-600BP-FN3_B04_002     CCCTAGAGATAGGGCTTCT-CCTTCGGGAGCAGAGTGACAGGTGGTGCATGGTTGTCGTC 236
No1>3-600BP-FN3_C04_003     CCCTAGAGATAGGGCTTCT-CCTTCGGGAGCAgAgTGACAGGTGGTGCATGGtTGTCGtC 229
                         * *  ***   *  * * ********  ****** **   *
```

```
No10>5-600BP-FN3_E04_001     AGATCGtGtcgTGAGATGTTGGgTTAAgTCCcgCAACGAGCGCAaCCCTTGATCTTAgTT 290
No14>14-600BP-Fn3         AGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGTAACGAGCGCAACCCTTGTCCTTAGTT 277
No8>13-600BP-FN3_D05_004     AGCTCGTGTCGTGAGATGTTGGgTTAAgTCCCGCAACGAGCGCAaCACTTGATCTTAGTT 295
No7>8-600BP-FN3_H04_004     AGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTT 298
No13>1-600BP-FN3_A04_001     AGcTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTT 297
No3>4-600BP-FN3_D04_004     AGcTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTT 299
No9>12-600BP-FN3_C05_003     AGCTcGTGTCGTGAGATGTTGGgTTAAGTCCCGCAaCGAGCGCAACCCTTGATCTTAGTT 292
No6>11-600BP-FN3_B05_002     AGCTCGTGTCGTGAGATGTTGGgTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTT 296
No11>7-600BP-FN3_G04_003     AGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTT 296
No12>6-600BP-FN3_F04_002     AGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAaCGAGCGCAaCCCTTGATCTTAGTT 289
No1>2-600BP-FN3_B04_002     AGCTcGTGTCGTGAGATGTTGGGTTAAGTCCCGcAaCgAGCGCAACCCTTGATCTTAGTT 296
No1>3-600BP-FN3_C04_003     AGCTcGTGTCGtGAGATGTTGGGTTAAGTCCCgcAaCgAGCGCAACCCTTGATCTTAGTT 289
                  ** ****************************** ************ ****  *******

No10>5-600BP-FN3_E04_001     GCCAgctT—TCAcTTGGGCACTtTAAGGtGACTGCCGGTGACAAACCGGAGGAAGGTGG 348
No14>14-600BP-Fn3         ACCAGCACGTTATGGTGGGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGGAAGGTGG 337
No8>13-600BP-FN3_D05_004     GCCATCaT—TTAGTTGgGCACTCTAaGGTGACTGCCGGTGACAAACCGGAGGAAGGTGt 353
No7>8-600BP-FN3_H04_004     GCCATCAT—TTAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGG 356
No13>1-600BP-FN3_A04_001     GCCATCAT—TTAGTTGGGCACTCTAAGGTGACTGCcGGTGACAAACCGGAGGAAGGtGG 355
No3>4-600BP-FN3_D04_004     GCCATCAT—TTAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGtGG 357
No9>12-600BP-FN3_C05_003     GCCAGCAT--TCAGTTGGGCACTCTAAGGTGACTGCCGGTGACAaACcGGAGGAGGGTG- 349
No6>11-600BP-FN3_B05_002     GCCATCAT—TAAGTTGGGCACTCTAAGGTGACTGCCGGTGACAaACCGGAGGAAGGTGg 354
No11>7-600BP-FN3_G04_003     GCCAGCAT—TCAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGG 354
No12>6-600BP-FN3_F04_002     GCCATCAT—TAAGTTGGcCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGG 347
No1>2-600BP-FN3_B04_002     GCCATCAT—TAAGTTGGGCACTCTAAGGTGACTGCCGGTGACAAACCGGAGGAAGGTGG 354
No1>3-600BP-FN3_C04_003     GCCATCAT—TAAGTTGGgCACTCTAAGGTGACTGCCGGtGACAAACCGGAGGAAGGTGG 347
                   *** *   *    *** **** ***** ************************* ****

No10>5-600BP-FN3_E04_001     GGATGACGTCAAATCATCATGCCCgTTATGACCTGGgCTACACACgTGCTACAATGGATG 408
No14>14-600BP-Fn3         GGATGACGTCAAGTCATCATGGCCCTTACGGCCTGGGCTACACACGTGCTACAATGGTCG 397
No8>13-600BP-FN3_D05_004     ggATGACGTCaaATCATCATGCCCSTTAYGACCTGTTCTAcACACGTGCTACAcTGGACg 413
No7>8-600BP-FN3_H04_004     GGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGACG 416
No13>1-600BP-FN3_A04_001     GgATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACgTGCTACAATGGACG 415
No3>4-600BP-FN3_D04_004     GgATGACgTCAAATCATCATGCCCcTTATGACCTGGGCTACACACgTGCTACAATGGACG 417
No9>12-600BP-FN3_C05_003     GGATGACGTCAAATCATCATGCCCCTTATGACCTGg-CTACACACGTGCTACAATGGACA 408
No6>11-600BP-FN3_B05_002     gGATGACGTCAAATCATCATGCCCCTTATGACCTGG-CTACACACGTGCTACAATGGACG 413
No11>7-600BP-FN3_G04_003     GGATGACGTCAAATCATCATGCCCCTTATGACCtGGGCTACACACGTGCTACAATGGACA 414
No12>6-600BP-FN3_F04_002     GGATgACGTCAAATCATCATGCCCcTTATGACCTGGGCTACACACGTgCTACAATGGACG 407
No1>2-600BP-FN3_B04_002     GGATGACaTCAAATCATCATGCCCcTTATGACCTGGGCTACACACGTGCTACAATGGACG 414
No1>3-600BP-FN3_C04_003     GGATGACaTCAAATcATCATGCCccTTATGACCTGtGCTACACACGTGCTACAcTGGACG 407
                  ******* **** ******** ** *** * ****  **************** ***

No10>5-600BP-FN3_E04_001     GaACAAAGAGCTGCGaACC-CGCGAGGGTAAGCGAATCTCATAAAGCCATTCTCAGTTCG 467
No14>14-600BP-Fn3         GTACAAAGGGTTGCCAAGC-CGCGAGGTGGAGCTAATCCCATAAAACCGATCGTAGTCCG 456
No8>13-600BP-FN3_D05_004     ATACAcACGGcTGCCA-aCTCGCGAGAgCcAgCTAATCCGATAAAGtCGTTCTCAgTTCG 472
No7>8-600BP-FN3_H04_004     ATACAAACGGTTGCCAACCTCGCGAGAGGGAgCTAATCCGATAAAGTCGTTCTCAGTTCG 476
No13>1-600BP-FN3_A04_001     ATACAAACGGTTGCCa-ACTCGCGAGAGGGAGCTAATCCGATAAAGTCGTTCTCAGTTCG 474
No3>4-600BP-FN3_D04_004     ATaCAAACGGtTGCCA-ACTCGCGAGAGGGAGCTAATCCGATAAAGTCGTTCTCAgTTCG 476
No9>12-600BP-FN3_C05_003     GAACAGAGGGCTGCaAgACCGCc-GGTTc-AgCCAATCcCATACATCTGTTCTCAGTTCG 466
No6>11-600BP-FN3_B05_002     GtACAAAGAGCTGCAagCCCCGCGAGGTG-CAtCTAATCTCATAaAcCGTTCTCAGTTCG 472
No11>7-600BP-FN3_G04_003     GAACAAAGGGCAGCGaAaCCGCGAGGTTA-AGCCAATCCCACAAATCTGTTCTCAGTTCG 473
No12>6-600BP-FN3_F04_002     GtACAAAGAGCTGCAagsCCGCGAGGTGG-AtCTAATCTCATAAAACCGTTCTCAGTTCG 466
No1>2-600BP-FN3_B04_002     GtaCAAAGAGCTGCAacACCGCGAGGTGG-AGCTAATCTCATAAAACCGTTCTCAgTTCG 473
No1>3-600BP-FN3_C04_003     GtaCAAAGAGCTGCAagACCGCgAGgTGg-AgCTAATCTCATAAAACCGTTCTCAgTTCG 466
                   *** * * ** * *    *      *     *    ** *** **

No10>5-600BP-FN3_E04_001     GATTGtaGGCTGCAACTCGCCTACATGAAGCCGGTAATCGCTAGTAATCGCtGtATCAGC 527
No14>14-600BP-Fn3         GATCGCAGTCTGCAACTCGACTGCGTGAAGTCGGAATCGCT-AGTAATCGCGAATCAGAA 515
No8>13-600BP-FN3_D05_004     GAtTGtagGCTGcTACTCGCCTACATGAAGCCGGAAtCGCT-AGtAATCGCGtATCAGcA 531
No7>8-600BP-FN3_H04_004     GATTGTAGGCTGCAACTCGCCTACATGAAGCCGGAT-CGCT-AGTAATCGCGGATCAGCA 534
No13>1-600BP-FN3_A04_001     GATTGTAGGCTGCAACTCGCCTACATGAAGCCGGAATCGCT-AGtAATCGCGtATCAGCA 533
No3>4-600BP-FN3_D04_004     GATTGTaGGCTGcAACTCGCCTACATGAAGCCGgAATCGCT-AGTAATCGCGtATCAcCA 535
No9>12-600BP-FN3_C05_003     GATCGCAGTCTGCaACTCGACTGCGTGAAGCTGTAT-CGCT-AGTAATCGCGA-TCAGCA 523
No6>11-600BP-FN3_B05_002     GATTGTAGGCTGCaACTCGCCTACATGAAGCTG-GATCGCT-AGTAATCSCGGATCAGCA 530
No11>7-600BP-FN3_G04_003     GATCGCAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGCT-AGTAATCGCGGATCAGCA 532
No12>6-600BP-FN3_F04_002     GATTGTaGGCTGCAACTCGCCTACATGAAGCTGGAATCGCT-AGTAATCGCGGATCAGCA 525
No1>2-600BP-FN3_B04_002     GATTGTaGGCTGCAACTcGCCTACATGAAGCTGGAATCGCT-AGTAATCGCGGATCAgCA 532
No1>3-600BP-FN3_C04_003     GATTGTAGGCTGCAACTCGCCTACATGAAGCTGtAATcGCT-AGTAATCGCGtATCActa 525
                  *** * ** **** ***** ** * *****  *       ******* *
```

```
No10>5-600BP-FN3_E04_001    TTGCCaGCTGgTGAaATACAcTTCgCGGGaCCatGtgCACACtCgcCCGTCaG--          580
No14>14-600BP-Fn3         TGTCGCGGTGAATA-----CGTTCCCGGGCCTTGTA-CACACC-GCCCGTC---- 559
No8>13-600BP-FN3_D05_004    TGCCGCGaTcGAaT-----ACGTTCGCGG-------------------------    555
No7>8-600BP-FN3_H04_004     TGCCGCGTGAAT-A-----CGTTCCCGGGCCTTGTACACACCG-CCCCGT-----       577
No13>1-600BP-FN3_A04_001    TGCCGCGGTGAATA-----CGTTCCCGGGCCTTGTACACACCG-CCCGT------       576
No3>4-600BP-FN3_D04_004     TGCCGYGGTGAATA-----CGTTCCCGGGCcCTTGTaCACAC-------------  572
No9>12-600BP-FN3_C05_003    TGCCGCGTGAATAC-----GTTCC------------------------------  542
No6>11-600BP-FN3_B05_002    TGCCGYGTGAATTA-----CGTTCCSCGGGCCTTGTACACACC-SGCCCGTCGAG     579
No11>7-600BP-FN3_G04_003    TGCCGCGGTGAATA-----CGTTCCCGGGCCCTTGTACACACC-GCCCGTCAG--     579
No12>6-600BP-FN3_F04_002    TGCCGCGGTGAATA-----CGTTCCCGGGCCTTGTACACACCG-CCCCGTCA---     571
No1>2-600BP-FN3_B04_002     TGCCGcGGTGAATA-----CGTtCCCGGGCCCTTGTaCACACC-GCCC-------     574
No1>3-600BP-FN3_C04_003     TGCCGtGaTGAATA-----CtTcCCCGGGaCC----------------------    552
                  * * *
```